

PEER REVIEW HISTORY

BMJ Open publishes all reviews undertaken for accepted manuscripts. Reviewers are asked to complete a checklist review form (<http://bmjopen.bmj.com/site/about/resources/checklist.pdf>) and are provided with free text boxes to elaborate on their assessment. These free text comments are reproduced below.

ARTICLE DETAILS

TITLE (PROVISIONAL)	Evaluating the importance of policy amenable factors in explaining influenza vaccination: a cross-sectional multinational study
AUTHORS	Wheelock, Ana; Miraldo, Marisa; Thomson, Angus; Vincent, Charles; Sevdalis, Nick

VERSION 1 - REVIEW

REVIEWER	Cath Jackson University of York, UK and Valid Research Ltd, UK
REVIEW RETURNED	31-Oct-2016

GENERAL COMMENTS	<p>This is very written paper. It is quite a hard read in places as there is so much detail to cover but I am not sure that can be avoided. I have a few suggestions for minor edits to aid reader understanding and 2 more significant suggestions/queries. I am unable to comment on the appropriateness of the statistics analysis so have read the paper assuming this is all correct.</p> <p>Article summary (and elsewhere in paper) - I am not sure what you mean by "individual and condition perceptions"</p> <p>Methods - I got slightly confused about whether where surveyed online or interviewed. You use the word "surveyed" in the abstract and "interviewed" in study sample.</p> <p>More significantly:</p> <ol style="list-style-type: none">1. You acknowledge the limitations of a cross sectional study in the discussion however throughout the paper you write as if you are predicting vaccination behaviour. I think this is misleading to the reader as your measure is of past behaviour. I would suggest rewording to tone this down or making clearer much earlier in the paper that you are not predicting future behaviour.2. You write that you have achieved nationally representative samples. I would have thought that the samples would ideally be representative of those who are eligible for the flu vaccine rather than nationally representative. Having an eligible health condition/age not surprisingly differentiated between vaccinated/unvaccinated participants, do you think that your study findings would look different if you have only include those people?
-------------------------	---

REVIEWER	David Elliman Consultant in Community Child Health Whittington Health United Kingdom
REVIEW RETURNED	20-Nov-2016

GENERAL COMMENTS	<p>This paper aims to address the reliability of previously recognised factors in the uptake of influenza vaccine that might be amenable to policy change. This is an important issue as influenza vaccine uptake around the world is well short of the minimum standard set by WHO.</p> <p>The authors claim that socio-psychological variables are inherently amenable to policy change (page 3) and later to policy and behaviour change (page 5) implying the two are different. A definition would be helpful of what is meant by each of the terms. The strength of this paper is supposed to be the use of robust multivariate analysis. The statistical analysis is therefore described in considerable detail. This would be beyond the understanding of many, if not the majority, of the proposed audience and may result in many not getting to the essence of the article. I wonder whether a lot of this could belong in an appendix for those who wish to read it, while only the main points are in the body of the article.</p> <p>The authors state that "studies frequently include past vaccination or intention to vaccinate as predictors thereby artificially boosting the explanatory ability of the model without necessarily explaining real-world behaviours" (page 5). I do not understand this. Relying on intention is indeed suspect, so one can only use past vaccination history to know what people have actually done. What alternative are the authors suggesting, or do they mean distant past as opposed to immediate past? If so, they should make this clear. What is a "non-probability online panel" (page 6)?</p> <p>What is the author's evidence that 'speeders' response (page 6), where non-homogeneous, is any less accurate than a slow responder? They could just be more certain in their views. Taking them out could introduce bias. What is the basis of an "optimal survey length of 20 minutes"? If optimal is 20 and average is 16 minutes, that implies many people were far from optimal, either too fast or too slow. Does this mean the results are likely to be inaccurate.</p> <p>If the authors are correct in stating "This is because when predictors are correlated, as it is often the case, the order of variable entry can have an effect on the estimated model parameters. Thus, variables were entered in a sequence according to their conceptual importance: variables which had been consistently associated with vaccination uptake in the past were entered first and those which had been explored less were entered last." It naturally follows that as they have decided the order of entry of variables into the model, they are potentially influencing the results. There is nothing to refute this in the paper.</p> <p>The authors state there were no significant differences in demographics between the sample and the general population (page 10), yet, in table 1, in the US population there is a much larger proportion with a university degree in the sample than in the population.</p> <p>When the survey was conducted, influenza vaccine was recommended for all the population ≥ 6 months old in US. In UK, the vaccine was only offered to those adults less than 65 years old. I don't know what the policy was in France. Do the uptake figures, for those under 65 years, in the table, apply to all those given the</p>
-------------------------	---

	<p>vaccine or only those with risk factors?</p> <p>Bearing in mind that, in the UK, vaccine is offered universally to the elderly and only selectively to those with chronic conditions under 65 years old. It is inevitable that those vaccinated will be older (page 12, line 19). Would it not be more useful to separate the 65 years old and those younger, when it comes to analysis, at least in those countries where the policy is different depending on age?</p> <p>The large paragraph on page 13 is extremely difficult to read as it is a list. As the p values are in the tables, they could probably be omitted from the text.</p> <p>Contrary to what is implied on page 22, influenza vaccine is available to all those 65 or over, free in UK, as well as those with eligible health conditions.</p> <p>UK GPs are already paid for each influenza vaccine given, so what is the evidence that adding them to QoF would make a difference (page 24)? It is only likely to do so if payment was only target related. This would be extremely difficult to introduce.</p> <p>The authors were reassured by the fact that pilot results did not feel the survey was long (page 26). This was a selected group. 4-6% of survey participants were speeders or flat-liners, both groups that may indicate fatigue, and an additional number didn't complete the survey. 5-10% may have found it too long on this basis.</p>
--	--

REVIEWER	<p>Jing Huang Department of Biostatistics and Epidemiology Perelman School of Medicine University of Pennsylvania Philadelphia, PA 19104-6021 USA</p>
REVIEW RETURNED	26-Dec-2016

GENERAL COMMENTS	<p>Comments to editor:</p> <p>In this manuscript, the authors conducted a cross-sectional multinational study to evaluate the association between socio-psychological variables and vaccination decisions after adjusting for demographic variables, socio-economic status and health conditions. The research gap and significance of the study was well established, the study was carefully designed, and the manuscript was well-written. I recommend minor revision since I have several concerns.</p> <p>Comments to authors:</p> <p>In this manuscript, the authors conducted a cross-sectional multinational study to evaluate the association between socio-psychological variables and vaccination decisions after adjusting for demographic variables, socio-economic status and health conditions. The study subjects were sampled using a stratified random sampling technique and multivariable logistic regression model was used to analyze data. The results suggested that socio-psychological variables can explain vaccination behavior better than demographic, socio-economic and health variables. The authors gave a balanced discussion on the strengths and limitations of the study.</p> <p>Major concerns:</p> <p>1. As mentioned by the authors, the study was subject to reporting bias. The influence of the reporting bias to the study results should</p>
-------------------------	--

	<p>be discussed.</p> <p>2. In page 8 line 12, the authors mentioned they dichotomized the continuous variables with missing values. This action is very subjective. The authors should justify why the dichotomization will not influence the results and statistical power of the analysis. The authors should also mention which variables were dichotomized and which were not, and discuss what was the potential consequence due to the dichotomization.</p> <p>3. After building model 1, the authors “manually removed one” insignificant variable “at a time”, (in page 8 line 34). The authors should describe what was the the sequence of removing variables and justify the criteria used to determine the sequence.</p> <p>4. The authors claimed “We developed robust regression models” in page 3 line 7. This statement is overly strong and misleading. The authors did not develop statistical methods but they used “robustness” check to select variables when building the final model. It should be clarified.</p> <p>5. In page 8 line 38, please provide details about what was of the specific criteria of “robustness” based on the change of regression coefficients.</p> <p>6. The authors assumed correlation coefficient between dependent and independent variables was 0.1 to calculate the sample size. What was the observed correlation coefficient based on the data? Please discuss.</p>
--	---

REVIEWER	Eleni Verykouki Aristotle University of Thessaloniki, Greece
REVIEW RETURNED	22-Jan-2017

GENERAL COMMENTS	<p>This is an interesting survey evaluating policy factors associated with influenza vaccination. The article is well written and of interest to the general audience of the journal. Statistics are accurately described and presented. However, I have a few comments:</p> <ul style="list-style-type: none"> - Page 6, "Study Sample": Sample size calculation is rather superficial based on an arbitrary assumption. Authors should better justify this choice that merely looks as an effort to justify a convenience sample and also take into account the multi-regional setting. - There is a wealth of p-value based statistical testing. Some type -I error adjustment would have been useful given that authors base their conclusions from unadjusted p-values which are marginal in many cases. - Given that this is a multi-regional study there are issues pertinent to questionnaire standardization to local populations/different languages. Authors must show some evidence of questionnaire standardization. - Some internal consistency measures such as Cronbach's alpha must be given.
-------------------------	--

VERSION 1 – AUTHOR RESPONSE

Reviewer 1

This is very written paper. It is quite a hard read in places as there is so much detail to cover but I am not sure that can be avoided. I have a few suggestions for minor edits to aid reader understanding and 2 more significant suggestions/queries. I am unable to comment on the appropriateness of the statistics analysis so have read the paper assuming this is all correct.

Thank you for your positive response to our work. As you will see, we have taken your points fully into consideration in our revisions. We opted to share the detail with the readers so that the scale, methods and overall nature of this multi-country study can be fully appraised.

1. Article summary (and elsewhere in paper) - I am not sure what you mean by "individual and condition perceptions"

Thank you for noting this – on reflection this is a point that required linguistic clarification. We explain in Methods that “socio-psychological measures were disease or vaccine-specific to avoid misinterpretation, they aimed to capture individuals’ own perceived risk rather than the risk they would judge in the general population (e.g. “With no flu vaccine, I would feel very vulnerable to the flu”). We further aimed to capture how people perceive their own risk depending upon having had the vaccination or not (e.g. “Without the flu vaccine, I am sure I would get influenza this winter”). The latter are judgements or perceptions conditional upon an individual’s vaccination status, hence the original term. To avoid confusion, we have provided an improved version of this explanation earlier on in the Introduction (p. 5) and referred to it further in the Methods (p. 8).

2. Methods - I got slightly confused about whether where surveyed online or interviewed. You use the word "surveyed" in the abstract and "interviewed" in study sample.

Participants were both surveyed online and interviewed via the phone, depending on socio-demographic characteristics (age and socio-economic strata), as explained in Methods. For clarity, we have now used “surveyed” when referring to the whole sample and “interviewed” when only referring to the telephone interviews.

More significantly:

3. You acknowledge the limitations of a cross sectional study in the discussion however throughout the paper you write as if you are predicting vaccination behaviour. I think this is misleading to the reader as your measure is of past behaviour. I would suggest rewording to tone this down or making clearer much earlier in the paper that you are not predicting future behaviour.

We understand your concern and we have reflected on this within our team for a while. Consequently, we have added the term “past” to vaccination or vaccination behaviour and “recent” to vaccination uptake throughout the manuscript when referring to our outcome measure. We have also added in the limitations an explicit statement that the design precludes any attempt to predict future behaviours.

4. You write that you have achieved nationally representative samples. I would have thought that the samples would ideally be representative of those who are eligible for the flu vaccine rather than nationally representative. Having an eligible health condition/age not surprisingly differentiated between vaccinated/unvaccinated participants, do you think that your study findings would look different if you have only include those people?

Although influenza vaccination is generally available free of charge to vulnerable populations, mainly for cost-effectiveness, it is also available at a cost to all adults who wish to receive the vaccine (with

rare exceptions for medical reasons). In the US, for example, around a third of not at risk adults get vaccinated every year and the CDC recommends the vaccine to all adults (in the UK and France, this statistic is not publicly available). Importantly, it is now recognised that seasonal influenza vaccines are more effective in healthy, younger adults, and thus by vaccinating, they can also provide indirect protection to those at risk. Therefore, we believe there is value in understanding coverage among non-eligible adults and what factors deter or motivate the general adult population to vaccinate against influenza. If we had focused on those at higher risk of developing complications, we would expect some of the findings to look different. This point is addressed below in response to a related comment.

We thank you for raising this issue, as it was indeed not clear why we were using a representative sample. We have now made it explicit in the Introduction (p. 4 and 5) and address the issue in the Discussion (p. 20, 22 and 26).

Reviewer: 2

This paper aims to address the reliability of previously recognised factors in the uptake of influenza vaccine that might be amenable to policy change. This is an important issue as influenza vaccine uptake around the world is well short of the minimum standard set by WHO.

Thank you for your positive view of our study – we also believe that better understanding of the social and behavioural factors that are linked to the uptake of influenza vaccination globally is critical for the delivery of effective public health programmes. Hence, this is a core element of our ongoing research programme.

5. The authors claim that socio-psychological variables are inherently amenable to policy change (page 3) and later to policy and behaviour change (page 5) implying the two are different. A definition would be helpful of what is meant by each of the terms.

We have modified the sentence you refer to and specified that “they are inherently amenable to policy and interventions to change behaviour” (p. 4). We think it is now clearer that these factors can be affected by policy in general, whether or not it is purposively designed to influence behaviour, and specific interventions seeking to change health behaviours. Thank you very much for raising this for clarification.

6. The strength of this paper is supposed to be the use of robust multivariate analysis. The statistical analysis is therefore described in considerable detail. This would be beyond the understanding of many, if not the majority, of the proposed audience and may result in many not getting to the essence of the article. I wonder whether a lot of this could belong in an appendix for those who wish to read it, while only the main points are in the body of the article.

We understand your concern and we have spent some time discussing within our team how best to trim the main body of the paper to ensure readability, whilst preserving the detail for the interested reader. We have moved a good part of the analysis to Supplementary material, as you suggested.

7. The authors state that "studies frequently include past vaccination or intention to vaccinate as predictors thereby artificially boosting the explanatory ability of the model without necessarily explaining real-world behaviours" (page 5). I do not understand this. Relying on intention is indeed suspect, so one can only use past vaccination history to know what people have actually done. What alternative are the authors suggesting, or do they mean distant past as opposed to immediate past? If so, they should make this clear.

Thank you for your comment. We agree that measuring intention is often less accurate than self-reported behaviour. Yet, a number of theories of health behaviour and empirical studies have used intention as their outcome measure. The problem we are highlighting, however, is a different and perhaps more relevant drawback; thus, it is important to us that it is clearly understood. Some studies trying to understand vaccination behaviour include as independent variables past behaviour (e.g. historical uptake when the outcome measure is recent vaccination) or intention to vaccinate during the following season when the outcome measure of interest is recent vaccination. The problem with this strategy is that these independent variables (historical vaccination and intention to vaccinate in future) are known to be proxies of vaccination, i.e. people who have vaccinated against influenza are more likely to do so again in the future. Thus, including these variables in a model aimed to explain why people vaccinate or do not vaccinate does not make a lot of sense, since all these variables are doing is increasing the amount of variance the model can explain (because past and future vaccination are significantly correlated, and are potentially endogenous, i.e. third factors determine the statistical relationship between both) without really shedding light of the reasons why people do so (e.g. they feel vulnerable every year). We have provided some examples in the Introduction to clarify this point (p. 5).

8. What is a “non-probability online panel” (page 6)?

A non-probability online panel is a panel of participants (usually large – over 1 million people), which is not representative of the whole population of a country. This is because such panels include those who can and are interested in participating, usually for a fee, and do not normally include people who cannot or are less able to use the internet. That is why we decided to employ a combined recruitment strategy to access the latter segments using telephone interviews. Market research agencies work with these types of panels and aim to draw representative samples from their panel universe. The use of probability online panels, such as the American Trends panel, is preferable, yet they are expensive to access. We have added a box (Box S2) in supplementary material with this explanation, for the benefit of the readers.

9. What is the author’s evidence that ‘speeders’ response (page 6), where non-homogeneous, is any less accurate than a slow responder? They could just be more certain in their views. Taking them out could introduce bias. What is the basis of an “optimal survey length of 20 minutes”? If optimal is 20 and average is 16 minutes, that implies many people were far from optimal, either too fast or too slow. Does this mean the results are likely to be inaccurate.

Thank you for seeking clarification on this issue. Given that research shows that the length of a survey depends on a number of factors such as the contents and format of the survey, as well as participant-related factors, the length we used as a reference was the approximate time participants needed to complete the survey when we piloted the materials. We have now included this information in the Methods and have also specified the number of pilot interviews per country, as we felt it was not clear we had conducted pilots in all three study countries.

One of the problems of using online panels is that a fraction of the participants are “professional respondents”. Some earn a living by answering as many surveys as possible and tend to lie in order to become eligible. These participants will usually ‘flat-line’ (e.g. choose 1 for all the Likert scale questions) and speed (complete surveys significantly faster than the average responder, but try to avoid “looking” like a flat-liner). In light of such practices, we excluded these participants as a quality measure. Market research companies use different methods to detect these respondents. A typical strategy is only excluding flat-liners without considering the amount of time they took to respond (for more information please see Menictas C, Wang P, Fine B. Assessing flat-lining response style bias in online research. *Australasian Journal of Market & Social Research* 2011;19(2):34-44). We also excluded speeders because, although their responses were not always homogeneous, it is highly

unlikely that a person will be able to accurately respond to a 20 minutes survey in 8 minutes (for more details please see Toepoel V, Das M and Van Soest A (2008). Effects of Design in Web Surveys: Comparing Trained and Fresh Respondents. Public Opin Quart 72(5): 985-1007). There is no conclusive evidence that people who are slower at responding provide less accurate information. Someone who took more time to respond than the average may have needed more time to think, may have needed more breaks or may have been interrupted more. For clarity, we have added a reference in Methods to substantiate our decision (p. 7).

10. If the authors are correct in stating “This is because when predictors are correlated, as it is often the case, the order of variable entry can have an effect on the estimated model parameters. Thus, variables were entered in a sequence according to their conceptual importance: variables which had been consistently associated with vaccination uptake in the past were entered first and those which had been explored less were entered last.” It naturally follows that as they have decided the order of entry of variables into the model, they are potentially influencing the results. There is nothing to refute this in the paper.

It is certainly correct that when using hierarchical approaches, the order in which variables are entered can affect the results. That is precisely why we are advocating for this order to be evidence-based, as has also been recommended by established statistical experts (we refer to their work in the manuscript). The number and pertinence of the variables entered can also affect the results. Therefore, our paper highlights the importance of reporting the variance in behaviour explained by the model, both overall and per block, not only the significance of the factors. To prove our point, as explained in Methods: “Separate analyses entering the blocks of variables in reverse order were also performed (i.e. from block 7 to block 1) to evaluate whether the order in which variables were entered significantly modified our results”. As mentioned in Results: “When blocks were entered in reverse order, demographic, socio-economic and health variables contributed little to the variance in past vaccination behaviour – 3% (US), 1% (UK) and 0% (France). This is not surprising, since people’s characteristics have an effect on their perceptions, thus they explain some of the same variance. This result further proves that poorly specified models – which are not evidence-based – that either omit important control variables or lack a robustness assessment of different model specifications lead to biased estimates.” Encouragingly, as we now have noted in Results (p. 19) “The magnitude and significance of the relationship between independent and dependent variables varied little between the first models (M1) – where all the variables were entered at the same time – and the last models (M6-M8) – where a reduced number of variables were entered in blocks –, which is a further indication of the robustness of our findings.”

11. The authors state there were no significant differences in demographics between the sample and the general population (page 10), yet, in table 1, in the US population there is a much larger proportion with a university degree in the sample than in the population.

You are absolutely right, thank you for spotting the mistake. We apologise for this oversight. We have added in Table 1 the significant results of the Fisher’s test and have also referred to this potential bias in the Limitations section. Moreover, we further describe in Methods the quotas used to attain our sample, to explain why education did not match that of the general US population (p. 6).

12. When the survey was conducted, influenza vaccine was recommended for all the population ≥ 6 months old in US. In UK, the vaccine was only offered to those adults less than 65 years old. I don’t know what the policy was in France. Do the uptake figures, for those under 65 years, in the table, apply to all those given the vaccine or only those with risk factors?

As our study examines the adult general population, the category “<65 vaccinated” includes all adults (both healthy and with eligible chronic conditions) under 65 years old. The US publishes that statistic,

whereas in the UK and France it is not publicly available. Our response to comment number 4 (first reviewer) details why we aimed to use adult nationally representative samples and not only study risk-groups. Additionally, we have provided more information about the vaccine recommendations in the Introduction (p. 4) and have addressed in the Discussion why we believe it is important to study the motivation of healthy adults to get vaccinated (p. 20, 22 and 26).

13. Bearing in mind that, in the UK, vaccine is offered universally to the elderly and only selectively to those with chronic conditions under 65 years old. It is inevitable that those vaccinated will be older (page 12, line 19). Would it not be more useful to separate the 65 years old and those younger, when it comes to analysis, at least in those countries where the policy is different depending on age?

We understand your point and have partly addressed it above. We agree that it would be desirable to model the responses of at-risk groups and to compare them with those of healthy adults. However, if we split the sample we will not have sufficient power to run the analyses confidently (some cells had less than 5 observations). Yet, by including demographics and health conditions as controls, our analyses show that, taken together, socio-psychological factors are driving vaccination, more so than age and having an eligible health condition. Further, although in the UK the vaccine is not recommended to all adults, unlike in the US, we found that 27% of healthy adults vaccinated. In our opinion, this finding alone warrants further studying coverage and vaccination drivers among healthy working adults in the UK. We have addressed this issue in the Discussion and state that future research testing our findings among adequately powered samples of eligible people is warranted (p. 20 and 22).

14. The large paragraph on page 13 is extremely difficult to read as it is a list. As the p values are in the tables, they could probably be omitted from the text.

Thank you for your suggestion. We have omitted the paragraph.

15. Contrary to what is implied on page 22, influenza vaccine is available to all those 65 or over, free in UK, as well as those with eligible health conditions.

You are absolutely right, we should have been more specific. We have used “high-risk groups” instead and have changed the order of “free at the point of delivery or affordable for most” to reflect that the first statement refers to the UK and the second to France (p. 22).

16. UK GPs are already paid for each influenza vaccine given, so what is the evidence that adding them to QoF would make a difference (page 24)? It is only likely to do so if payment was only target related. This would be extremely difficult to introduce.

Yes, indeed QoF already rewards for achieving certain immunisation targets. However, the incentive for influenza vaccination does not vary depending on people’s likelihood of getting vaccinated. What we are trying to argue is that one possible way of increasing vaccination rates would be to differentiate the financial incentives in QoF according to people’s propensity to vaccinate. Specifically, the financial incentive should be higher for eligible population groups that are less likely to vaccinate, such as younger people with chronic illnesses. However, we do agree that this would require additional resources. We have now amended the paragraph to further clarify this suggestion and the challenges of its implementation (p. 25).

17. The authors were reassured by the fact that pilot results did not feel the survey was long (page 26). This was a selected group. 4-6% of survey participants were speeders or flat-liners, both groups that may indicate fatigue, and an additional number didn’t complete the survey. 5-10% may have found it too long on this basis.

We believe we have partly addressed your concern in our response to comment number 9. It is true that there is a chance that people who did not complete the survey may have found it too long. We have referred to this possibility in the Limitations, but have now added a more detailed explanation. It is also true that pilots are not usually statistically representative, but we aimed to interview participants with different socio-demographic characteristics and none felt the survey was too long, which was indeed reassuring (p. 27).

Reviewer: 3

Comments to editor:

In this manuscript, the authors conducted a cross-sectional multinational study to evaluate the association between socio-psychological variables and vaccination decisions after adjusting for demographic variables, socio-economic status and health conditions. The research gap and significance of the study was well established, the study was carefully designed, and the manuscript was well-written.

I recommend minor revision since I have several concerns.

Many thanks for the positive reception of our work.

Comments to authors:

In this manuscript, the authors conducted a cross-sectional multinational study to evaluate the association between socio-psychological variables and vaccination decisions after adjusting for demographic variables, socio-economic status and health conditions. The study subjects were sampled using a stratified random sampling technique and multivariable logistic regression model was used to analyze data. The results suggested that socio-psychological variables can explain vaccination behavior better than demographic, socio-economic and health variables. The authors gave a balanced discussion on the strengths and limitations of the study.

Major concerns:

18. As mentioned by the authors, the study was subject to reporting bias. The influence of the reporting bias to the study results should be discussed.

Thank you for the suggestion. We have now further discussed the possible influence of the reporting bias to the study results in Limitations (p. 27).

19. In page 8 line 12, the authors mentioned they dichotomized the continuous variables with missing values. This action is very subjective. The authors should justify why the dichotomization will not influence the results and statistical power of the analysis. The authors should also mention which variables were dichotomized and which were not, and discuss what was the potential consequence due to the dichotomization.

Thank you for your comment. We agree that dichotomising variables is not desirable, as it may lead to a reduction in the statistical power to detect an association between the variable and the outcome measure and underestimate the extent of variation in outcome between groups. We indeed mention dichotomisation as a limitation. However, other frequently used strategies such as excluding variables with missing or "I don't know/not applicable" responses or to fill in or "impute" missing values can lead to different kinds of bias. In our case, since "I do not know/not applicable" responses are not missing values as such (i.e. participants chose to answer "I do not know/not applicable"), we felt it was not appropriate to either exclude or impute these values. Dichotomisation was, therefore, considered an acceptable alternative.

We only dichotomised 4 continuous variables which had a significant number of “I do not know/not applicable” responses. We identify which variables were dichotomised in Tables S2-S4 in Supplementary material, as initially stated in Methods. These were: “Physician thinks I should vaccinate”, “Relatives think I should vaccinate”, “perceived knowledge of the vaccine (to make informed decisions)” and “the vaccine is painful”. The first two variables lend themselves to be dichotomised, as those who may influence vaccination decisions would normally be either supportive or unsupportive of their patient or relative being vaccinated (or their opinion is unknown, which we included in the “other than yes” category). We have provided an expanded justification in Limitations (p. 27).

20. After building model 1, the authors “manually removed one” insignificant variable “at a time”, (in page 8 line 34). The authors should describe what was the the sequence of removing variables and justify the criteria used to determine the sequence.

This part was not clear, as we realised in retrospect. We did not remove the variables in a sequence, as this procedure could have confounded the results. We developed models which included all the significant variables and all the non-significant variables except for one. We did this for each of the non-significant variables, one at a time. Hence, we built 12 different models for the US (each model excluded one of the 12 non-significant variables), 11 for the UK and 21 for France. Our aim was to examine the effect of the missing variable on the rest of the variables, one at a time. We have re-worded the procedure for clarity (it is now in supplementary material as suggested by reviewer 2) and have excluded the sentence “one at a time”, as it may be causing the confusion.

21. The authors claimed “We developed robust regression models” in page 3 line 7. This statement is overly strong and misleading. The authors did not develop statistical methods but they used “robustness” check to select variables when building the final model. It should be clarified.

Thank you for your suggestion. We have removed the terms “developed” and “robust” when referring to the models and only use “robust” when describing the robustness checks.

22. In page 8 line 38, please provide details about what was of the specific criteria of “robustness” based on the change of regression coefficients.

We have run different specifications to assess whether the significance of each variable was robust across specifications. The correlation of an independent variable with the dependent variable was considered robust if it remained statistically significant across specifications. We have made this explicit in Methods (Box S2 in Supplementary material) and also noted in Results that “The magnitude and significance of the relationship between independent and dependent variables varied little between the first models (M1) – where all the variables were entered at the same time – and the last models (M6-M8) – where a reduced number of variables were entered in blocks –, which is a further indication of the robustness of our findings.” (p. 19).

23. The authors assumed correlation coefficient between dependent and independent variables was 0.1 to calculate the sample size. What was the observed correlation coefficient based on the data? Please discuss.

Thank you for your comment. All the correlation coefficients between the dependent and independent variables were higher than 0.1, except for two variables which were tested for the first time in this study: “Bad experience vaccines – child” ($r = -0.082$, $p < 0.05$ in the UK; $r = 0.040$, $p > 0.05$ in the US; and $r = -0.064$, $p > 0.05$ in France) and “Scary health experience – child” ($r = 0.090$, $p < 0.05$ in the US). Given that, to our knowledge, this is the first time childhood experiences have been investigated quantitatively in the context of adult vaccination, further testing these results across different samples

would be desirable to ensure that the link (or lack thereof) between these variables and the outcome measure is a true one. We have referred to this point in Methods (p. 8), Results (p. 19) and Discussion (p. 24).

Reviewer: 4

This is an interesting survey evaluating policy factors associated with influenza vaccination. The article is well written and of interest to the general audience of the journal. Statistics are accurately described and presented. However, I have a few comments:

Thank you for the positive comments on our study.

24. Page 6, "Study Sample": Sample size calculation is rather superficial based on an arbitrary assumption. Authors should better justify this choice that merely looks as an effort to justify a convenience sample and also take into account the multi-regional setting. Although there is a wealth of evidence in this area, as highlighted in the introduction, the methods used are disparate and, as a result, findings are not consistent. For example, a meta-analysis by Brewer et al. (referenced throughout the paper), which illustrates this heterogeneity, showed that the effect sizes of the perceived severity, likelihood and susceptibility to influenza range from ($z = -0.18$ to $z = 0.42$). Moreover, we sought to test some factors which had not been previously studied (please see response to question 23). Therefore, when calculating the size of our sample, we used a conservative assumption (detailed in methods). We made an effort to attain samples that were representative of the population of interest (adults) and have provided more details about our sampling strategy in Methods (p. 6).

25. There is a wealth of p-value based statistical testing. Some type -I error adjustment would have been useful given that authors base their conclusions from unadjusted p-values which are marginal in many cases.

There is debate about the pertinence or benefit of using p-value based statistical testing in regression analysis, which can reduce power and increase the risk of type II error (Perneger, T.V. What's wrong with Bonferroni adjustments. *BMJ* 1998; 316 : 1236; Rothman, K.J. No adjustments are needed for multiple comparisons. *Epidemiology* 1990; Vol. 1, No. 1 : 43-46). Strategies to reduce spuriousness, which we employed in our analyses, are to include plausible and evidence-based variables, but also to ensure there are 20 or more subjects or observations per variable (Harrell, Frank. *Regression modeling strategies: with applications to linear models, logistic and ordinal regression, and survival analysis*. Springer, 2015)]. As a result, the overall p-value of our models was significant and the final models explained an important proportion of the variance in vaccination behaviour (64 to 80%), which are a good indication of the trueness of the relationship between independent and dependent variables. We discussed and debated this within the team with input from a number of statisticians and we would respectfully propose that we maintain the reporting of our data with the unadjusted p-values. We feel that given the nature of the study and data we report unadjusted values allow easier interpretation as well as comparison of the pattern that we obtained with other studies in the field. We would also suggest that we report the p-values as clearly as we possibly can (as we have now done in the Discussion), to reduce any chance of misinterpretation by readers – thus not making excessive interpretation of weak results (p. 22 and 24).

26. Given that this is a multi-regional study there are issues pertinent to questionnaire standardization to local populations/different languages. Authors must show some evidence of questionnaire standardization. Some internal consistency measures such as Cronbach's alpha must be given.

Thank you for your suggestion. We have performed internal consistency analyses of socio-

psychological scales in each country and our results show the internal consistency of the scales is comparable across countries, except for the “trust in vaccination stakeholders” scale in France. This result may be influenced by the ongoing controversy around the influenza vaccine, which has compromised trust in vaccination stakeholders and has had a lasting negative effect on vaccination rates. We have referred to these analyses in Methods (p. 8), Results (p. 19) and Discussion (p. 22).

VERSION 2 – REVIEW

REVIEWER	David Elliman Consultant in Community Child Health Public Health England/Great Ormond Street Hospital United Kingdom
REVIEW RETURNED	30-Mar-2017

GENERAL COMMENTS	<p>Thank you for giving me the opportunity to see this again. The authors have a difficult task. The main novel approach of this study is the statistical methodology and there are few new findings or recommendations. Most of the readers will be clinicians and I still feel that this is a difficult read for a non-researcher/non-statistician. I am not sure what the answer is to address this.</p> <p>One of the limitations the authors ascribe to previous studies asking people why they weren't vaccinated was to "report post decisional rationalisations of their behaviours". This is mentioned, though not by name, in the final paragraph and suggests further prospective research to explore this. As it is considered, quite rightly, a significant limitation and I imagine was identified before initiating the study, I am puzzled as to why the study was not designed with this in mind. Many people were probably not immunised because they didn't get around to it and would then, as the authors say, rationalise their action with a statement that is more an excuse than a reason. This limitation would seem important enough to mention in the summary.</p> <p>Factors that effect uptake of influenza vaccine will include perception of its effectiveness, as mentioned by the authors, and any scares, of which there have been around the vaccine, both in adults and children. These obviously cannot be controlled, but should be described as present or absent around the time of vaccination in the different countries and therefore a possible factor in the study. It could easily mean different factors were important in different years and so when the survey is performed in relation to this is important. This study was related to the 2013/4 season, so it would be useful for there to be mention of any (or no) relevant factors that might have related to that season in particular. This is mentioned for France but not the other two countries.</p> <p>I don't understand the sentence on page 6 lines 28-36. Historical immunisation is likely to be quite accurate in that it is memory of an event that only happened less than 6 months ago and later on the authors do mention there being good evidence of a high correlation with medical records..</p> <p>The response rate of the email survey used opened emails as the denominator. I don't understand why. If I receive an email that is not of interest to me from the title, I may well bin it without opening it. Does that not make me a non-responder? If my logic is correct, that means the true response rate may be much lower than that quoted. Why not use emails sent and not bounced? Would that be more equivalent to how other response rates are calculated?</p>
-------------------------	--

	<p>On page 70 line 31-34, is recorded the proportion of healthy adults under 65 receiving the vaccine. This of course is the proportion of respondents to the survey. The figure for UK seems high. Is there any way of validating this? As far as I am aware this data is not collected routinely, so it may not be possible, but it is important. The figure for France is 16% which is exactly the same as the figure for all under 65s in Table 1. Were there no adults under 65 with health problems immunised in France? Are the figures in the text incorrect or should the word "healthy" be removed?</p> <p>On page 71, it is stated that "Although the UK and France have healthcare systems which are free at the point of delivery or affordable for most, the influenza vaccine is only free of charge for high risk groups, which may explain the association between health insurance and past vaccination in both countries – albeit weak in France." This could be tested by analysing the data in UK by those under 65 v older adults. If health insurance is important, there should be no effect in 65s+ as, in UK, the vaccine is free to all of them.</p> <p>In table 1, there are no estimates for income in the general population yet in the notes to the table it is stated that inter alia data from income is from the US Census Bureau. Is the note incorrect or should there be data in the table. Likewise there must be figures for the UK population.</p>
--	--

REVIEWER	Jing Huang University of Pennsylvania
REVIEW RETURNED	03-Apr-2017

GENERAL COMMENTS	The revised manuscript has addressed all my previous concerns. I have no more comments.
-------------------------	---

REVIEWER	Eleni Verykoui Medical School, Aristotle University of Thessaloniki, Greece
REVIEW RETURNED	24-Mar-2017

GENERAL COMMENTS	The authors have adequately addressed all concerns and came up with a very well elaborated revision. I have no further comments and recommend publication.
-------------------------	--

VERSION 2 – AUTHOR RESPONSE

Reviewer 2

1. Thank you for giving me the opportunity to see this again. The authors have a difficult task. The main novel approach of this study is the statistical methodology and there are few new findings or recommendations. Most of the readers will be clinicians and I still feel that this is a difficult read for a non-researcher/non-statistician. I am not sure what the answer is to address this.

We appreciate your concern. Yet, in the first round of reviews the other three reviewers advised us quite strongly to provide even more technical details and analyses, which we have done to their satisfaction. You very rightly point out that the focus of our paper is to propose improved ways to measure and analyse data on vaccination uptake, building upon an important body of literature on the

subject – this is precisely what we are seeking to do; to advance our understanding and improve our response to concerns regarding vaccination.

In writing and revising the manuscript, we have endeavoured to both achieve our aims and improve and clarify our work in light of the expert commentary (including yours), while taking into account the perspectives and interests of BMJ Open's diverse readership. Given the purpose of our study and the word limit we must comply with, we have tried our best to strike a balance between methodological detail and implications for policy and practice. We would certainly be delighted to discuss this in live conversation with the editorial team and yourself, if this would provide a more suitable means to consider the nature of this paper further.

2. One of the limitations the authors ascribe to previous studies asking people why they weren't vaccinated was to "report post-decisional rationalisations of their behaviours". This is mentioned, though not by name, in the final paragraph and suggests further prospective research to explore this. As it is considered, quite rightly, a significant limitation and I imagine was identified before initiating the study, I am puzzled as to why the study was not designed with this in mind. Many people were probably not immunised because they didn't get around to it and would then, as the authors say, rationalise their action with a statement that is more an excuse than a reason. This limitation would seem important enough to mention in the summary.

Thank you for your suggestion. In retrospective research, there is always a risk of capturing post-decisional rationalisations, which we have tried to minimise as much as possible within the limitations of this research design. In the case of the influenza vaccine, however, given that it is administered periodically (annually), the only way to solve this problem is to target respondents who have never vaccinated before and to survey them twice: before the influenza season (to assess their pre-vaccination views) and after (to record their actual vaccination status). We do appreciate this methodology would have offered a more robust assessment of the determinants of vaccination, but we were faced, as many colleagues in this line of research, with a pragmatic concern. This approach requires considerable financial resources and time – particularly if an aim is to achieve a representative sample (as we tried, across 3 countries) – which, unfortunately, were not available to us. We have further explained this in Limitations and have specifically referred to the term "post-decisional rationalisations" to enhance understanding of this methodological aspect of our work.

3. Factors that effect uptake of influenza vaccine will include perception of its effectiveness, as mentioned by the authors, and any scares, of which there have been around the vaccine, both in adults and children. These obviously cannot be controlled, but should be described as present or absent around the time of vaccination in the different countries and therefore a possible factor in the study. It could easily mean different factors were important in different years and so when the survey is performed in relation to this is important. This study was related to the 2013/4 season, so it would be useful for there to be mention of any (or no) relevant factors that might have related to that season in particular. This is mentioned for France but not the other two countries.

Thank you for this comment. Based upon the available literature and news searches, to our knowledge, there were no major controversies around the influenza vaccine in the US and the UK during the 2013/4 season or the previous season, which is reflected on their fairly constant vaccination rates for more than a decade up to 2014. To ensure we are clear on this, we have added this information in the Discussion and have provided references about the US and UK historical vaccination rates.

4. I don't understand the sentence on page 6 lines 28-36. Historical immunisation is likely to be quite accurate in that it is memory of an event that only happened less than 6 months ago and later on the authors do mention there being good evidence of a high correlation with medical records.

We used the term “historical” prompted by one of your comments in the first round of reviews. In our earlier revision of the paper, we agreed with you that using “past” vaccination could be confused with the most recent season, and hence the term “historical” better described past vaccination seasons (not including the most recent). It is precisely because of the strong correlation between historical, intended (future) and recent vaccination, that we are critiquing research which includes proxies of recent vaccination (e.g. historical or future vaccination) as determinants of recent vaccination. Unsurprisingly, when proxies of recent vaccination are included in models as independent variables, they account for a large proportion of the variance in recent vaccination behaviour, yet this information does not explain what factors underpin successive vaccination, only that it is happening. We have tried to further explain this in the introduction, for clarity on this point.

5. The response rate of the email survey used opened emails as the denominator. I don't understand why. If I receive an email that is not of interest to me from the title, I may well bin it without opening it. Does that not make me a non-responder? If my logic is correct, that means the true response rate may be much lower than that quoted. Why not use emails sent and not bounced? Would that be more equivalent to how other response rates are calculated?

We felt that it was more informative to use opened e-mails as the denominator, an approach often used in market research, given that online surveys may remain unopened for a number of reasons, including email servers filtering valid survey invitations as spam. However, for comparability with other recruitment methods and simplicity, we have followed your suggestion and changed the formula used to calculate the response rates. Given that online survey methods and response rates benchmarks are continuously changing, we have also added a more recent reference on non-probability panels to contextualise the achieved response rates. Further, we have recently become aware of an emerging body of research from established scholars showing that higher response rates are not associated with more accuracy, and sometimes high response rates can yield less accurate results. We have included this information and added a relevant reference in Limitations.

6. On page 70 line 31-34, is recorded the proportion of healthy adults under 65 receiving the vaccine. This of course is the proportion of respondents to the survey. The figure for UK seems high. Is there any way of validating this? As far as I am aware this data is not collected routinely, so it may not be possible, but it is important. The figure for France is 16% which is exactly the same as the figure for all under 65s in Table 1. Were there no adults under 65 with health problems immunised in France? Are the figures in the text incorrect or should the word "healthy" be removed?

Regarding the first question, we agree that the proportion of under 65s who vaccinated against influenza is higher than we expected. Unfortunately, to our knowledge, there is no available vaccination coverage data among this group which can be used as a benchmark. These are both important findings.

Thank you for the second comment; we have changed the phrasing to “both with and without eligible chronic conditions”, as our intention was to refer to both healthy and eligible under 65s across the three countries.

7. On page 71, it is stated that "Although the UK and France have healthcare systems which are free at the point of delivery or affordable for most, the influenza vaccine is only free of charge for high risk groups, which may explain the association between health insurance and past vaccination in both countries – albeit weak in France." This could be tested by analysing the data in UK by those under 65 v older adults. If health insurance is important, there should be no effect in 65s+ as, in UK, the vaccine is free to all of them.

As mentioned in the last round of reviews, we certainly agree that it would be desirable to better understand what motivates (or deters) under and over 65s to vaccinate, and to compare the results. However, since the focus of this research was the general adult population (and not a particular sub-group), our sample size precludes us from performing reliable sub-group analyses.

8. In table 1, there are no estimates for income in the general population yet in the notes to the table it is stated that inter alia data from income is from the US Census Bureau. Is the note incorrect or should there be data in the table. Likewise there must be figures for the UK population.

Thank you for giving us the opportunity to explain this further. To facilitate survey completion and improve data accuracy, we used a limited number of income bands relevant to each country to collect household income data. This is because bias and error due to lack of knowledge, recall problems, sensitivity, etc., are common in income reporting (Moore JC, Welniak EJ. Income measurement error in surveys: A review. Journal of Official Statistics 2000 1;16(4):331), thus limiting the number of income categories participants are presented with can reduce some of these sources of inaccuracy. Given that countries calculate median household income using more granular data (e.g. the US uses more than 40 bands), we could not directly compare the median income of our samples with that of each of the populations. Therefore, we only report the proportion of people earning below and above the band that was closest to the population median household income and state as a note that “The reference income band was the closest to the US/UK/France’s median household income”. For clarity, we have now provided an expanded explanation in the Participants section of Results and have also included the potential difference (given that do we do not know the exact difference) between our samples and the corresponding populations’ income in Limitations.

VERSION 3 – REVIEW

REVIEWER	David Elliman Great Ormond Street Hospital, London, UK
REVIEW RETURNED	28-May-2017

GENERAL COMMENTS	Most of the issues raised have been addressed.
-------------------------	--