PEER REVIEW HISTORY

ARTICLE DETAILS

TITLE (PROVISIONAL)	Public acceptability of financial incentives for smoking cessation in pregnancy and breastfeeding: a survey of the British public
AUTHORS	Hoddinott, Pat; Morgan, Heather; Maclennan, Graeme; Sewel, Kate; Thomson, Gillian; Bauld, Linda; Yi, Deokhee; Ludbrook, Anne; Campbell, Marion

VERSION 1 - REVIEW

REVIEWER	Dr Marita Lynagh University of Newcastle Australia
REVIEW RETURNED	08-May-2014

GENERAL COMMENTS	There are lots (ie. too many) data tables. Some of these (eg. Figure 1 & Table 2 - essentially display the same information). COuld the authors consider reducing the number of those to only those most relevant & critical to the study findings. Table 3 title (Summary of independent predictors...) is confusing . Is it agreement OR disagreement? It can't be both. This is a well written manuscript describing a methodologically sound assessment of public opinion regarding the use of incentives for smoking cessation in pregnancy & breastfeeding. The authors have evaluted a 'shortlist or seven promising incentive strategies'. This list is limited only to shopping vouchers for women & additional funding for providers. Given the mixed findings or the study, it would be useful for the authors to comment on whether the 'limited' range of options may have influenced the findings. Did the authors consider other 'forms' of incentives such as cash incentives or voucher for baby-related products (eg. diapers etc)? There is no reference to previous studies in the Introduction or Discussion that have examined public acceptability of financial incentives & how the findings from this study compare to those. The manuscript would be strengthened by reference to these studies (ie. Long et al, 2008, J G Int Med; Lynagh et al, 2011, Nic & Tob Res).

REVIEWER	Jane Sandall King's College, London
REVIEW RETURNED	26-May-2014

GENERAL COMMENTS	This study aimed to assess public attitudes to a range of promising incentives to 1) reduce smoking in childbearing women, and 2) improve breastfeeding. This is an important issue which as the

	authors state, should be assessed prior to developing intervention studies to ensure that the intervention is acceptable to the target population and does not end up widening inequalities. There are two issues with the paper that require further consideration. First, that of addressing smoking cessation and breastfeeding in the same survey. On first reading I thought that the study was about interventions to reduce smoking cessation during pregnancy and the breastfeeding period, and I wonder how many of the respondents were equally confused. Public attitudes to smoking and breastfeeding are complex, different to each issue, and as the authors have found out, differ by a range of demographics and are influenced by personal experience. It is possible there is a serious risk of confounding response by addressing the two issues in the same survey. Indeed the authors found framing effects around which issue was asked about first. I think this makes the results more difficult to interpret and would welcome more discussion on this possibility, and should be mentioned as a study limitation. The second issue is of generalisability. Ipsos Mori used a quota sample. Quota samples although commonly used are not as scientifically robust as random sampling. I would like to know how the demographics of the sample in Table 1 compares with those of the wider population from which it is drawn, and if different, would like the authors to address the issue of generalisability in the discussion, and mention as a study limitation if appropriate. More detailed points are as follows. The title of the paper is not clear that two issues are being addressed simultaneously. Article summary Findings are reported in the article summary, I am not sure this is a strength or limitation? P5:23 were all respondents asked about acceptable amount, or a subsample. It is not clear from the wording. P5:42 please provide a reference for quota sampling. P6:3 please provide a reference for sample size calculation, and what was this based on? P7:46 is there a way of presenting results in table 3 in a clearer format, these are quite confusing. P9:22 Can the wording be changed so a negative is not used, and change ‘who did not disagree’ to ‘who agreed’? Ditto, next para. This wording is hard to interpret. Overall, the results are presented for the general public, but the target population are parents of childbearing age. Some results have been presented for this group, but not all. It would be helpful to the reader if the results for the subgroup were consistently presented. Discussion It would be helpful if the discussion focused firstly on the attitudes of the general public and secondly of the target group of parents of childbearing age. The authors only discuss women of childbearing age, but surely such interventions would need to be acceptable to men of childbearing age as well? Finally, it is not clear whether ethical approval and consent was sought for this study. This needs to be discussed.

VERSION 1 – AUTHOR RESPONSE

Dr Marita Lynagh

1. Too many tables. We agree that Figure 1 and Table 2 present the same data in different ways. However we believe that readers will take away different messages from the two. Figure 1 illustrates the spread of responses over the 5 categories, whereas Table 3 summarises the data. We will accept the editors' final decision on this. We support open access to data, and have therefore provided extensive data tables for web linked publication.
2. Table 3 title. The results were modelled on agreement and the title has been changed to reflect this.
3. Limited range of incentives in survey. On p6, we have described some of the other incentives that we considered and explain that a full justification for the shortlist is described in detail in the HTA monograph which is now in press.
4. Additional references: Lynagh M, Bonevski B, Symonds I, Sanson-Fisher RW. Paying women to quit smoking during pregnancy? Acceptability among pregnant women. Nicotine Tobacco Res 2011;13:1029-36. This study is included in our evidence synthesis but was not included in our paper as the survey sample included pregnant women only. However, we have added a sentence to discuss this on p 10, as it is consistent with our findings. The reference: Long, J. A., Helwig-Larsen, M., & Volpp, K. G. (2008). Patient opinions regarding 'pay for performance for patients'. Journal of General Internal Medicine, 23, 1647–1652 was excluded as it investigates patients in waiting rooms of 2 university based clinics in America and does not focus on incentives around childbirth. The questions asked about paying people to quit smoking, lose weight, control their blood pressure or control their diabetes. It is debatable how relevant this paper is to our survey. We know from the Diepeveen et al. systematic review (reference 1) that incentives to improve the health of children are more acceptable than to adults in general. We therefore stand by our decision not to include this. We have corrected the reference order as we noticed an error with reference 22 appearing in the text after reference 23.

Jane Sandall

1. To address Jane's confusion, we have clarified the aims and objectives to reduce the likelihood of confusion for readers by changing the phrase" incentives for smoking cessation and breastfeeding" to: "incentives for smoking cessation and FOR breastfeeding". We have done this in the title, the abstract, the article summary box and on page 4.
2. The issue of confounding by addressing incentives for two behaviours in the same survey. We consider that researching 2 behaviours concurrently can be considered as both a strength and a limitation. We have added text to the article summary: p4, and to page 10 to expand this, with addition of two new references (20,21) reporting the association between breastfeeding and smoking cessation, highlighting that they confound each other. Our view is that health related behaviours in pregnancy should be investigated concurrently as they are complexly inter-related for women and their social networks. The overall findings from our mixed methods study (reference 16), which include qualitative data, support a less reductionist approach to understanding behaviour change.
3. Generalisability and quota sampling. The merits of quota sampling versus random probability sampling have been debated for many years and a number of studies have compared the results from high quality, well controlled quota surveys with those from random probability surveys and other trusted data sources (see Moser and Stuart (1953), Stephenson (1979), Marsh and Scarbrough (1990), Orton (1994), Myant and Hope (2006)). An independent assessment of the Scottish Environmental Attitudes and Behaviours Survey by Raab (2009) concluded that "the [quota] survey design has had very little effect on the precision of estimates compared to what would have been achieved from a simple random sample." The overwhelming message from these studies is that estimates from high quality, well controlled quota and random probability samples are generally comparable: most comparisons reported in the above studies showed no or small differences in estimates and precision between sample types. Both Stephenson (1979) and Orton (1994) present evidence suggesting that differences arising from comparisons between probability sample results and quota sample results are in-line with chance expectation. We have added text on p9 to address

this and a new reference 18. Raab (2009)
4. How does our sample (Table 1) compare with the general population? Firstly we have changed UK to British in several places, to emphasise that this sample did not include Northern Ireland (p2, p3, p8, p10, p22). We have clarified on p6, that quota samples were set based on a combination of the National Readership Survey (http://www.nrs.co.uk/) and census data which is considered an appropriate reference database for the British population. Additional tables comparing weighted and un-weighted data have been included as a new Web 3 document, with re-numbering of subsequent Web documents. Comparing our sample with the Adult Health in Great Britain Survey 2012 http://www.ons.gov.uk/ons/rel/ghs/opinions-and-lifestyle-survey/adult-health-in-great-britain--2012/stb-health-2012.html a similar proportion smoked (1 in 5); there is no comparable population data for any children breastfed. We therefore consider our sample to be representative of the British public.
5. The title is unclear. We have revised this and would be happy to consider any other suggestions that the editorial team has.
6. Article summary reports findings. The findings mentioned are to support our statement about the originality of our research, which we consider to be a strength.
7. P5, line 23. Only the subsample who answered strongly agree, agree or neither agree, nor disagree were asked about the value of incentives. Asking people about the value of an incentive when they disagree with providing them would produce results that would be difficult to interpret. We have changed the text to improve clarity.
8. P5, line42. We have provided a new reference 18 (Raab 2009) which describes an independent assessment of the approach to quota sampling used by IPSOS MORI
9. P6, line 3. The a priori sample size calculation was stated in our NIHR/HTA grant application and in the study protocol. It was based on standard survey size calculations for estimating proportions based on the desired margin of error and level of confidence, 3% and 95% respectively.
10. P7, line 46. We suggest another way to present Table 3 below and we will be guided by the editor's advice on which to include in the paper.
11. P9, lines 8 and 22. "did not disagree" has been clarrified as the subgroup where respondents answered strongly agree, agree and neither agree, nor disagree.
12. Consistency of reporting the results for the subgroup who were of childbearing age. We do not fully understand this comment. Our overall aim was to assess the acceptability to the British general public. We therefore analysed results by participant age and have presented the results consistently compared to the other subgroups (Table 3). We chose age >=65 as the reference group. The age subgroups 18<=24; 25<=34; 35<=44 were all found to be independent predictors of agreement with all incentive strategies (see full results tables in Web 4). Our aim was not to assess the acceptability of incentives to British people of childbearing age.
13. The authors only discuss women of childbearing age. We agree that incentives need to be acceptable to both men and women. We have tried to be explicit that our survey findings report the views of women (any age) compared to men (any age). We did not conduct a subgroup analysis for women of childbearing age (<=44), compared to men of childbearing age (<=44), although it might be interesting to do so. To improve clarity we have changed "being of childbearing age" to "men and women of childbearing age" in two places: p2, line50; p9, line38.
14. Ethics approvals were for the protocol which described this survey.

Marsh, C. and Scarbrough, E. (1990). Testing nine hypotheses about quota sampling. Journal of the Market Research Society, vol. 32 no.4
Moser C. and Stuart, A. (1953) An experimental study of quota sampling. Journal of the Royal Statistical Society, series A, vol. 116, no. 4.
Myant K. and Hope, S. (2006). "A comparison of quota and random samples for measuring sport participation." Research report for the Scottish Executive.
http://www.scotland.gov.uk/Publications/2009/03/25155151/9

Orton, S. (1994). Evidence of the efficiency of quota samples. Survey methods newsletter, vol. 15, no. 1

Raab, G. (2009). Technical Annex to Davidson S., Martin C., Treanor S. Scottish Environmental Attitudes and Behaviours Survey 2008: Technical Report. Scottish Executive. (Accessed June 2014) http://www.scotland.gov.uk/Resource/Doc/265441/0079493.pdf
Stephenson, C., B. (1979) Probability sampling with quotas: wan experiment. Public Opinion Quarterly, vol. 43, no. 4.