PEER REVIEW HISTORY

ARTICLE DETAILS

TITLE (PROVISIONAL)	The effectiveness of general practice-based health checks on health behaviour and incidence on non-communicable diseases in individuals with low socioeconomic position: a randomised controlled trial in Denmark
AUTHORS	Kamstrup-Larsen, Nina; Dalton, Susanne; Grønbæk, Morten; Broholm-Jørgensen, Marie; Thomsen, Janus Laust; Larsen, Lars; Johansen, Christoffer; Tolstrup, Janne

VERSION 1 - REVIEW

REVIEWER	David Blanco Universitat Politècnica de Catalunya
REVIEW RETURNED	08-Feb-2019

GENERAL COMMENTS	This report shows the results of an evaluation of the consistency between the CONSORT checklist you submitted and the information that was reported in the manuscript. Please, make the following revisions: • For CONSORT Item 6a ("Completely defined pre-specified primary and secondary outcome measures, including how and when they were assessed"), please include in the paragraph about secondary outcomes details on how self-efficacy and perceived stress were measured as you did with the other secondary outcomes. • For CONSORT Item 9a (“Mechanism used to implement the random allocation sequence, describing any steps taken to conceal the sequence until interventions were assigned”), please explain how the allocation system was set up so that the person enrolling participants did not know in advance which treatment the next person was going to get. • For CONSORT Item 11a (“If done, who was blinded after assignment to interventions (for example, participants, care providers, those assessing outcomes) and how") please report the blinding status of the GPs participating in the study. • For CONSORT Item 13a (“For each group, losses and exclusions after randomisation, together with reasons”), please make sure in the CONSORT flow diagram that the numbers in the “Allocated to Check-in” box are adequate since 364+2+186 is not equal to 549. If the numbers are correct, please clarify the possible confusion. Moreover, please provide the number of care providers performing the intervention in each group and the number of patients treated by each care provider. For more details on this, please see Item

	13a of the CONSORT extension for non-pharmacological interventions (https://annals.org/aim/fullarticle/2633220/consort-statement-randomized-trials-nonpharmacologic-treatments-2017-update-consort-extension). • For CONSORT Item 17a ("For each primary and secondary outcome, results for each group, and the estimated effect size and its precision (such as 95% confidence interval)"), please report the results at 12-month follow up for the secondary outcome "self-efficacy" since this is listed as one of the secondary outcomes of the study but not mentioned in the "Results" section. Furthermore, in Table 3, please include for each outcome a measure of the effect size and its precision, as you did in Table 2.

REVIEWER	Dr Richard Lowrie NHS Greater Glasgow and Clyde /University of Glasgow Department of General Practice and Primary Care Scotland, UK
REVIEW RETURNED	15-Feb-2019

GENERAL COMMENTS	I think this is a well conducted study on a group of patients who are difficult to recruit in a pragmatic trial of this nature. The authors are to be commended for their rigorous approach and for describing this study and presenting it in a clear and meaningful way. I have a few points / questions that can hopefully be addressed in a revision, which I would be happy to review in due course. Most of these are around clarification of features of the methods, that will help inform other research in this area: 1. Patients. Can the authors elaborate on the validity and appropriateness, or evidence of "no schooling beyond lower secondary school" being a measure of low socioeconomic position? 2. Practices. Can the authors revise the consort diagram to include numbers of practices recruited and followed through at each stage of the trial? I think this may be stipulated in the consort recommendations for cluster RCTs. Also, for ease of understanding by readers. 3. Flow diagram would benefit from being annotated with date ranges corresponding to each phase of the trial. 4. Power calculation. Needing 150 smokers in each arm to detect 15% difference is based on assumption that half are motivated to quit. This seems a critical assumption but by design there is no way of determining whether the 50% rule applied. Can the authors discuss this in limitations section? 5. Follow up. 10% a-priori loss to follow up also appears conservative (I understand this is difficult, given the target group) when the loss was actually around 24% (from flow chart). Can the authors discuss this point. 6. Intervention. I am not clear why the "Invitation" phase constitutes part of the intervention. Note the text of this section seems to be a repeat of the text in the previous section 'Recruitment and participants'. 7. Intervention. Can the authors describe how many of the assessments were conducted by GPs and how many by other practice staff? Much is said about the importance of GP-patient relationship, and knowing the proportions of patients seen by GPs

	in context of this trial would illuminate the importance of this or otherwise. 8. Intervention. Is it not a major weakness of the intervention, that some patients were referred to another place for their follow up consultation, given the reluctance of patients from lower socioeconomic groups to present to healthcare providers when asymptomatic? Would it be worth mentioning that if there was, for example, home visits, of continuity of care with same GP, then the effect of the intervention might have been greater? 9. Ethics. If there was no need for ethical approval for practice and patient participation, would it not have been theoretically possible to follow the outcomes of non responders through remote data linkage? Minor points: would the title read better if "on" was replaced with "of"?. Page 8 line 21 change "age" to "ago" and page 15 line 13 change "illustrate" to "illustrates".

REVIEWER	Martin Gulliford King's College London UK
REVIEW RETURNED	15-Feb-2019

GENERAL COMMENTS	This is a well reported account of a trial of health checks in Denmark. The authors have given a clear account of the study. the design and analysis of the study are generally appropriate. the lack of a theoretically motivated and well standardised intervention is a weakness that is discussed. Per protocol and ITT analyses are reported, the latter being based on multiple imputation. There are two stages of attrition - attendance for the check and return of follow-up data. The report would have benefited from presenting additional information about differential attrition of smokers during follow-up. It would also be better to present the per protocol and ITT analyses in separate tables, as in Table 2 it is not clear why 94 is 31% of the intervention trial arm and 147 is 36% of the control trial arm. Specific comments In the Abstract, Background: where it reads 'health checks aimed at certain groups at high risk may reduce adverse health behaviour and identify persons with metabolic risk factors and non-communicable diseases (NCDs).' It appears that the conclusion of the study is being stated as the introduction. In the Introduction, it may be worth commenting on the occurrence of multiple risk factors as these are strongly associated with mortality (BMJ 2008;337:a1440) but difficult to modify (https://bmjopen.bmj.com/content/7/6/e015375). Methods page 5 line 40: please give details of the response rate to the questionnaire about education. Page 6, line 20: give details of allocation concealment (not blinding). Page 6, line 21; where it refers to eligible patients, it is not clear what the eligibility criteria were. were patients already treated for

	risk factors or chronic diseases excluded? Presumably they did not need a health check? Page 7 line 1 where it says 'Before the health check the GPs received results from the patient-reported questionnaire in the GPs electronic patient record in the form of an electronic data interchange (EDI) message including summed scores and categorization of items from the baseline questionnaire (see supplementary).' It is not clear what questionnaire and scores are referred to. Up to now we have only been told that participants were sent a questionnaire about their educational attainment. Explain whether the check was free and whether conducted during work hours. With regard to the imputation, were the data missing at random? would a last observation carried forward approach have been more reasonable? Page 12, based on the ASA recommendations on the use of P values the term statistically significant should not be used. https://amstat.tandfonline.com/doi/abs/10.1080/00031305.2016.1154108#.XGbTUej7SUk Interpretation should be based on the estimated effects not the P values, the latter only gauging the strength of evidence. It is not clear initially whether the random intercepts model was the primary analysis or whether that was only done later as a sensitivity analysis. Table 3: the 'median regression' requires a footnote of explanation. Figure 1: it is not made clear why 425 responders gives 303 in a per protocol analysis in the intervention group and 422 and 407 in the control group. It would be beneficial to include a table comparing the characteristics of those lost to follow-up, with those not lost, by trial arm. Presumably smokers were more likely to be lost, sorry if I missed that info. It would have been beneficial to have had a standard intervention package such as prescription of nicotine replacement or failing that some process evaluation information about what interventions were offered. The intervention does not appear to be well theorised. Why did we think that the participants would change their behaviours based on what was offered? It would be beneficial to address the TIDIER checklist as well as CONSORT.

REVIEWER	Rie Goto Department of Archaeology and Anthropology University of Cambridge United Kingdom (left in 2018)
REVIEW RETURNED	16-Feb-2019

GENERAL COMMENTS	1) The authors should explain the definition and means of 'check-in' in the text. The medical system is different by countries. 2) The authors should discuss the possible bias of the check-in based study design. 3) The intervention was based on the health consultation provided by GPs following the standard medical practices and suggested to additional health check after 6 months. Is this 'intervention' or 'general medical practices'? How difference of the practices between intervention and control groups and how the intervention strongly designed to effect on health behaviour? Please explain more details. 4) The authors used only two variables to identify the low socioeconomic position - no formal education beyond secondary school and cohabitation status (no details how low cohabitation status was justified). It needs to explain and justify why the two variables were used. 5) Smoking status was described as 5 different status and also re-categorised into two 'daily' and 'not-daily' smokers, but Table 1 showed 3 categories (daily smoker, occasional and ex-smoker and never smoker). Not clear. 6) Table 4 – prevalence of depression were significantly higher in check-in (5%) than usual care group (2%). It is due to the small sample size with a large range of 95% CI. Please discuss it more carefully. 7) The title should include the location of the study.

VERSION 1 – AUTHOR RESPONSE

Reviewer #1:

Reviewer Name: David Blanco

Institution and Country: Universitat Politècnica de Catalunya

Please state any competing interests or state 'None declared': None declared

This report shows the results of an evaluation of the consistency between the CONSORT checklist you submitted and the information that was reported in the manuscript.

Please, make the following revisions:

For CONSORT Item 6a ("Completely defined pre-specified primary and secondary outcome measures, including how and when they were assessed"), please include in the paragraph about secondary outcomes details on how self-efficacy and perceived stress were measured as you did with the other secondary outcomes.

Thank you for commenting on our paper. Following is included to the Method and material section:

“Stress during the past month was assessed by the perceived stress scale (PSS) (score range 0-40) (Cohen, 1983). The person’s belief in their innate ability to achieve goals was assessed using general self-efficacy (score range 10-40) (Bandura 1977;Mikkelsen 1999).”

For CONSORT Item 9a (“Mechanism used to implement the random allocation sequence, describing any steps taken to conceal the sequence until interventions were assigned”), please explain how the allocation system was set up so that the person enrolling participants did not know in advance which treatment the next person was going to get.

The randomization in Check-In was conducted in the statistical programme SAS by a data manager at the National Institute of Public Health. The general practitioners were not part of the randomisation process. The phrase about the randomisation is rewritten to clarify this and no state:

“Eligible patients were randomised in SAS by a data manager at the National Institute of Public Health to either Check-In or usual care in a 1:1 allocation. The randomisation was stratified by gender and 5-year age group. Couples living together were allocated to the same group to avoid contamination.”

For CONSORT Item 11a (“If done, who was blinded after assignment to interventions (for example, participants, care providers, those assessing outcomes) and how") please report the blinding status of the GPs participating in the study.

To meet this comment, we added the following to the Method and material section:

“Double-blinded, meaning that both patients and GPs were blinded to the allocation of group, would have been ideal (Rothman 2008); nevertheless, due to real-life setting, blinding of participants was only possible in the control group and not among GPs.”

For CONSORT Item 13a (“For each group, losses and exclusions after randomisation, together with reasons”), please make sure in the CONSORT flow diagram that the numbers in the “Allocated to Check-in” box are adequate since 364+2+186 is not equal to 549. If the numbers are correct, please clarify the possible confusion. Moreover, please provide the number of care providers performing the intervention in each group and the number of patients treated by each care provider. For more details on this, please see Item 13a of the CONSORT extension for non-pharmacological interventions (https://annals.org/aim/fullarticle/2633220/consort-statement-randomized-trials-nonpharmacologic-treatments-2017-update-consort-extension).

The numbers in the CONSORT flow diagram is corrected.

The randomisation was conducted at individual level meaning that the number of GPs were the same in the two groups – the 32 general practice clinics.

Furthermore, the range of individuals per general practice is indicated in the result section where it now state the following:

“Of the 8,508 (49%) who responded to the baseline questionnaire, 1,104 met the inclusion criteria regarding level of education and marked that they could be contacted again (range per general practice clinic: 12-110 individuals; median=18).”

For CONSORT Item 17a (“For each primary and secondary outcome, results for each group, and the estimated effect size and its precision (such as 95% confidence interval)”), please report the results at 12-month follow up for the secondary outcome “self-efficacy” since this is listed as one of the secondary outcomes of the study but not mentioned in the “Results” section. Furthermore, in Table 3, please include for each outcome a measure of the effect size and its precision, as you did in Table 2.

Results for self-efficacy at follow-up are included in the result section which now state the following: “Further, no difference between the two groups were found regarding self-efficacy where both groups had a median at 29 (IQR for ‘Check-In’ 25,33; IQR for usual care 24,34) (data not shown).”

Furthermore, we added a measure of the effect size and its precision for each of the outcomes in Table 3.

Reviewer #2:

Reviewer Name: Dr Richard Lowrie

Institution and Country: NHS Greater Glasgow and Clyde /University of Glasgow Department of General Practice and Primary Care, Scotland, UK

Please state any competing interests or state ‘None declared’: None declared

I think this is a well conducted study on a group of patients who are difficult to recruit in a pragmatic trial of this nature. The authors are to be commended for their rigorous approach and for describing this study and presenting it in a clear and meaningful way.

I have a few points / questions that can hopefully be addressed in a revision, which I would be happy to review in due course. Most of these are around clarification of features of the methods, that will help inform other research in this area:

1. Patients. Can the authors elaborate on the validity and appropriateness, or evidence of "no schooling beyond lower secondary school" being a measure of low socioeconomic position?

Thank you for the encouraging comments on our paper. To meet the first comment the following is included to the background section:

“Short education was used as measure for low SEP as educational level captures the influence of resources on health and the knowledge and skills attained through education may affect an individual’s cognitive functioning, make individuals more receptive to health education messages, or more able to communicate with and access appropriate health services (Galobardes, 2006)”

2. Practices. Can the authors revise the consort diagram to include numbers of practices recruited and followed through at each stage of the trial? I think this may be stipulated in the consort recommendations for cluster RCTs. Also, for ease of understanding by readers.

The recruitment of general practitioner clinics is included in the CONSORT diagram.

3. Flow diagram would benefit from being annotated with date ranges corresponding to each phase of the trial.

The Check-In RCT was conducted in five rounds which make it difficult to indicate the date ranges for each of the phases. Dates for enrolment phase (January 2014-September 2015) and the follow-up phase January 2015-September 2016 is included in the flow diagram. Furthermore, it is explained in the Method and material section that the trial had five rounds:

“The recruitment of the GPs was, however, challenged due to a break down in the collective bargaining between the Danish Regions Salary and Rate Board (RTLN) and the Organisation of General Practitioners (PLO) in late 2012 (Nexøe, 2013). In all, Check-In ended up having five rounds between January 2014 to September 2016.”

4. Power calculation. Needing 150 smokers in each arm to detect 15% difference is based on assumption that half are motivated to quit. This seems a critical assumption but by design there is no way of determining whether the 50% rule applied. Can the authors discuss this in limitations section?

This is a good and important point which now is discussed in the limitation section where the following is stated:

“Furthermore, our sample calculations were based on several assumptions which can be discussed. The assumption that half are motivated to quit smoking can seem high and cannot be verified in the design. This assumption is, however, supported by the literature where 63% of daily smokers in Denmark with no education beyond lower secondary school are found to be motivated to quit smoking (40). Moreover, the 10% a-priori loss to follow-up was conservative when compared to the fact that the actually loss was 24%. Nevertheless, in total, 228 and 225 daily smokers were enrolled in the ‘Check-In’ intervention and control group, respectively, which exceeded the sample size calculations that indicated that we needed 150 daily smokers in each group. This indicates that despite a higher loss to follow-up than expected the sample was most likely large enough to detect had there been any effect of ‘Check-In’ regarding adverse health behaviour.”

5. Follow up. 10% a-priori loss to follow up also appears conservative (I understand this is difficult, given the target group) when the loss was actually around 24% (from flow chart). Can the authors discuss this point.

See respond to the comment above.

6. Intervention. I am not clear why the "Invitation" phase constitutes part of the intervention. Note the text of this section seems to be a repeat of the text in the previous section 'Recruitment and participants'.

The invitation was part of the intervention as this proactive approach from the GPs was belief to have significant meaning for the reach of the intervention. To clarify the possible confusing, we have changed the heading of the ‘Recruitment and participants’ section, which describes the identification of the study population, to “Identification of the study population”. This should make the distinction between the two aspects clearer.

7. Intervention. Can the authors describe how many of the assessments were conducted by GPs and how many by other practice staff? Much is said about the importance of GP-patient relationship, and knowing the proportions of patients seen by GPs in context of this trial would illuminate the importance of this or otherwise.

The GP-patient relationship is of great significant and this was why it was decided that in Check-In the health consultation after the health check had to be performed by the GP. On the other hand, it was also important that Check-In could be implemented in a busy working day in general practice, so it was prioritised that the health check could be conducted “by either the GP or other health staff at the clinic as per usual clinical practice”. Unfortunately, we do not know the actually proportion of health checks conducted by the GPs, but all the health consultations were conducted by the GPs.

8. Intervention. Is it not a major weakness of the intervention, that some patients were referred to another place for their follow up consultation, given the reluctance of patients from lower socioeconomic groups to present to healthcare providers when asymptomatic? Would it be worth mentioning that if there was, for example, home visits, of continuity of care with same GP, then the effect of the intervention might have been greater?

This is a good point and, in the following, I will clarify the possible confusing.

All the follow up consultations were conducted at the patient’s own GP. At the follow up consultation the GP could refer the patients to behaviour-change programmes at the municipalities if necessary, however, the patient could also “follow the medical standards for general practice on procedures for diagnostics and treatment”.

9. Ethics. If there was no need for ethical approval for practice and patient participation, would it not have been theoretically possible to follow the outcomes of non responders through remote data linkage?

This is a good point and in the following I will clarify why we chose not to do so. Theoretically, this would have been possible, however, only for the outcomes regarding the metabolic risk factors and NCDs and not for the patient-reported health behaviour, thus we did not to do so.

Minor points: would the title read better if "on" was replaced with "of"?. Page 8 line 21 change "age" to "ago" and page 15 line 13 change "illustrate" to "illustrates".

This is corrected – except for the title.

Reviewer #3:

Reviewer Name: Martin Gulliford

Institution and Country: King's College London UK

Please state any competing interests or state ‘None declared’: None

Please leave your comments for the authors below

This is a well reported account of a trial of health checks in Denmark. The authors have given a clear account of the study. The design and analysis of the study are generally appropriate. The lack of a theoretically motivated and well standardised intervention is a weakness that is discussed. Per protocol and ITT analyses are reported, the latter being based on multiple imputation. There are two stages of attrition - attendance for the check and return of follow-up data. The report would have benefited from presenting additional information about differential attrition of smokers during follow-up. It would also be better to present the per protocol and ITT analyses in separate tables, as in Table 2 it is not clear why 94 is 31% of the intervention trial arm and 147 is 36% of the control trial arm.

Specific comments:

In the Abstract, Background: where it reads 'health checks aimed at certain groups at high risk may reduce adverse health behaviour and identify persons with metabolic risk factors and non-communicable diseases (NCDs).' It appears that the conclusion of the study is being stated as the introduction.

We thank for the thorough reading and valuable comments. To meet the first comment, we have changed the sentence in the abstract to the following:

“However, it is not clear whether health checks aimed at certain groups at high risk may reduce adverse health behaviour and identify persons with metabolic risk factors and non-communicable diseases (NCDs).”

In the Introduction, it may be worth commenting on the occurrence of multiple risk factors as these are strongly associated with mortality (BMJ 2008;337:a1440) but difficult to modify (https://bmjopen.bmj.com/content/7/6/e015375).

This is a good point and the following is added to the introduction section:

“Furthermore, the occurrence of multiple of these adverse health behaviours are strongly associated with mortality (Van Dam, 2008) but are difficult to modify (Alageel, 2017).”

Methods page 5 line 40: please give details of the response rate to the questionnaire about education.

The details of the response rate to the questionnaire are included in the result section in the paragraph named “Participant flow” where it is stated:

“Of the 8,508 (49%) who responded to the baseline questionnaire, 1,104 met the inclusion criteria regarding level of education and marked that they could be contacted again.”

Page 6, line 20: give details of allocation concealment (not blinding).

The paragraph about allocation concealment is modified to clarify the possible confusing and now state the following:

“Double-blinded, meaning that both patients and GPs were blinded to the allocation of group, would have been ideal (Rothman, 2008); nevertheless, due to real-life setting, blinding of participants was only possible in the control group and not in the intervention group and among GPs.”

Page 6, line 21; where it refers to eligible patients, it is not clear what the eligibility criteria were. were patients already treated for risk factors or chronic diseases excluded? Presumably they did not need a health check?

To meet this comment and clarify for possible confusing we have rewritten the information about eligible patients which now state the following:

“Eligible patients met the inclusion criteria which were no formal education beyond lower secondary school and consent to be contacted for research purpose. No exclusion criteria were implied.”

Page 7 line 1 where it says 'Before the health check the GPs received results from the patient-reported questionnaire in the GPs electronic patient record in the form of an electronic data interchange (EDI) message including summed scores and categorization of items from the baseline questionnaire (see supplementary).' It is not clear what questionnaire and scores are referred to. Up to now we have only been told that participants were sent a questionnaire about their educational attainment. Explain whether the check was free and whether conducted during work hours.

This is a good point and information about items included in the questionnaire is added to the section describing the baseline questionnaire which now state the following:

“Therefore, to identify the study population baseline questionnaires in Danish (including items about sex, date of birth, cohabitation status, highest educational level achieved, height and weight, smoking status, alcohol consumption, physical activity, diet, general self-efficacy, perceived stress and family disposition of NCDs) were sent out to all individuals aged 45-64 years, who lived in Copenhagen and who were on the participating GPs’ patient lists.”

Further, in the section describing the health check it is added that the health checks were free of charged and that these were conducted during the opening hour of the general practice clinic:

“The health check was free of charge and took place during the opening hour of the general practice clinic to which the patient was registered and was conducted by either the GP or other health staff at the clinic as per usual clinical practice.”

With regard to the imputation, were the data missing at random? would a last observation carried forward approach have been more reasonable?

It is impossible to test whether the missing at random condition is satisfied (Allison PD, 2001; Missing Data).

To account for this relevant baseline characteristics were used in the imputation and we believe that this was the best way to handle missing data in this study.

Page 12, based on the ASA recommendations on the use of P values the term statistically significant should not be used.
https://amstat.tandfonline.com/doi/abs/10.1080/00031305.2016.1154108#.XGbTUej7SUk

Interpretation should be based on the estimated effects not the P values, the latter only gauging the strength of evidence.

This is changed throughout the paper. For instance, for estimates regarding health behaviour is now stated like the following:

“After 12 months of follow-up no difference was found between the Check-In and usual care group on daily smoking (ITT: OR=0.99; 95% CI:0.58-1.09), binge drinking (ITT: OR=0.82; 95% CI:0.59-1.14), physical inactivity (ITT: OR=0.97; 95% CI:0.74-1.27) or obesity (ITT: OR=0.90; 95% CI:0.67-1.21) (Table 2)”

It is not clear initially whether the random intercepts model was the primary analysis or whether that was only done later as a sensitivity analysis.

As stated at page 11 the random intercepts model was only done later as a sensitivity analysis:

“To evaluate the stability of our results an Interclass Coefficient (ICC) was estimated within a two-level model with patients (level 1) nested within general practices (level 2), and all estimates were calculated in the model including the condition variable and general practices as random intercept, allowing for correlation between patients from the same general practice. Furthermore, sensitivity analysis including age and sex in the logistic regression were carried out.”

Table 3: the 'median regression' requires a footnote of explanation.

The meet this comment we added the following explanation as a footnote to table 3:

“Median regression estimates the median of the dependent variable.”

Figure 1: it is not made clear why 425 responders gives 303 in a per protocol analysis in the intervention group and 422 and 407 in the control group.

This is a good point and to clarify the possible confusing we included a footnote to Figure 1 in which the numbers are explained:

“In the per protocol analyses are only included individuals who responded the questionnaire and who followed the ‘treatment’ for the allocated group (for individuals allocated to intervention this meant attending the health check and responding to the questionnaire; for individuals allocated to usual care this meant responding to the questionnaire). Hence, of the 425 responders in the intervention group 303 individuals attended the health check and could be included in the per protocol analyse. Of the 422 responders in the usual care group 407 answered the questions regarding the smoking status and could be included in the per protocol analyse.”

It would be beneficial to include a table comparing the characteristics of those lost to follow-up, with those not lost, by trial arm. Presumably smokers were more likely to be lost, sorry if I missed that info.

This is a good point and a such table is included as supplementary. Further, the following is included to the result section:

“In addition, the baseline characteristics for those lost to follow-up and those not lost to follow-up were comparable, however, the proportion of daily smokers and physical inactive were higher among those lost to follow-up compared to those not lost to follow-up (see supplementary).”

It would have been beneficial to have had a standard intervention package such as prescription of nicotine replacement or failing that some process evaluation information about what interventions were offered.

With the Check-In RCT, we aimed to explore the realistic potential for preventive health checks in general practice. It was important that Check-In could be implemented in a busy working day in general practice.

The intervention does not appear to be well theorised. Why did we think that the participants would change their behaviours based on what was offered? It would be beneficial to address the TIDIER checklist as well as CONSORT.

To meet this comment the following is included to the introduction section:

“It was developed in response to health-behaviour models in which increased awareness about the causes, consequences and cures for a particular health behaviour or health problem is expected to increase the likelihood for change (Prochaska, 1983) and in which knowledge is expected to lead to action (Champion, 2008). A preventive health check at the GP has the potential to confront the patient with a problem and provide feedback about both adverse health behaviour and the consequences of continuing the injurious behaviour. For example, poor lung function measure can demonstrate the health consequences of smoking and lead to a discussion about the adverse effects of smoking which may increase the chance for smoking cessation.”

Reviewer #4:

Reviewer Name: Rie Goto

Institution and Country: Department of Archaeology and Anthropology, University of Cambridge, United Kingdom (left in 2018)

Please state any competing interests or state ‘None declared’: None declared

Please leave your comments for the authors below

1) The authors should explain the definition and means of ‘check-in’ in the text. The medical system is different by countries.

Thank you for commenting on our paper.

‘Check-In’ was the name of the intervention. To clarify the possible confusing the name of the intervention is now written ‘Check-In’ throughout the paper.

2) The authors should discuss the possible bias of the check-in based study design.

It is important to address possible bias in the study which are done in the discussion section. This is for instance done with the following points in the discussion:

“The idea in ‘Check-In’ was that patients with adverse health behaviour amenable to intervention at the health check should be offered a referral to the municipality health centre for a free lifestyle change program. However, project data indicated that the opportunity of a referral may have been under-utilized as some of the patients rejected a referral to the municipality, and in some cases, the GP considered a referral to be irrelevant. The result was a low level of intensity of the part of the intervention targeting adverse health behaviour.”

and

“A potential limitation in the study was contamination between groups, which potentially occurred if patients in the usual care group had treatment beyond usual care.”

and

“The lack of effectiveness of ‘Check-In’ regarding more new hospital contacts and prescription medication for metabolic risk factors and NCDs can be ascribed to the fact that more than 60% of individuals included in the study were known with one or more NCDs at baseline. Most had visited their GP within the last year with a median number of contacts to the GP of 7 and 8 in the ‘Check-In’ and usual care group respectively (Table 1). Patients with a known NCD may, as such most likely, already be in some kind of scheduled treatment at their GP.”

Furthermore, the following is included in the discussion section:

“Even so, we cannot completely rule out that the effectiveness regarding depression was due to chance because of the small sample size in ‘Check-In’.”

3) The intervention was based on the health consultation provided by GPs following the standard medical practices and suggested to additional health check after 6 months. Is this ‘intervention’ or ‘general medical practices’? How difference of the practices between intervention and control groups and how the intervention strongly designed to effect on health behaviour? Please explain more details.

This is a good point and important to discuss. The intervention was designed so if it was effective it could easily be implemented in a busy working day in general practice. It is possible that the dose of intervention was too low and too similar to the control group to achieve sufficient change of adverse health behaviour which was also discussed in the discussion section where the following is stated:

“However, the intensity of the intervention might have been too low to achieve sufficient change of adverse health behaviour among individuals with low socioeconomic position, which may have contributed to the lack of measurable behavioural change in ‘Check-In’.”

4) The authors used only two variables to identify the low socioeconomic position - no formal education beyond secondary school and cohabitation status (no details how low cohabitation status was justified). It needs to explain and justify why the two variables were used.

The point about using educational level as a proxy for with the information about educational level for the non-respondents were also mentioned by the first reviewer. To meet this comment, we included the following to the background section:

“Short education was used as measure for low SEP as educational level captures the influence of resources on health and the knowledge and skills attained through education may affect an individual’s cognitive functioning, make individuals more receptive to health education messages, or more able to communicate with and access appropriate health services (Galobardes B (part 1). 2006)”

5) Smoking status was described as 5 different status and also re-categorised into two ‘daily’ and ‘not-daily’ smokers, but Table 1 showed 3 categories (daily smoker, occasional and ex-smoker and never smoker). Not clear.

This is a good point and the categories are changed in Table 1 which now only include ‘daily smoker’ and ‘not daily smoker’.

6) Table 4 – prevalence of depression were significantly higher in check-in (5%) than usual care group (2%). It is due to the small sample size with a large range of 95% CI. Please discuss it more carefully.

To meet this comment, we added the following to the discussion section:

“Even so, we cannot completely rule out that the effectiveness regarding depression was due to chance because of the small sample size in Check-In.”

7) The title should include the location of the study.

This was also commented by the editor and the title now states the following:

“The effectiveness of general practice-based health checks on health behaviour and incidence on non-communicable diseases in individuals with low socioeconomic position: a randomized controlled trial in Denmark”

VERSION 2 – REVIEW

REVIEWER	Martin Gulliford King's College London
REVIEW RETURNED	19-Aug-2019

GENERAL COMMENTS	Thank you, I believe the authors have given a satisfactory response to my review comments.