PEER REVIEW HISTORY

ARTICLE DETAILS

TITLE (PROVISIONAL)	Randomised controlled non-inferiority trial of primary care-based facilitated access to an alcohol reduction website: cost effectiveness analysis
AUTHORS	Hunter, Rachael; Wallace, Paul; Struzzo, Pierliugio; Della Vedova, Roberto; Scafuri, Francesca; Tersar, Costanza; Lygidakis, Charilaos; Mc Gregor, Richard; Scafato, Emanuele; Freemantle, Nick

VERSION 1 – REVIEW

REVIEWER	Dominic Conroy Birkbeck, University of London U.K.
REVIEW RETURNED	16-Oct-2016

GENERAL COMMENTS	Thank you for the opportunity to review this paper. The research described in this paper concerns an RCT intended to assess whether a self-directed access to a hazardous drinking online resource holds health and cost benefits relative to conventional face to face practitioner input. This was an interesting paper, and the research described seems valuable and of international relevance. I suggest that a statistician with a background in health economics should be approached to provide expert opinion in reviewing this paper, if such an opinion has not already been sought be the editorial team. Expertise aside, I have some suggestions below which should help strengthen the paper. To a large extent these concern the manuscripts written style, which could sometimes be clearer/more concise. - The term 'facilitated access' could be better defined at some point in the paper; I am still not entirely clear what this branch of the intervention involved. Specifically, I am not clear what it was that was doing the facilitating- is the facilitator the online resources or does a practitioner introduce them to the online resources? I assume the term has a technical meaning that will be more relevant to other readers. That said, if possible, I would consider alternative phrasing for this intervention group to ensure clearer meaning. - There are some typos on p.2 in the Interventions part of the abstract ("… or to to access via their GP…") - I would condense down the content in the results subsection of the abstract on p.2; there is some unnecessary content and repetition in this section. In this same subsection I would consider an alternative word for 'dominated' to ensure clear meaning and include some more concrete details in this sentence (e.g. what was the lower cost).

	- P4, L27 onwards Some rewriting of the middle paragraph would be good- for instance could refer to 'practitioners' rather than refer to 'GP and other clinical staff'. - P5 Final paragraph referring to how UK data is being transposed to an Italian setting. I was not entirely clear how/why this was being done- could this paragraph be re-written so as to be clearer and more brief? - P6, L12- could PPSRU be dropped or spelt out? - P6, L33-P7, L9- the description of QALYs here could (as with other places) be much more clearly and briefly outlined. - P8, L5- the description of your analytic approach could be outlined more clearly here so that your reader understands what your analytic approach was and how it should be interpreted. I was unclear what "were used to populate" meant in this sentence, for example. - P8, L18- sentence beginning "No discount rate…". I did not understand your intended meaning in this sentence and suggest it is re-written and also combined with an adjacent paragraph. - P8 sentence starting "Information on the duration…". As these details have appeared in your methods section I would omit these here or present them in an abbreviated fashion. - P9. The four paragraphs here describing the cost-related findings were very difficult to follow. It is hard to get a sense of the key 'headline' detail from the rather note like written style of this section, while the short paragraphing gives a rather fragmented feel to the page. I suggest rewriting this section aiming to cast things in a much clearer narrative form- explain to the reader what is going on in bold strokes. - P9. I suggest summarizing Table 1 in the text as there is not really enough complex information here to warrant a stand-alone table. - P10, L17. I suggest that the term 'intention-to-treat' would be the technically accurate term to use in place of 'available cases' in this section. - P11, L14-23- I would break this sentence up into 2-3 sentences. This is because it was very long but also because the meaning could be made clearer.

REVIEWER	Mark Deady UNSW, Australia
REVIEW RETURNED	28-Oct-2016

GENERAL COMMENTS	The current paper presents the cost-effectiveness of a randomised 1:1 non-inferiority trial comparing facilitated access to a website for hazardous drinkers with a standard face-to-face brief intervention (BI). I credit the authors with their efforts in this important area. Overall I feel the paper is generally well-written although there are areas in which at least elucidation is needed. This issue is primarily found regarding the outcome measures and the resulting cost savings. I have a number of concerns with the primary and secondary outcomes (which I thought could benefit from clearer delineation). Firstly, the use of the UK NHS etc for most of the costing is of major concern, although the authors speak briefly about this in the limitations I think more thought needs to go into the precise differences between the INHS and NHS that may affect generalisation.

	For instance the cost per visit is so strikingly different that the assumption that the services provided are for all relevant intents and purposes the same (eg time commitment etc) is somewhat alarming. Again the lack of Italian norms regarding the QALY outcome is relevant. The use of the AUDIT is somewhat unclear. It appears to be used at 3 and 12months but the method section suggests it may only have occurred at 3- and 6-months and have been extrapolated to 12-months. The Wallace paper is constantly cited but without access to this paper, this reviewer was confused by this section regarding cases of hazardous /harmful drinking prevented. Furthermore, as the scale is largely based on 12-month responding it is unclear how interpretation of these data can really occur (ie the respondent has given 12month data (dating back to preintervention) at both 3- and 6-months). Additionally I found the reporting of savings on page 9 confusing and could be better articulated for clarity. (also is webhosting etc incorporating in the cost of the facilitated intervention?). The benefits stated per 1000 patients are also confusing, for instance the non-inferiority seems to hold but is discussed as superiority of facilitated care. Eg the 2 QALYs gained is not reflected in this section of the results rather that the two are not significantly different. Furthermore, the 158 additional hazardous drinkers mentioned is not adequately discussed in the discussion. Why was this case? Is this due largely to item 10 on the AUDIT? Finally the first 2 paragraphs of the discussion require proof reading and clarification as a number of sentences could be more clearly described. There are also attempts made to unfairly stretch the results to fit a British population, which is something of a concern throughout the paper. Although the generalisability of the results is certainly a moot point. The authors' should refrain from overstatement of the findings and interpretation in a British context.

REVIEWER	Borko Jovanovic Northwestern U USA
REVIEW RETURNED	09-Dec-2016

GENERAL COMMENTS	I don't like this paper, it seems to piggyback on the previous paper and address an issue I do not see as important. In addition, I don't think "equivalence" is established properly. Also, the result is really not surprising, and to some extent is irrelevant - doctors are more expensive that software - we know that. Not sure what this all is about. Sorry to be so blunt!

REVIEWER	Christopher K. Haddock National Development and Research Institutes, Inc., USA No competing interests
REVIEW RETURNED	06-Feb-2017

GENERAL COMMENTS	I focused on the methods and statistical analysis for this article in my review. The author's research focuses on an important topic and their Facilitated Intervention undoubtedly holds promise. However, there are methodological limitations which should be addressed. (1) It is not clear why this was conceptualized as a non-inferiority trial. According to the rationale provided by the authors the purpose was: "to evaluate the short term cost savings and potential additional short term benefits to the INHS of facilitated access to a website for hazardous and harmful drinking compared to a standard face-to-face brief intervention (BI) over 12 months. Although a non-inferiority trial can be analyzed as a superiority trial, if this description is accurate the study was originally conceptualized as a superiority trial and no non-inferiority analysis is offered in the paper. In fact, the authors assumed non-inferiority. (2) If there was a 1:1 allocation of participants to groups why were the N's so different (i.e., 347 vs. 416, or a 69 person imbalance)? There are ways to randomly allocate participants to groups equally. (3) The authors assume that the average appointment length for an INHS GP is 11 minutes. No support is provided for this pivotal assumption. Studies in the USA of primary care office visits have varied but are typically over 15 minutes (e.g., https://www.ncbi.nlm.nih.gov/pmc/articles/PMC2254573/). If the actual average length is 15 minutes, the cost per minute is decreased by 27% from that assumed in this paper. This would significant impact cost analyses. (4) The authors use reference 13 to support their claim that the average GP office visit in the UK is 11 minutes. However, there is a newer version of that annual report (http://www.pssru.ac.uk/project-pages/unit-costs/2015/) which appears to suggest that the average clinic visit for a General Practitioner is 17.2 minutes (Table 10.8a, p. 176). (5) QALY models were primarily based on complete case analysis (i.e., cases with no missing data), which could bias the results of the comparison. Typically, the primary model in randomized trials is an intention-to-treat analysis with careful consideration of missing data (e.g., multiple imputation). There is substantial loss-to-follow up in the study (17.9% in the Facilitated Access group, 19.5% in the Face-to-Face group) which highlights the problem with a complete case analysis. In Table 3 the authors report the "Available Case" analysis which used imputation and the QALYs (adjusted) is 0.0006 (versus 0.002 for the complete case analysis). This is a substantial difference (not "marginally smaller" as it is described) in outcomes. (6) The authors present the results as if Facilitated access had better outcomes than the Face-to-Face intervention although results of their analysis suggest they do not have evidence for this claim. Merely including the caveat "but not significantly so" obscures the actual study outcomes. The 95% CI for QALYs for both the complete case and (more appropriate) multiple imputation analyses are wide (suggesting imprecision in the estimates) and include zero. Given the very small point estimates and insignificant results of the hypothesis tests the only conclusion is that there were no differences in outcomes between the two groups. The language in the paper should not imply otherwise.

	(7) Table 1 is titled “duration of appointments”. However, according to the methods this is actually the duration of the face-to-face intervention, which may be less than the time required for the appointment. Even given the authors conservative estimate of 11 minutes per appointment by GPs, the fact that the time reported for 56.3% of those in face-to-face intervention was less than 5 minutes suggests the appointment duration was longer than the intervention duration. (8) The time required for GPs for the Facilitated Intervention group appears to have been assessed less precisely than for the Face-to-Face group. Did you use the same 3-category assessment items (i.e., < 5 min, 5-10 minutes, > 10 minutes) for the Facilitated Group as you did the Face-to-Face condition? If so, did all mark the < 5-minute category? Did you ask the physicians about all 347 cases in the Facilitated intervention group and compute an average time or was the 2-minute assumption a guess? You assume that all 347 of the encounters for the Face-to-Face group where the GP marked “< 5 minutes” took exactly 5 minutes (i.e., the most conservative interpretation possible) but assume that for the Facilitated Group it was 2 minutes. The entire results section for Costs appears to be based on guesses which are biased toward the Facilitated Access condition.

REVIEWER	Melissa Tracy Department of Epidemiology and Biostatistics University at Albany School of Public Health State University of New York United States
REVIEW RETURNED	15-Feb-2017

GENERAL COMMENTS	This paper reports a cost-effectiveness analysis comparing costs per minute, costs per patient, and overall costs for face-to-face brief intervention vs. facilitated access to a website for hazardous drinkers. The organization of the Methods section could be a bit more streamlined, with information further clarified. The section on “costs” states that the average appointment length was assumed to be 11 minutes (line 56, page 5) with costs per minute then estimated. This assumption is restated in the “hypothesis testing” section (lines 30-32, page 8) but I think it would be better to restate the average cost per minute that is being assumed, since the average duration of the brief interventions vs. facilitated access are estimated in the study based on information reported by the GPs (pp. 8-9). Furthermore, it would be helpful to discuss how individuals who did not receive the brief intervention were handled in cost analyses – are they assigned a time of zero minutes in the lower-bound analysis? It would also be helpful for readers to clarify how utility scores are calculated in the discussion of QALYs in the Methods sections (p. 7) so it is no surprise when getting to the Results section. I was unsure about the sample sizes reported in the complete case analysis (Table 2, as they don’t exactly match the 12-month sample sizes reported in Figure 1). I would also clarify the table titles and descriptions (i.e., “difference in health utility” implies a comparison of facilitated access vs. brief intervention, so positive values indicate greater utility for facilitated access)

	The CEP and CEAC analyses are appropriate, but perhaps this information should be included last to parallel the order of the Methods section. In describing the analysis of benefits per 1000 patients referred (p. 11), it is not immediately clear that the reported odds ratio of 0.94 refers to the odds of harmful or hazardous drinking in the facilitated access group vs. the face-to-face group so this should be clarified or the percentage of both group groups who met criteria for harmful or hazardous drinking should be reported, as in the following paragraph. Ultimately, the conclusions regarding the cost effectiveness of facilitated access vs. face-to-face brief intervention rest heavily on the assumptions made in analyses, and, although this is acknowledged in the article discussion, this point could probably be made more forcefully.

REVIEWER	Swarna Weerasinghe Dalhousie University, Dept. Community Health and Epidemiology, Halifax, Nova Scotia, Canada
REVIEW RETURNED	15-Feb-2017

GENERAL COMMENTS	This study addresses an important area in community health. However, the study needs major revisions as follows. (1) Include a literature review on the topic and the rationale for the study (2) Include a section on data analysis by clearly identifying what variables are used in the regression and what type of variables. Given this is a randomized trails your intervention should be the major predictor. What are the other predictors that you adjusted for. What type of regression model did you fit? This section is very important. (3) Results need to include beta coefficients of regression and the confidence intervals. You need to include means and confidence intervals for the intervention (web based) and the other group. (4) Discussion needs to be strengthen by comparing with the outcomes of research already conducted elsewhere. (5) Given above limitations conclusions are vague and too strong. (6) This paper needs a major rewrite and the team should get the help of a statistician. Thorough literature review should be done.

VERSION 1 – AUTHOR RESPONSE

Reviewer 1

Comment: The term ‘facilitated access’ could be better defined at some point in the paper; I am still not entirely clear what this branch of the intervention involved. Specifically, I am not clear what it was that was doing the facilitating- is the facilitator the online resources or does a practitioner introduce them to the online resources? I assume the term has a technical meaning that will be more relevant to other readers. That said, if possible, I would consider alternative phrasing for this intervention group to ensure clearer meaning.

“Facilitated access” is short hand for “facilitated access to an interactive website for reducing hazardous and harmful drinking” as described in the first sentence in the Methods. It is the GP that facilitates access by referring patients to the website by handing them a leaflet and suggesting that they access the website Down Your Drink website. We have amended the description of this throughout the document to make it clearer.

There are some typos on p.2 in the Interventions part of the abstract (“… or to to access via their GP…”)

Response: Thank you for bringing this to our attention. This has now been amended.

Comment: I would condense down the content in the results subsection of the abstract on p.2; there is some unnecessary content and repetition in this section. In this same subsection I would consider an alternative word for ‘dominated’ to ensure clear meaning and include some more concrete details in this sentence (e.g. what was the lower cost).

Response: We have removed the last sentence in the results section of the abstract following this advice from the reviewer and given this is repeated in the first sentence of the conclusion.

Comment: P4, L27 onwards Some rewriting of the middle paragraph would be good- for instance could refer to ‘practitioners’ rather than refer to ‘GP and other clinical staff’.

Response: This paragraph has now been amended.

Comment: P5 Final paragraph referring to how UK data is being transposed to an Italian setting. I was not entirely clear how/why this was being done- could this paragraph be re-written so as to be clearer and more brief?

Response: This paragraph has now been rewritten as requested.

Comment: P6, L12- could PPSRU be dropped or spelt out?

Response: This reference has now been removed.

Comment: P6, L33-P7, L9- the description of QALYs here could (as with other places) be much more clearly and briefly outlined.

Response: This section has now been rewritten.

Comment: P8, L5- the description of your analytic approach could be outlined more clearly here so that your reader understands what your analytic approach was and how it should be interpreted. I was unclear what “were used to populate” meant in this sentence, for example.

Response: The wording in this sentence has been changed and a reference to the standard methodology has been included.

Comment: P8, L18- sentence beginning “No discount rate…”. I did not understand your intended meaning in this sentence and suggest it is re-written and also combined with an adjacent paragraph.

Response: This is information that needs to be reported as set out in the CHEERS statement “Report the choice of discount rate used.” This is standard terminology for when no discount rate is used.

Comment; P8 sentence starting “Information on the duration…”. As these details have appeared in your methods section I would omit these here or present them in an abbreviated fashion.

Response: This has now been amended.

Comment: P9. The four paragraphs here describing the cost-related findings were very difficult to follow. It is hard to get a sense of the key ‘headline’ detail from the rather note like written style of this section, while the short paragraphing gives a rather fragmented feel to the page. I suggest rewriting this section aiming to cast things in a much clearer narrative form- explain to the reader what is going on in bold strokes.

Response: The text has now been amended to be clearer.

Comment: P9. I suggest summarizing Table 1 in the text as there is not really enough complex information here to warrant a stand-alone table.

Response: This has now been incorporated into the text.

Comment: P10, L17. I suggest that the term ‘intention-to-treat’ would be the technically accurate term to use in place of ‘available cases’ in this section.

Response: The term intention-to-treat implies that the data chosen were based on trial randomisation. This is equally the case in the complete case and available case analyses. Only the multiple imputation analysis includes all patients in the trial. Instead the “available case” term here is used to contrast the results with the complete case analysis. We have renamed the available case analysis “incomplete case analysis” to make this point clearer.

Comment: P11, L14-23- I would break this sentence up into 2-3 sentences. This is because it was very long but also because the meaning could be made clearer.

Response: This has now been broken up based on the reviewers recommendation.

Reviewer 2

Comment: The current paper presents the cost-effectiveness of a randomised 1:1 non-inferiority trial comparing facilitated access to a website for hazardous drinkers with a standard face-to-face brief intervention (BI). I credit the authors with their efforts in this important area. Overall I feel the paper is generally well-written although there are areas in which at least elucidation is needed. This issue is primarily found regarding the outcome measures and the resulting cost savings.
I have a number of concerns with the primary and secondary outcomes (which I thought could benefit from clearer delineation). Firstly, the use of the UK NHS etc for most of the costing is of major concern, although the authors speak briefly about this in the limitations I think more thought needs to go into the precise differences between the INHS and NHS that may affect generalisation. For instance the cost per visit is so strikingly different that the assumption that the services provided are for all relevant intents and purposes the same (eg time commitment etc) is somewhat alarming.

Response: Italian NHS costs were used for the costing. This has now been made clearer in the Methods section. English NHS was only used for inflation uplift of costs and any assumptions about the duration of an appointment to calculate the cost per minute. The total INHS cost though was used to generate the cost-effectiveness acceptability curve so allow for the uncertainty in that value to be modelled.

Comment: Again the lack of Italian norms regarding the QALY outcome is relevant.

Response: There are no Italian values for the EQ-5D-5L and the UK values are the most relevant. We have noted this issue further in the limitations and cited research that has compared UK and Italian EQ-5D-3L tariff scores to give some understanding of the implications of not having an Italian specific tariff.

Comment: The use of the AUDIT is somewhat unclear. It appears to be used at 3 and 12months but the method section suggests it may only have occurred at 3- and 6-months and have been extrapolated to 12-months.

Response: Apologies – there was a typo on page 7 where 6 months was entered instead of 12 months. The AUDIT was administered at 3 months and 12 months. This has now been amended to the correct value.

Comment: The Wallace paper is constantly cited but without access to this paper, this reviewer was confused by this section regarding cases of hazardous /harmful drinking prevented. Furthermore, as the scale is largely based on 12-month responding it is unclear how interpretation of these data can really occur (ie the respondent has given 12month data (dating back to preintervention) at both 3- and 6-months).

Response: We agree that this paper may require a rewrite if not published alongside the Wallace et al paper. Given that no decision has been taken on this paper yet we will only undertake this rewrite following a steer from the Editor.

Comment: Additionally I found the reporting of savings on page 9 confusing and could be better articulated for clarity. (also is webhosting etc incorporating in the cost of the facilitated intervention?). The benefits stated per 1000 patients are also confusing, for instance the non-inferiority seems to hold but is discussed as superiority of facilitated care. Eg the 2 QALYs gained is not reflected in this section of the results rather that the two are not significantly different. Furthermore, the 158 additional hazardous drinkers mentioned is not adequately discussed in the discussion. Why was this case? Is this due largely to item 10 on the AUDIT?

Response: The QALYs have now been removed given the confusion this caused. The discussion of the issues with the AUDIT would be clearer if published alongside the companion paper. Again this may need to be rewritten dependent on the status of the other paper and we would appreciate a steer from the Editor on this.

Comment: Finally the first 2 paragraphs of the discussion require proof reading and clarification as a number of sentences could be more clearly described. There are also attempts made to unfairly stretch the results to fit a British population, which is something of a concern throughout the paper. Although the generalisability of the results is certainly a moot point. The authors' should refrain from overstatement of the findings and interpretation in a British context.

Response: We appreciate the reviewers comments on this and hence now toned down the findings in England to better reflect the data.

Reviewer 3

Comment: I don't like this paper, it seems to piggyback on the previous paper and address an issue I do not see as important. In addition, I don't think "equivalence" is established properly. Also, the result is really not surprising, and to some extent is irrelevant - doctors are more expensive that software - we know that. Not sure what this all is about. Sorry to be so blunt!

Response: Although we appreciate the authors candour, and recognise that this paper does require the context of its sister paper, the issues are still important. Even if we accept that doctors are less expensive than software (although this is not a given considering the additional costs of website development and maintenance) the issue is then if they are less expensive, is this at potential loss of clinical effectiveness? This paper argues that the potential for an increase in the implementation of interventions for hazardous drinking using facilitated access for websites is likely due to the reduced time it takes the GP, and at no loss of quality of care.

Reviewer 4

Comment 1: It is not clear why this was conceptualized as a non-inferiority trial. According to the rationale
provided by the authors the purpose was: "to evaluate the short term cost savings and potential additional short term benefits to the INHS of facilitated access to a website for hazardous and harmful drinking compared to a standard face-to-face brief intervention (BI) over 12 months.
Although a non-inferiority trial can be analyzed as a superiority trial, if this description is accurate the study was originally conceptualized as a superiority trial and no non-inferiority analysis is offered in the paper. In fact, the authors assumed non-inferiority.

Response: The aim of the trial as indicated in the trial protocol attached to the submission was non-inferiority. This relates to (i) the primary outcome of AUDIT-10, which is the outcome used to determine the sample size; (ii) the EQ-5D as used to calculate QALYs. Our hypothesis regarding costs is that facilitated access would be cost saving given that it takes less GP time. It is perfectly reasonable that a trial assume non-inferiority in the benefits but that it is cost-saving, particularly for interventions that are justified on the basis of being cheaper but resulting in no worse outcome to patients. We have amended the aim reported at the end of the Introduction to make this clearer.

Comment 2: If there was a 1:1 allocation of participants to groups why were the N's so different (i.e., 347 vs. 416, or a 69 person imbalance)? There are ways to randomly allocate participants to groups equally.

Response: This issue is addressed in the main paper for the trial:
"Randomisation led to a chance imbalance between the numbers of subjects in the two groups, but checks at several points during the trial confirmed that the imbalance was not due to a programming error and we were able to confirm that the software was operating correctly."

Comment: The authors assume that the average appointment length for an INHS GP is 11 minutes. No support is provided for this pivotal assumption. Studies in the USA of primary care office visits have varied but are typically over 15 minutes (e.g.,
https://www.ncbi.nlm.nih.gov/pmc/articles/PMC2254573/). If the actual
average length is 15 minutes, the cost per minute is decreased by 27% from that assumed in this paper.
This would significant impact cost analyses.

Response: Since this paper was submitted for peer review additional evidence for the duration of GP appointments has been published. Primarily, a study carried out by Hobbs et al. (2016) of 398 English general practices (101.8 million consultations) between April 2007 and March 2014. They found that the average duration of an appointment was 9 minutes. We have updated the results to reflect this new finding. The assumptions made about the duration of appointments though has had limited impact on the main findings.

Comment: The authors use reference 13 to support their claim that the average GP office visit in the UK is 11
minutes. However, there is a newer version of that annual report
(http://www.pssru.ac.uk/projectpages/
unit-costs/2015/) which appears to suggest that the average clinic visit for a General Practitioner is 17.2 minutes (Table 10.8a, p. 176).

Response: Please see response above that there is now more up to date evidence. To note also that the correct value to use is the surgery time of 11.7 minutes in the document referenced by the reviewer, not the clinic time.

Comment 5: QALY models were primarily based on complete case analysis (i.e., cases with no missing data), which could bias the results of the comparison. Typically, the primary model in randomized trials is an intention-to-treat analysis with careful consideration of missing data (e.g., multiple imputation). There is substantial loss-to-follow up in the study (17.9% in the Facilitated Access group, 19.5% in the Face-to- Face group) which highlights the problem with a complete case analysis. In Table 3 the authors report the "Available Case" analysis which used imputation and the QALYs (adjusted) is 0.0006 (versus 0.002 for the complete case analysis). This is a substantial difference (not "marginally smaller" as it is described) in outcomes.

Response: Our statistical and health economic analysis plan pre specified that the primary analysis of both the primary outcome and the primary analysis of the health economics data would be a complete case analysis and that we would assume that the mechanism of missing data was at random. A secondary analysis using multiple imputation to account for missing data was specified. We have removed the term “marginally” and included more detail on the implications of the multiple imputation analysis.

Comment: 6 The authors present the results as if Facilitated access had better outcomes than the Face-to-Face intervention although results of their analysis suggest they do not have evidence for this claim. Merely including the caveat “but not significantly so” obscures the actual study outcomes. The 95% CI for QALYs for both the complete case and (more appropriate) multiple imputation analyses are wide (suggesting imprecision in the estimates) and include zero. Given the very small point estimates and insignificant results of the hypothesis tests the only conclusion is that there were no differences in outcomes between the two groups. The language in the paper should not imply otherwise.

Response: We have removed the one reference to the facilitated access group resulting in better outcomes, which occurs in the conclusion of the abstract.

Comment 7: Table 1 is titled “duration of appointments”. However, according to the methods this is actually the duration of the face-to-face intervention, which may be less than the time required for the appointment. Even given the authors conservative estimate of 11 minutes per appointment by GPs, the fact that the time reported for 56.3% of those in face-to-face intervention was less than 5 minutes suggests the appointment duration was longer than the intervention duration.

Response: This table has now been removed (see response to reviewer 1). This is correct that face-to-face BIs are commonly less than a total appointment length given that the need to be fit in with the rest of a GP appointment.

Comment 8: The time required for GPs for the Facilitated Intervention group appears to have been assessed less precisely than for the Face-to-Face group. Did you use the same 3-category assessment items (i.e., < 5 min, 5-10 minutes, > 10 minutes) for the Facilitated Group as you did the Face-to-Face condition? If so, did all mark the < 5-minute category? Did you ask the physicians about all 347 cases in the Facilitated intervention group and compute an average time or was the 2-minute assumption a guess? You assume that all 347 of the encounters for the Face-to-Face group where the GP marked “< 5 minutes” took exactly 5 minutes (i.e., the most conservative interpretation possible) but assume that for the Facilitated Group it was 2 minutes. The entire results section for Costs appears to be based on guesses which are biased toward the Facilitated Access condition.

Response: GPs were asked a general question about the duration of facilitated access, but not specifically for each appointment. We appreciate the reviewers point though that we have assumed that the “<5 minutes” category took 5 minutes in face-to-face group and hence have increased the value in the facilitated access group to 5 minutes to fit with this assumption. None of the GPs reported facilitated access taking longer than 5 minutes.

Reviewer 5

Comment: The organization of the Methods section could be a bit more streamlined, with information further clarified. The section on “costs” states that the average appointment length was assumed to be 11 minutes (line 56, page 5) with costs per minute then estimated. This assumption is restated in the “hypothesis testing” section (lines 30-32, page 8) but I think it would be better to restate the average cost per minute that is being assumed, since the average duration of the brief interventions vs. facilitated access are estimated in the study based on information reported by the GPs (pp. 8-9).

Response: The cost per minute is not used in the hypothesis test and hence it is not mentioned. Instead we report the number of additional appointments as a result of facilitated access compared to face-to-face interventions.

Comment: Furthermore, it would be helpful to discuss how individuals who did not receive the brief intervention were handled in cost analyses – are they assigned a time of zero minutes in the lower-bound analysis?

Response: It’s not clear what additional information this would provide the reader. Patients that did not receive the intervention are included as a 0 if they are included and removed if they are not included. It’s not clear how else this should be handled.

Comment: It would also be helpful for readers to clarify how utility scores are calculated in the discussion of QALYs in the Methods sections (p. 7) so it is no surprise when getting to the Results section.

Response: This has now been made clearer.

Comment: I was unsure about the sample sizes reported in the complete case analysis (Table 2, as they don’t exactly match the 12-month sample sizes reported in Figure 1).

Response: These will differ as for complete cases this will only be for individuals with complete data at all 3 time points.

Comment: I would also clarify the table titles and descriptions (i.e., “difference in health utility” implies a comparison of facilitated access vs. brief intervention, so positive values indicate greater utility for facilitated access)

Response: Additional clarity has been added.

Comment: The CEP and CEAC analyses are appropriate, but perhaps this information should be included last to parallel the order of the Methods section.

Response: The methods section and results do mirror each other. The results per 1000 patients are last in the methods and the results.

Comment: In describing the analysis of benefits per 1000 patients referred (p. 11), it is not immediately clear that the reported odds ratio of 0.94 refers to the odds of harmful or hazardous drinking in the facilitated access group vs. the face-to-face group so this should be clarified or the percentage of both group groups who met criteria for harmful or hazardous drinking should be reported, as in the following paragraph.

Response: This section may need to be re-written dependent on the editorial decision regarding the companion paper. Again we would appreciate an Editorial steer on the best way to proceed with this.

Comment: Ultimately, the conclusions regarding the cost effectiveness of facilitated access vs. face-to-face brief intervention rest heavily on the assumptions made in analyses, and, although this is acknowledged in the article discussion, this point could probably be made more forcefully.

Response: We have added additional information to make this clearer.

Reviewer 6

below This study addresses an important area in community health. However, the study needs major revisions as follows.

(1) Include a literature review on the topic and the rationale for the study

Response: A literature review on the topic was not the intended aim of this study. Instead we have drawn attention to evidence available in this area.

(2) Include a section on data analysis by clearly identifying what variables are used in the regression and what type of variables. Given this is a randomized trails your intervention should be the major predictor. What are the other predictors that you adjusted for. What type of regression model did you fit? This section is very important.

Response: This has now been made clearer. In addition to the randomisation variable the only other variable adjusted for in the analysis is baseline EQ-5D-5L utility score.

(3) Results need to include beta coefficients of regression and the confidence intervals. You need to include means and confidence intervals for the intervention (web based) and the other group.

Response: The beta coefficients are reported in what is now Table 2. Means are reported in Table 1 and now include confidence intervals.

(4) Discussion needs to be strengthen by comparing with the outcomes of research already conducted elsewhere.

Response: The only research we are aware of that has been conducted elsewhere has been discussed in the discussion. There is limited research conducted in this field and hence the limited number of studies to draw on.

(5) Given above limitations conclusions are vague and too strong.

(6) This paper needs a major rewrite and the team should get the help of a statistician. Thorough literature review should be done.

We respectfully disagree that this paper requires additional statistical assistance. The last author on this paper is an internationally renowned Chair in Epidemiology and Biostatistics who assisted with the design of the study, the interpretation of the results and write up of the paper.
As discussed, the aim of this paper was not a thorough literature review and as mentioned above we have drawn on the limited literature available in this area. As a result it is not clear why a literature review is required.

VERSION 2 – REVIEW

REVIEWER	Mark Deady UNSW, Australia
REVIEW RETURNED	25-May-2017

GENERAL COMMENTS	my notes are in essence for the editor and not the author, all my comments have been responded to but the issue remains regarding the Wallace paper that is constantly cited but without access to this paper, some aspects of this paper are unclear. I shall defer to the editor in this regard.

REVIEWER	Christopher Haddock National Development & Research Institutes
REVIEW RETURNED	30-May-2017

GENERAL COMMENTS	The authors were modestly responsive to the reviewer feedback. I agree with Reviewer 5 that the paper could use a biostatistical collaborator, despite the response to the contrary. Several aspects of the study suggest that it a statistician was needed, such as the highly unbalanced sample sizes per group despite the plan for 1:1 randomization. The response to this critique suggests all of the authors were unaware how balance in sample size is accomplished using block randomization techniques. Also, because the RCT protocol suggest this was a noninferiority trial doesn't mean it was analyzed as one. Unfortunately, many of the issues with the study cannot be corrected now that it is completed.

REVIEWER	Melissa Tracy Assistant Professor Department of Epidemiology and Biostatistics University at Albany School of Public Health State University of New York United States
REVIEW RETURNED	30-May-2017

GENERAL COMMENTS	The revisions the authors made in the text and table titles have addressed many of my previous concerns with respect to clarity of the Methods and Results, particularly regarding the comparisons being made. I would note, however, that the revisions have also introduced some confusion into the paper, so further editing is required. For example, while the text now states that the difference in time between facilitated access and face-to-face intervention would allow for an additional GP appointment for every *three* patients referred to the website (p. 9), the abstract still states that time for an additional appointment would be made available for every *two* patients

	referred. I was also confused by the cost assumptions listed for the most and least conservative scenarios: in particular, wouldn't the most conservative scenario include the higher estimated cost per patient of the face-to-face intervention (€11.10 rather than €10.16), or is the goal of the most conservative scenario to minimize the difference in costs between the two groups? In any case, the total estimated cost of €64 reported for the face-to-face intervention in the least conservative scenario (p. 9) doesn't make sense given a training cost of €60 and time cost of €10-11 per patient. It's also not clear where the estimated probability of 84% for cost-effectiveness (listed in the first sentence of the Discussion on p. 12) is coming from since it wasn't mentioned in the revised results section nor is it apparent from Figure 2.

VERSION 2 – AUTHOR RESPONSE

Reviewer 2

My notes are in essence for the editor and not the author, all my comments have been responded to but the issue remains regarding the Wallace paper that is constantly cited but without access to this paper, some aspects of this paper are unclear. I shall defer to the editor in this regard.

Thank you. This has been addressed by the Editor above.

Reviewer 5

Comment: The revisions the authors made in the text and table titles have addressed many of my previous concerns with respect to clarity of the Methods and Results, particularly regarding the comparisons being made.
I would note, however, that the revisions have also introduced some confusion into the paper, so further editing is required. For example, while the text now states that the difference in time between facilitated access and face-to-face intervention would allow for an additional GP appointment for every *three* patients referred to the website (p. 9), the abstract still states that time for an additional appointment would be made available for every *two* patients referred.

Response: Thank you for identifying the error in the abstract. We have now updated the abstract to the correct figure of "three", same as the text on page 9.

Comment: I was also confused by the cost assumptions listed for the most and least conservative scenarios: in particular, wouldn't the most conservative scenario include the higher estimated cost per patient of the face-to-face intervention (€11.10 rather than €10.16), or is the goal of the most conservative scenario to minimize the difference in costs between the two groups?

Response: As correctly identified by the reviewer, the goal of the most conservative analysis is to minimise the cost differences between the two groups. Specifically the aim of this analysis is to test what the lowest possible value of the cost-savings associated with facilitated access are. If we were to choose the higher value for face-to-face access this would be acting in favour of facilitated access and would increase the cost-savings.
We have now made this clearer in the paper by adding "lowest possible cost difference" and "highest possible cost difference" in brackets after "most conservative analysis" and "least conservative analysis" respectively

Comment: In any case, the total estimated cost of €64 reported for the face-to-face intervention in the least conservative scenario (p. 9) doesn't make sense given a training cost of €60 and time cost of €10-11 per patient.

Response: Thank you for identifying this error. This has now been updated to the correct value of €71 for the cost of training and face-to-face for a total cost difference of €58 (€71-€13). In light of the reviewer's comments we have revised this section and double checked all of our figures.

Comment: It's also not clear where the estimated probability of 84% for cost-effectiveness (listed in the first sentence of the Discussion on p. 12) is coming from since it wasn't mentioned in the revised results section nor is it apparent from Figure 2.

Response: This is for INHS costs including training only (no website costs). We have included the text "84% probability if only the cost of training (excluding website development costs) are included" on page 11 in the CEAC section of the results to make this clearer.

Reviewer 4

Comment: The authors were modestly responsive to the reviewer feedback. I agree with Reviewer 5 that the paper could use a biostatistical collaborator, despite the response to the contrary. Several aspects of the study suggest that it a statistician was needed, such as the highly unbalanced sample sizes per group despite the plan for 1:1 randomization. The response to this critique suggests all of the authors were unaware how balance in sample size is accomplished using block randomization techniques.

Response: The referee is correct to identify that blocking could avoid the kind of imbalance that we have between the randomised conditions (which occurred by chance using a simple, and extensively tested, randomisation algorithm). However the referee appears unaware that such blocking processes are a source of bias in open trials (while appropriately used in blinded ones). This is because future allocations are at least in part predictive based on past ones. The statistical inefficiency of the imbalance is very modest, and the approach is optimally unbiased.

Comment; Also, because the RCT protocol suggest this was a noninferiority trial doesn't mean it was analyzed as one.

Response: The trial has been analysed as a noninferiority trial as specified in the protocol.

Comment: Unfortunately, many of the issues with the study cannot be corrected now that it is completed.

VERSION 3 – REVIEW

REVIEWER	Melissa Tracy Dept of Epidemiology & Biostatistics University at Albany School of Public Health State University of New York United States
REVIEW RETURNED	01-Sep-2017

GENERAL COMMENTS	I am satisfied with the authors' revisions.