PEER REVIEW HISTORY

BMJ Open publishes all reviews undertaken for accepted manuscripts. Reviewers are asked to complete a checklist review form (see an example) and are provided with free text boxes to elaborate on their assessment. These free text comments are reproduced below. Some articles will have been accepted based in part or entirely on reviews undertaken for other BMJ Group journals. These will be reproduced where possible.

ARTICLE DETAILS

TITLE (PROVISIONAL)	Prediction of initiation and cessation of breast feeding from late pregnancy to 16 weeks: The Feeding Your Baby (FYB) cohort study
AUTHORS	Donnan, Peter; Dalzell, Janet; Symon, Andrew; Rauchhaus, Petra; Monteith-Hodge, Ewa; Kellett, Gillian; Wyatt, Jeremy; Whitford, Heather

VERSION 1 - REVIEW

REVIEWER	Dr Vivien Swanson, Senior Lecturer Psychology Division, School of Natural Sciences, University of Stirling, Stirling K9 4LA UK I have research interests in the psychology of breastfeeding.
REVIEW RETURNED	04-Jun-2013

THE STUDY	Exclusion criteria are not clearly described. Also reasons for exclusion of some who returned postcards not explained.
GENERAL COMMENTS	This is an interesting paper identifying demographic and psychological antenatal predictors of breastfeeding duration and cessation. The findings that socio-economic status, parity, attitudes and intentions predict women's decision to initiate breastfeeding is not surprising, and indeed has been found in several such studies. However this study is novel in that it uses this information to identify an algorithm which allows health professionals to predict whether or not individual women will breastfeed, and gives us some clues about continuation. It is a weakness of this study that only antenatal data (attitudes etc) is available to predict cessation - but I assume that the proposed follow-up RCT will account for postnatal changes in psychological variables due to women's experience of breastfeeding. The study has a reasonable sample size, and socio-demographic spread, which makes the findings generalisable. Although the authors suggest that ethnicity may be a factor they have missed, this study does replicate the findings of many others, suggesting this may not be a major weakness. The study is also novel because it uses SMS text messaging to collect longitudinal data in an unobtrusive manner which appears acceptable to participants. I was not altogether clear as to how the reliability of this method was checked, as the counter-checks used by the researchers also involved self-report. Nevertheless the researchers claim that their 'real time' assessment may be more accurate than other methods of self-report data collection. The study is theoretically underpinned using a Theory of Planned

	Behaviour (TPB) approach. Again this has been successfully used in several studies of breastfeeding so is an appropriate and useful model. This is clearly a strength of the study. I did have some queries in respect of how the TPB data was used in the study. The description of the measurement of TPB variables in the methods section is sketchy. How were TPB attitudes measured, and was there overlap between these and the IIFAS? It is surprising that TPB attitudes are not predictive of behaviour, as this has frequently been the case in previous research - although this may be a statistical artifact of inclusion of the IIFAS. It was also surprising that self-efficacy or perceived control wasn't predictive of breastfeeding, as has been the case in many studies - although this may be because this variable was only measured antenatally. Additionally, since intention is a major predictor, it is important to provide more detail about how this was measured. The point made in the introduction and discussion about the failure in current policy to discuss women's intention is well-made, since this leads health professionals to miss a key opportunity for behavioural intervention by discussing their beliefs and attitudes towards different feeding methods. The rationale for this has related to the importance of promoting a culture of breastfeeding as 'normal', so not raising the possibility of formula feeding - but in reality may have left some women feeling their views and concerns about different feeding methods have not been acknowledged or addressed. There are some very minor points below which it would be helpful for the authors to address, but in general I thought this was an interesting piece of research which merits publication. Minor issues: 1. P3, Lines 26,27 the meaning of this isn't clear ('points to targeting...') 2. P4, Lines 36,37. It would be useful to clarify the impact of the BFI on both initiation and continuation of breastfeeding rather than just talking about 'breastfeeding rates'. 3. P5, Line30 As discussed above - clarify how TPB variables were measured. 4. P5, Line 31 - In what way were the questions informed by self-efficacy theory? 5. P5, Line 12. Are there any exclusions? 6. P6, line 16 onwards - I dont this the information about recruitment to focus groups needs to be included here - and the topic guide isn't included in Appendix 1! 7. P6, Line 35, How does repeat texting ensure reliability? 8. P7, line 9 onwards. What was the basis for selection of participants from the 639 postcards returned? Was there any drop-out from the texting process? 9. P7. Why was the duration of follow-up 16 weeks, and why was the 'end point' assessment 4 weeks after that? 10. P7. Again, exclude description of procedure for focus groups. 11. P7 There is no heading for the start of the 'results' section. 12. P8, line 7, What is 'major' occupational group? and how is occupation coded - there is no information on this. 13. P8, line 41. The authors could acknowledge potential for a 'hawthorne' effect which may have influenced breastfeeding rates, or at least the reporting of these.

	14. P9, line 22. Explain how the study findings can be used 'to identify women who need additional support' - not sure how they will do this from the existing data. 15. P9, line 40 This study doesnt tell us why women stopped breastfeeding, so there will indeed be a need to explore these factors in future studies. 16. I was slightly confused by the breastfeeding time series data. The scale on Figures 3a and 3b suggests that some women stopped breastfeeding long after the 14 week + 4 weeks end points of the study (max 126 days). Is this data extrapolated in some way? 17. P17, line 16 onwards. I wasn't clear about the difference between the categorical and continuous measures of intention. 18. P20, Table 4a. I didnt understand what 'Total IIFAS score + 10 units' meant?

REVIEWER	Kingsley. E. Agho, PhD, MPH Research Fellow in Epidemiology and Biostatistics School of Medicine University of Western Sydney Australia
REVIEW RETURNED	10-Jun-2013

THE STUDY	If the study is a prospective cohort, the authors should report relative risk and not odds ratio – is there any reason for reporting Odd ratios instead of relative risk. In table 3 what is the rational for choosing the following weeks – 6weeks, 8 weeks and 16 weeks ?. The WHO recommended further disaggregate for reporting EBF as: 0-1 months; 2-3 months , 0-3 months and in this case, 0-4 months. In table 1, mean should be in one decimal place (1 dp) In table 2 remove x2 test and Parous - any bF vs primiparous can't be significant because the confidence interval (CI) around the OR includes 1.0. I think it was a typo error. In table 2, i would like to see both adjusted and unadjusted Odd ratios reported and the OR and the CI around the OR should be in 1 dp. Combine 4a & b Remove Appendix 1 – it doesn't add anything to the paper – it makes the paper look ugly. The plots in figs 3a & b indicated that there are statistical differences between Parous-any BF, Parous – No BF and Primiparous but the hazard ratios reported in tables 4a & b revealed that Parous-any BF, Parous – No BF and Primiparous did not differ statistically. Why are these results not consistent? Could you plots 3a & b for both adjusted and unadjusted and then remove tables 4a & b. In the Article focus section of the paper, the authors indicated that,

	"Assess the critical time points for the discontinuation of BF". I don't understand why the authors did not adjust for mother's working status (and leave it the final model) because the challenges posed by BF was that infant mothers may be forced to return to full time work with a shorter breastfeeding span given impressive economic growth in many high-income countries including New Zealand. I don't understand why you need predictive model to determine the best time to discontinue BF because table 3 revealed that rate of Exclusive BF (EBF) decreases as the time (or child's age) increases – see above for one of the possible reasons.

VERSION 1 – AUTHOR RESPONSE

Reviewer: Dr Vivien Swanson,
Senior Lecturer
Psychology Division, School of Natural Sciences, University of Stirling, Stirling K9 4LA UK

I have research interests in the psychology of breastfeeding.

Exclusion criteria are not clearly described. Also reasons for exclusion of some who returned postcards not explained.
We have included a sentence on page 5 to explain the exclusion criteria more clearly 'There were no exclusions based on feeding intention or maternal history.' We have also added Appendix 2 giving full details of the reasons why some who returned postcards were not included. The main reasons were problems with the phone number given or incomplete information.

This is an interesting paper identifying demographic and psychological antenatal predictors of breastfeeding duration and cessation. The findings that socio-economic status, parity, attitudes and intentions predict women's decision to initiate breastfeeding is not surprising, and indeed has been found in several such studies. However this study is novel in that it uses this information to identify an algorithm which allows health professionals to predict whether or not individual women will breastfeed, and gives us some clues about continuation. It is a weakness of this study that only antenatal data (attitudes etc) is available to predict cessation - but I assume that the proposed follow-up RCT will account for postnatal changes in psychological variables due to women's experience of breastfeeding.

The study has a reasonable sample size, and socio-demographic spread, which makes the findings generalisable. Although the authors suggest that ethnicity may be a factor they have missed, this study does replicate the findings of many others, suggesting this may not be a major weakness.

The study is also novel because it uses SMS text messaging to collect longitudinal data in an unobtrusive manner which appears acceptable to participants. I was not altogether clear as to how the reliability of this method was checked, as the counter-checks used by the researchers also involved self-report. Nevertheless the researchers claim that their 'real time' assessment may be more accurate than other methods of self-report data collection.
Further information about the validity check has been included on Page 6: 'and by comparison with data collected by the health visitor.'

The study is theoretically underpinned using a Theory of Planned Behaviour (TPB) approach. Again this has been successfully used in several studies of breastfeeding so is an appropriate and useful model. This is clearly a strength of the study.

I did have some queries in respect of how the TPB data was used in the study. The description of the measurement of TPB variables in the methods section is sketchy.
We have added an appendix 3 with more details of the TPB questions.

How were TPB attitudes measured, and was there overlap between these and the IIFAS? It is surprising that TPB attitudes are not predictive of behaviour, as this has frequently been the case in previous research - although this may be a statistical artifact of inclusion of the IIFAS. It was also surprising that self-efficacy or perceived control wasn't predictive of breastfeeding, as has been the case in many studies - although this may be because this variable was only measured antenatally.
All the TPB sub-scales were significantly associated with initiating breast feeding univariately. However, PBC and attitudes were strongly correlated with and predictive of intentions so when combined in the multivariate model intentions was the strongest independent predictor so the other TPB variables did not enter. We measured TPB antenatally but we take the point that it would be interesting to measure this over time as well. Comment on this and further references have been added in the discussion on page 9.

Additionally, since intention is a major predictor, it is important to provide more detail about how this was measured.
As above we have added more detail of TPB measurement.

The point made in the introduction and discussion about the failure in current policy to discuss women's intention is well-made, since this leads health professionals to miss a key opportunity for behavioural intervention by discussing their beliefs and attitudes towards different feeding methods. The rationale for this has related to the importance of promoting a culture of breastfeeding as 'normal', so not raising the possibility of formula feeding - but in reality may have left some women feeling their views and concerns about different feeding methods have not been acknowledged or addressed.
We agree with this point.
There are some very minor points below which it would be helpful for the authors to address, but in general I thought this was an interesting piece of research which merits publication.

Minor issues:

1. 1. P3, Lines 26,27 the meaning of this isn't clear ('points to targeting...')
This has been changed to ‘could allow the development and trialling of targeted interventions.’

2. P4, Lines 36,37. It would be useful to clarify the impact of the BFI on both initiation and continuation of breastfeeding rather than just talking about 'breastfeeding rates'.
We have altered this to say ‘initiation and continuation rates’.

3. P5, Line30 As discussed above - clarify how TPB variables were measured.
This has been included in appendix 3.

4. P5, Line 31 - In what way were the questions informed by self-efficacy theory?
The variables were based on these theories. It would take a large amount of text to give detailed justification for each variable. The details are given in appendix 3.

5. P5, Line 12. Are there any exclusions? ‘There were no exclusions based on feeding intention or maternal medical history’ inserted

6. P6, line 16 onwards - I dont this the information about recruitment to focus groups needs to be

included here - and the topic guide isn't included in Appendix 1!
We have removed most of this paragraph.

7. P6, Line 35, How does repeat texting ensure reliability?
This is the standard method of checking reliability by repeating the measurement soon after the first collection and then looking for agreement. This has previously been published in JAMIA as reference 24 in the paper. Additional information has been included on page 6 to clarify the procedure for checking the validity of texting: ‘and by comparison with data collected by the health visitor’.

8. P7, line 9 onwards. What was the basis for selection of participants from the 639 postcards returned? Was there any drop-out from the texting process?
Details of reasons for drop-out are now included in appendix 3.

9. P7. Why was the duration of follow-up 16 weeks, and why was the 'end point' assessment 4 weeks after that?
The cost and length of the study dictated some of the decisions about data collection, including the 16 week limit on texting participants and the 4 weeks for subsequent follow up. No amendment made to the paper.

10. P7. Again, exclude description of procedure for focus groups.
This paragraph does not report the procedure, but the numbers included in the qualitative arm of the study. We feel it is relevant to include this detail.

11. P7 There is no heading for the start of the 'results' section. Heading now included

12. P8, line 7, What is 'major' occupational group? and how is occupation coded - there is no information on this.
We have added a reference to the ONS classification system on page 5 which we then reclassified into the four broad groups as shown in the tables.

13. P8, line 41. The authors could acknowledge potential for a 'hawthorne' effect which may have influenced breastfeeding rates, or at least the reporting of these.
This is a good point and we have added the text ‘The texting in itself may have acted as an intervention to encourage continuation of breastfeeding.’ to page 8.

14. P9, line 22. Explain how the study findings can be used 'to identify women who need additional support' - not sure how they will do this from the existing data.
Added in ‘using a prediction model’ on page 9 to explain this comment.

15. P9, line 40 This study doesnt tell us why women stopped breastfeeding, so there will indeed be a need to explore these factors in future studies.
Added in ‘future studies could explore this issue.’

16. I was slightly confused by the breastfeeding time series data. The scale on Figures 3a and 3b suggests that some women stopped breastfeeding long after the 14 week + 4 weeks end points of the study (max 126 days). Is this data extrapolated in some way?
Difficulties with contacting some participants led to protracted follow up in several cases. We have added in ‘Some of this follow up was protracted due to difficulties in contacting several participants.’ on page 7.

17. P17, line 16 onwards. I wasn't clear about the difference between the categorical and continuous

measures of intention.
There was a clear dichotomy in the intentions variables and so it was sensible to treat it as categorical rather than continuous. Explanation has been included below table 1.

18. P20, Table 4a. I didnt understand what 'Total IIFAS score + 10 units' meant?
All variables are presented as ORs or RRs in software output for an increase in one unit. For some variables this gives values which are very close to 1 but nevertheless statistically significant. It is common practice to then give the OR or RR for a, say 10 unit increase (e.g. age) as a more meaningful unit of change. Hence, '+10 units' for the IIFAS score which is arbitrary but a meaningful change.

Reviewer: Kingsley. E. Agho, PhD, MPH
Research Fellow in Epidemiology and Biostatistics School of Medicine University of Western Sydney Australia

If the study is a prospective cohort, the authors should report relative risk and not odds ratio – is there any reason for reporting Odd ratios instead of relative risk.
We agree that Relative Risk would be preferable as this study was prospective

In table 3 what is the rational for choosing the following weeks – 6weeks, 8 weeks and 16 weeks ?. The WHO recommended further disaggregate for reporting EBF as: 0-1 months; 2-3 months , 0-3 months and in this case, 0-4 months.
We chose 6 and 8 weeks to allow comparison with routinely collected local and national data. National data collection at the time of the study did not reflect the WHO recommendations. Note that we can record precisely in time because of the SMS approach whereas the WHO categories are broad and lack precision.

In table 1, mean should be in one decimal place (1 dp)
We agree and have rounded the SD to one decimal place.

In table 2 remove x2 test and Parous - any bF vs primiparous can't be significant because the confidence interval (CI) around the OR includes 1.0. I think it was a typo error.
We have checked this result and it is exactly as SAS outputs. However, we have discovered that SAS does not necessarily produce RRs and CI in the way expected for categorical variables. We have redone this and now give corrected RRs and CIs for the categorical variables. The continuous variables are unchanged.

In table 2, i would like to see both adjusted and unadjusted Odd ratios reported and the OR and the CI around the OR should be in 1 dp.
We believe this would add to the complexity of the table without adding anything useful. We have changed the results to 2 dps as this is the standard way of presenting ORs and RRs. For example, with only 1 dp there would be no difference between 1.05 and 1.14, the former is a 5% increase while the latter is a 14 % increase.

Combine 4a & b
It is not clear what is meant by combine here. The 'exclusive' breastfeeding group are clearly a subset of the 'any' breastfeeding group and so the latter is the two combined. We feel it is important to present both as different papers do not make it clear which they mean. In any case, it is standard procedure to report both 'any' and 'exclusive' breast feeding as these are clearly not the same.

Remove Appendix 1 – it doesn't add anything to the paper – it makes the paper look ugly.
We believe that this algorithm could be useful to researchers and clinically useful and unsure of the definition of ‘ugly’.

The plots in figs 3a & b indicated that there are statistical differences between Parous-any BF, Parous – No BF and Primiparous but the hazard ratios reported in tables 4a & b revealed that Parous-any BF, Parous – No BF and Primiparous did not differ statistically. Why are these results not consistent?
These results are consistent. The figures are unadjusted plots for a single variable and hence demonstrate a univariate relationship. We should point out that no tests are presented in the figures here so the relationships may or may not be significant. The tables 4a and b add additional information on the adjusted effects of a number of variables.

Could you plots 3a & b for both adjusted and unadjusted and then remove tables 4a & b.
Tables 4a and b present the final model from which the algorithm is derived and demonstrate multivariate relationships and so are important to present. The plots are unadjusted and so not presenting the same information. Adjusted plots are difficult but they also could only show one variable at a time and makes assumptions that the other variables are not important.

In the Article focus section of the paper, the authors indicated that, “Assess the critical time points for the discontinuation of BF”. I don’t understand why the authors did not adjust for mother’s working status (and leave it the final model) because the challenges posed by BF was that infant mothers may be forced to return to full time work with a shorter breastfeeding span given impressive economic growth in many high-income countries including New Zealand.
Mother’s current working status was not measured. We did assess mother’s occupational status from baseline so we cannot draw any conclusions regarding the effect of returning to work on breastfeeding duration.

I don’t understand why you need predictive model to determine the best time to discontinue BF because table 3 revealed that rate of Exclusive BF (EBF) decreases as the time (or child’s age) increases – see above for one of the possible reasons.
Our predictive model determines the critical time points for discontinuation of breastfeeding. In most cases this is much shorter than the WHO recommendation of exclusive breastfeeding for the first 6 months. Our data could allow health care professionals to target interventions appropriately in order to support women to prolong breastfeeding and help some achieve the WHO target.