

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)	Reducing unnecessary vitamin testing in general practice: barriers and facilitators according to general practitioners and patients.
AUTHORS	Hofstede, Hetty; van der Burg, Rosalie; Mulder, Bob; Bohnen, Arthur; Bindels, Patrick; de Wit, Niek; de Schepper, E; van Vugt, Saskia

VERSION 1 - REVIEW

REVIEWER	Enrique Rodríguez Borja Laboratory of Clinical Biochemistry and Molecular Pathology. Hospital Clinico Universitario. Valencia (Spain)
REVIEW RETURNED	16-Feb-2019

GENERAL COMMENTS	<p>The authors state that there is little evidence strategies to reduce vitamin overtesting in general practice (Abstract, line 6). I suggest to the paper (Clin Chem Lab Med. 2018 Jun 27;56(7):1063-1070. doi: 10.1515/ccim-2017-1031.), where they show that "Send & hold" clinical decision supporting rules in electronic request are a promising informatics tool that can support in vitamins utilization management.</p> <p>Some of the conclusions of author's manuscript have been described before in previous works (barriers and facilitators). I recommend to review Salinas M & Lopez-Garrigos M et al. works regarding this subject. In fact, there are numerous references in the literature to educational interventions (verbal and written initiatives) as an strategy for test management. Unfortunately these strategies require exhaustive monitoring on the part of the laboratory since its effects decrease over time (short sustainability). I would recommend the authors to address this issue in their manuscript including proper references.</p> <p>The manuscript is quite interesting not only because we know the reason argued by GPs regarding overtesting but the patient's as well. REVERT study seems very promising but since it has not been published yet (it's under review I assume), I think it would be better to not include it in the references (reference 15).</p> <p>Congratulations for your manuscript.</p>
-------------------------	---

REVIEWER	Gorkem Sezgin Australian Institute of Health Innovation, Macquarie University, Australia
REVIEW RETURNED	19-Feb-2019

GENERAL COMMENTS	<p>Congratulations to the authors on completing this globally relevant and well-written study on the barriers for increasing healthcare quality through reduction of unnecessary testing of vitamins D and B12. This study will be a valuable contribution to our understanding of decision-making processes in general practice and the implementation of interventions to improve this process. Some minor points for the consideration of the authors:</p> <ul style="list-style-type: none"> - Page 3, line 28: et al should have a period after "al", i.e. et al. . - I have not seen the rationale in the text for the use of grounded-theory design, which is part of the SRQR guideline, point 5. The inclusion of this might benefit readers who might be unfamiliar with these themes, such as a quantitative researcher like myself, better understand the rationale. - Page 5, line 7-9: The authors have pointed that they expect interviews with 20 GPs and 20 patients to be sufficient for item saturation. I am unclear about the justification for this value. It is arguable whether the reference provided could be applicable in this case, as the reference states, "It is hard to say how generalizable our findings might be.". Furthermore, the reference provided suggests a sample of 12 to be sufficient, whilst providing references to other studies suggesting other arguable values. The authors might want to consider, alternatively, providing a more robust justification and explanation for their item saturation, which could simply be a sentence or two indicating whether they believe they have reached item saturation, and if applicable, when. The authors have indicated that they "monitor progress towards saturation", yet there is no indication of whether this is achieved. <p>I would like to once again congratulate the authors on the completion of their study, and extend my thanks for the opportunity to review this manuscript.</p>
-------------------------	--

REVIEWER	Elisabeth Björk Brämberg Karolinska Institutet Sweden
REVIEW RETURNED	10-Jun-2019

GENERAL COMMENTS	<p>Dear Editor,</p> <p>Thank you for the review assignment for BMJ Open. In my review I've addressed the method-section (as specified in the invitation letter). Further, I've some overall comments regarding the study's title and introduction. These are listed below.</p> <p>This is an interesting manuscript, and a good idea to add a qualitative perspective to the RCT mentioned in the manuscript.</p> <p>Title: It's a bit unclear to me what is meant with "Qualitative analysis" and "Grounded theory design". Please state whether this study includes empirical data (in my reading, a qualitative analysis indicates a re-use of already collected data). Please clarify the "design" in "grounded theory design".</p> <p>Introduction: The authors address a knowledge gap regarding previous research. In my opinion, the authors could address this knowledge gap by introducing a theoretical framework for example from the implementation area. The authors refer to Grol and</p>
-------------------------	--

	<p>Wensing in the method section, which is good but since there's a lack of clinical studies, a theoretical framework could strengthen the introduction (and discussion).</p> <p>One of the main concerns in the method-section is the weakness in describing the numbers of interviews that have been conducted and the numbers of interviews used in the analysis. The authors calculated that a total of 40 participants should be needed in order to reach saturation, but their analysis was "saturated" after 14, which is only 25% of the included participants.</p> <p>Method</p> <p>Design and setting: please clarify the study design which in the title is stated as a grounded theory design. In the current version the RCT and REVERT study is mention, but no information regarding the choice of a GT design and reasons for motivating this study design.</p> <p>Recruitment of participants: Which principles guided the selection of participants, a strategic sampling procedure, other (GPs as well as patients)? The GPs who recruited patients, what kind of information did they receive? Did the information provided to the GPs and patients include that your interest was "unnecessary"? In such case, consider the concept's impact on the participants' willingness to speak free during the interviews.</p> <p>Data collection: How was the grounded theory approach applied in the data collection process? Had the two master's student any prior experience or training in qualitative interviewing? Were the interview protocols developed by the authors or have the protocols and questions been used before? Were the questions validated? Barriers and facilitators are stated in the title and aim, were there specific questions about Bs and Fs?</p> <p>Please give a detailed explanation of the development of the questions and protocols including for example pilot-testing. The authors refer to Grol, Wensing and Bosch, it's a bit unclear how this reference was used and for what. The mean time for the patient-interviews was 15 minutes. Were all topics covered in these interviews?</p> <p>You expected to interview 40 participants (20 GPs and 20 patients) in order to reach saturation. In the analysis you mention that saturation was reach after coding about 14 interviews. These 14 include both GPs and patients. How many interviews were conducted? Did you exclude interviews from the analysis based on the principle of saturation? These issues need some clarification. The method section should be provided with a flow chart describing the selection (how many GPs/patients received an invitation, numbers that consented to participation, numbers and reasons for non-participation, numbers interviewed, and numbers analyzed).</p> <p>Data analysis: Please specify how the applied grounded theory method was used in the analysis. Clearly describe the two different approaches (deductive and inductive) used and how these different analyses were performed. NVivo is a software for structuring data, not analyzing. You coded 14 interviews from GPs and patients, this is only 25% of what you calculated as sufficient in the data collection (p 5, line 7). The concept of saturation is questionable. How did you use this concept? When (and why) was saturation reached? Did the authors use for example triangulation or research group reflexivity?</p>
--	--

VERSION 1 – AUTHOR RESPONSE

–Comments of reviewer 1

#1

The authors state that there is little evidence strategies to reduce vitamin overtesting in general practice (Abstract, line 6). I suggest to the paper (Clin Chem Lab Med. 2018 Jun 27;56(7):1063-1070. doi: 10.1515/cclm-2017-1031.), where they show that "Send & hold" clinical decision supporting rules in electronic request are a promising informatics tool that can support in vitamins utilization management.

Response: We thank this reviewer for his suggestion; we added this paper as a reference in the introduction of our manuscript and also included a remark in the Introduction of the revised manuscript.

Old text:

Understanding barriers to, and facilitators for, reducing over-testing is essential to develop a long-term strategy to tackle this problem.

New text (p.3 line 33-37):

So far, there is little evidence of effective strategies to reduce this over-testing in general practice, although clinical decision support rules seem promising.¹¹ Understanding barriers to, and facilitators for, reducing over-testing is essential to develop a long-term strategy to tackle this problem.¹⁰

#2

Some of the conclusions of author's manuscript have been described before in previous works (barriers and facilitators). I recommend to review Salinas M & Lopez-Garrigos M et al. works regarding this subject. In fact, there are numerous references in the literature to educational interventions (verbal and written initiatives) as a strategy for test management. Unfortunately these strategies require exhaustive monitoring on the part of the laboratory since its effects decrease over time (short sustainability). I would recommend the authors to address this issue in their manuscript including proper references.

Response: We thank this reviewer for the suggested literature. We added the following references about educational interventions as an strategy for test management in our discussion:

23. Salinas M, López-Garrigós M, Flores E, et al. Education and communication is the key for the successful management of vitamin D test requesting. *Biochem Med (Zagreb)* 2015; 25: 237–241.

24. Miyakis S, Karamanof G, Lontos M, et al. Factors contributing to inappropriate ordering of tests in an academic medical department and the effect of an educational feedback strategy. *Postgrad Med J.* 2006; 82:823–829. doi:10.1136/pgmj.2006.049551

Old text:

Patients and GPs having conflicting information was one of the main barriers to reducing unnecessary vitamin testing. In line with our results, previous research has highlighted that not only health professionals, but also the media, are key information providers on this topic for patients.²⁰ A media content analysis showed that news articles linked vitamin D to a wide range of health conditions without conclusive scientific evidence.²¹ As reflected by our study as well as previous research, this has resulted in confusion regarding the usefulness of vitamin testing, among both patients and GPs.^{9, 22} Moreover, GPs' information sources also present conflicting results, reinforcing this confusion. To counter this, previous research highlighted the need for clear information that reflects the actual state of knowledge and for ongoing research for both healthcare professionals and patients.^{9, 10, 20} Similarly, GPs in this study mentioned that clear guidelines for patients and GPs regarding vitamin testing would help them in discussions with their patients.

New text (p.14 line 29-40, p.15 line 1-3):

Patients and GPs having conflicting information was one of the main barriers to reducing unnecessary vitamin testing. In line with our results, previous research has highlighted that not only health professionals, but also the media, are key information providers on this topic for patients.²⁰ A media content analysis showed that news articles linked vitamin D to a wide range of health conditions without conclusive scientific evidence.²¹ As reflected by our study as well as previous research, this has resulted in confusion regarding the usefulness of vitamin testing, among both patients and GPs.^{9, 22} Moreover, GPs' information sources also present conflicting results, reinforcing this confusion. To counter this, previous research highlighted the need for clear information that reflects the actual state of knowledge and for ongoing research for both healthcare professionals and patients.^{9, 10, 20} Similarly, GPs in this study mentioned that clear guidelines for patients and GPs regarding vitamin testing would help them in discussions with their patients. In line with this, in this study education was found as one the most important facilitators for reducing vitamin testing. Previous research showed that education and communication through electronic educational codified comments might improve vitamin requests.²³ In addition, strategies for reducing unnecessary vitamin testing require continuous education, because the intervention-effect of education seems to decrease over time.²⁴

#3

The manuscript is quite interesting not only because we know the reason argued by GPs regarding overtesting but the patient's as well. REVERT study seems very promising but since it has not been published yet (it's under review I assume), I think it would be better to not include it in the references (reference 15).

Response: In line with the policy of BMJ Open, we decided to cite the article only in the text, not as a formal reference. We do find it important to cite this paper, as it points out the specific intervention the present study was part of.

Old text:

The REVERT study was an RCT assessing the effectiveness of a GP intervention programme including education, monitoring, and feedback on numbers in relation to ordering vitamin D and B12 tests. Four times a year, GPs received feedback on the number of tests they ordered. After randomisation, half of all participating practices also received patient information on vitamin testing.¹⁵ In total, 22 general practices (117 GPs with 134,000 patients) in the Utrecht region and 4 health centres (41 GPs and 61,000 patients) in the Rotterdam region participated in the REVERT study.

New text (p.4 line 27-34):

The REVERT study was an RCT assessing the effectiveness of a GP intervention programme including education, monitoring, and feedback on numbers in relation to ordering vitamin D and B12 tests. Four times a year, GPs received feedback on the number of tests they ordered. After randomisation, half of all participating practices also received patient information on vitamin testing. In total, 22 general practices (117 GPs with 134,000 patients) in the Utrecht region and 4 health centres (41 GPs and 61,000 patients) in the Rotterdam region participated in the REVERT study (van Vugt SF, de Schepper EIT, van Delft S. et al. Reducing vitamin test ordering in primary care: the effectiveness of a professional and patient oriented strategy).

Comments of reviewer 2

Congratulations to the authors on completing this globally relevant and well-written study on the barriers for increasing healthcare quality through reduction of unnecessary testing of vitamins D and B12. This study will be a valuable contribution to our understanding of decision-making processes in general practice and the implementation of interventions to improve this process.

Response: We thank this reviewer for the compliments on our manuscript, as well as agreeing to its relevance.

#1

Page 3, line 28: et al should have a period after "al", i.e. et al.

Response: We edited the manuscript as suggested by the reviewer.

#2

I have not seen the rationale in the text for the use of grounded-theory design, which is part of the SRQR guideline, point 5. The inclusion of this might benefit readers who might be unfamiliar with these themes, such as a quantitative researcher like myself, better understand the rationale.

Response: We agree with the reviewer that the choice for the grounded theory design can be explained more clearly. We have added a short explanation to the Methods section:

Old text:

SRQR reporting guidelines were used for this qualitative study.¹³ This qualitative study, based on a grounded theory design,¹⁴ used semi-structured interviews among GPs and patients from two primary care networks in the Netherlands that participated in the REVERT study (REducing Vitamin tEsting in pRimary care pracTice).

New text ((p.4 line 14-27):

SRQR reporting guidelines were used for this qualitative study.¹⁴ This qualitative study used a grounded theory design¹⁵, because this design is explicitly suited for examining how meanings in people's perceptions are related to their actions. Applied to our study, using grounded theory allowed us to study how meanings attached to vitamin testing interrelate to choices and actions regarding vitamin testing, for both GPs and patients. The aim is, ultimately, to develop new theoretical concepts, grounded in qualitative data, which represent barriers and facilitators for vitamin testing, currently not reported in the literature. These new theoretical concepts may be further developed and tested in future research.

Data were collected through semi-structured interviews among GPs and patients from two primary care networks in the Netherlands that participated in the REVERT study (REducing Vitamin tEsting in pRimary care pracTice).

#3

Page 5, line 7-9: The authors have pointed that they expect interviews with 20 GPs and 20 patients to be sufficient for item saturation. I am unclear about the justification for this value. It is arguable whether the reference provided could be applicable in this case, as the reference states, "It is hard to say how generalizable our findings might be.". Furthermore, the reference provided suggests a sample of 12 to be sufficient, whilst providing references to other studies suggesting other arguable values. The authors might want to consider, alternatively, providing a more robust justification and explanation for their item saturation, which could simply be a sentence or two indicating whether they believe they have reached item saturation, and if applicable, when. The authors have indicated that they "monitor progress towards saturation", yet there is no indication of whether this is achieved.

Response: We thank this reviewer for careful reading our manuscript. We agree that the numbers presented as 'sufficient', in ours as well as in cited papers, may appear to be arbitrary. In order to provide a stronger explanation for our claim to saturation, we decided to revise two paragraphs of the Methods section. This made us realize that we could improve the clarity of the data collection process, by mentioning the final number of interviews in this section instead of in the Results section. All in all, we have revised the manuscript in three places. The first section pertaining to data collection

has been changed as follows (Please note that this section also contains revisions in response to comments #2 and #7 by Reviewer 3):

Old text:

The interviews were performed during the last quartile of the intervention period of the REVERT study. All interviews were performed by two interviewers (HH, RB). The interviewers were two master's medical students with a background in medical research and/or qualitative research. The GP interviews were conducted face-to-face in the GPs' office, and the patient interviews were conducted by telephone. The interviews were semi-structured, and the content was developed collaboratively in a multidisciplinary team of researchers, GPs, and a psychologist (BM), using previous research about analysing de-implementation projects.¹⁶ The interviews lasted approximately 30 minutes and 15 minutes for GPs and patients, respectively, and consisted of four broad topic sections covering barriers and facilitators for reducing the number of (unnecessary) vitamin D and B12 laboratory tests ordered. The four topics were: 1) perceptions of, and reasons for, vitamin D and B12 testing; 2) cognitive, motivational, and social factors potentially influencing the number of vitamin tests ordered;¹⁶ 3) evaluation of the study intervention (e-module, education, and feedback); 4) ideas regarding a successful strategy for a durable reduction in vitamin test ordering. Baseline characteristics of GPs (sex, age, years working as GP, intervention group (de-implementation strategy 1 or 2), and patients (sex, age, and education level) were ascertained at the end of the interview. Data on number of patients per practice were retrieved by emailing the practices. In addition, data on socioeconomic status (SES) were retrieved from the Social and Cultural Planning Office (SCP) in the Netherlands and linked to our data through the four digits of the postal codes of the practice area. SCP calculates socioeconomic status scores based on information concerning education, income, and position in the labour market.¹⁷ We expected interviews with 20 GPs and 20 patients to be sufficient for item saturation.¹⁸ During data collection, interim meetings were held with the interviewers (HH, RB) and psychologist (BM) to discuss data and monitor progress towards saturation.

New text (p.5 line 12-40, p.6 line 1-2):

The interviews were performed by two interviewers (HH, RB), during the last quartile of the intervention period of the REVERT study. The interviewers were two master's medical students with a background in medical research and/or qualitative research, supported by a multidisciplinary team of researchers, GPs, and a psychologist specialized in communication research (BM). BM trained HH and RB in how to apply guidelines for doing in-depth interviews.

The GP interviews were conducted face-to-face in the GPs' office, and the patient interviews were conducted by telephone. Interviews lasted approximately 30 minutes and 15 minutes for GPs and patients, respectively, and were semi-structured using a list that covered four broad topics of barriers and facilitators for reducing the number of (unnecessary) vitamin D and B12 testing. The four topics were based on the framework by Grol and Wensing¹⁶, namely: 1) perceptions of, and reasons for, vitamin D and B12 testing; 2) cognitive, motivational, and social factors potentially influencing the number of vitamin tests ordered;¹⁶ 3) evaluation of the study intervention (e-module, education, and feedback); 4) ideas regarding a successful strategy for a durable reduction in vitamin test ordering. Baseline characteristics of GPs (sex, age, years working as GP, intervention group (de-implementation strategy 1 or 2), and patients (sex, age, and education level) were ascertained at the end of the interview. Data on number of patients per practice were retrieved by emailing the practices. In addition, data on socioeconomic status (SES) were retrieved from the Social and Cultural Planning Office (SCP) in the Netherlands and linked to our data through the four digits of the postal codes of

the practice area. SCP calculates socioeconomic status scores based on information concerning education, income, and position in the labour market.¹⁷ We expected interviews with 20 GPs and 20 patients to be sufficient for item saturation.¹⁸ During data collection, interim meetings were held with the interviewers (HH, RB) and psychologist (BM) to discuss data and monitor progress towards saturation.

Based on a previous study, we expected a minimum of approximately 12 interviews with GPs and 12 interviews with patients to be sufficient for saturation,¹⁸ although numbers mentioned in the literature vary, and thus cannot be taken as absolute indicators of saturation or any other criterium. To guarantee at least 12 interviews per group, the aim was to organise about 20 interviews with GPs and 20 interviews with patients. Twenty-one GPs from different practices were invited to participate. One GP declined, so in total 20 GPs agreed to participate in this study (5 GPs in Rotterdam and 15 GPs in Utrecht). Of the 22 patients who consented to participate in the study, 3 could not be reached by telephone by the researchers, resulting in 19 interviewed patients.

The second revision in the Methods section, to clarify what we took as an indication of saturation:

Old text:

The interviews were recorded on audiotape and transcribed verbatim. Next, these data were coded combining a deductive (i.e. Groll and Wensing's framework)¹⁶ and an inductive (i.e. data-driven) approach, using QSR NVivo (version 11).¹⁹ All interviews were coded independently by two researchers (HH and RB). The emerging themes were continuously compared with interview transcripts. After coding about 14 interviews for both the GP and the patient group, no new codes were added, indicating data saturation. The assigned codes and themes were discussed by the coding researchers until consensus was achieved. Three researchers (RB, HH, and BM) further discussed the themes and categorised them into interrelated topics.

New text (p.6 line 5-15):

The interviews were recorded on audiotape and transcribed verbatim. Next, these data were coded combining a deductive (i.e. Grol and Wensing's framework)¹⁶ and an inductive (i.e. data-driven) approach, using QSR NVivo (version 11).¹⁹ All interviews were coded independently by two researchers (HH and RB). The emerging themes were continuously compared with interview transcripts. During data collection, interim meetings were held with the interviewers (HH, RB) and communication researcher (BM) to discuss data collection and analysis, including emerging themes and how these interrelated. The assigned codes and themes were discussed by the coding researchers until consensus was achieved.

Data saturation was monitored and discussed as well. After coding 14 interviews for the GP group and 14 interviews for the patient group, no new codes were added, which means that data saturation was reached at that point.

Finally, the Results section was edited by moving the discussion of the number of participants to the Methods section:

Old text

In total, 21 GPs from different practices were invited to participate. One GP declined, so in total 20 GPs agreed to participate in this study (5 GPs in Rotterdam and 15 GPs in Utrecht). The GPs' characteristics are summarised in Table 1. Of the 22 patients who consented to participate in the study, 3 could not be reached by telephone by the researchers. The characteristics of the final 19 patients interviewed are also summarised in Table 1.

New text (p.6 line 24-25):

The characteristics of the 20 GPs and 19 patients who participated in the study are summarised in Table 1.

Comments of reviewer 3

Dear Editor,

Thank you for the review assignment for BMJ Open. In my review I've addressed the method-section (as specified in the invitation letter). Further, I've some overall comments regarding the study's title and introduction. These are listed below.

This is an interesting manuscript, and a good idea to add a qualitative perspective to the RCT mentioned in the manuscript.

#1

Title: It's a bit unclear to me what is meant with "Qualitative analysis" and "Grounded theory design". Please state whether this study includes empirical data (in my reading, a qualitative analysis indicates a re-use of already collected data). Please clarify the "design" in "grounded theory design".

Response: We appreciate this suggestion; however, we are not aware that the term 'qualitative analysis' may indicate a secondary analysis, as suggested by the reviewer. We do agree that the title is a bit too complex for readers who are not familiar with qualitative research designs. That is why we have reformulated the title to present a more general instead of 'technical' description of our work.

Old title: Barriers and facilitators for reducing unnecessary vitamin testing in general practice: a qualitative analysis based on a grounded theory design.

New title: Reducing unnecessary vitamin testing in general practice: barriers and facilitators according to general practitioners and patients.

Secondly, in response to comment #2 of Reviewer 2, we have revised our manuscript to provide a more comprehensive explanation of what a grounded theory design entails.

#2

Introduction: The authors address a knowledge gap regarding previous research. In my opinion, the authors could address this knowledge gap by introducing a theoretical framework for example from the implementation area. The authors refer to Grol and Wensing in the method section, which is good but since there's a lack of clinical studies, a theoretical framework could strengthen the introduction (and discussion).

Response: We agree with the reviewer regarding the relevance of a theoretical framework in this research domain. However, the lack of theory in this domain is exactly what warrants a grounded theory study. In order to strengthen our rationale for this study, we have revised a section in the Introduction:

Old text:

So far, no detailed information is available on the barriers and facilitators for rationalisation of vitamin test ordering in general practice. Therefore, we performed a qualitative assessment using semi-structured interviews among both GPs and patients to explore the barriers and facilitators for reducing the number of unnecessary vitamin D and B12 laboratory tests ordered.

New text (p. 4 line 5-8):

So far, theoretical perspectives as well as empirical studies on the barriers and facilitators of vitamin test ordering in general practice are lacking. Therefore, we performed a qualitative assessment using semi-structured interviews among both GPs and patients to explore the barriers and facilitators for reducing the number of unnecessary ordered vitamin D and B12 laboratory tests.

Secondly, we have revised the description of the data collection in the Methods section in response to comment #3 of Reviewer 2, to clarify that we used the Grol & Wensink framework to develop the interview topic list.

#3

One of the main concerns in the method-section is the weakness in describing the numbers of interviews that have been conducted and the numbers of interviews used in the analysis. The authors calculated that a total of 40 participants should be needed in order to reach saturation, but their analysis was "saturated" after 14, which is only 25% of the included participants.

Response: We fully agree with the reviewer that the description about the number of interviews that were aimed for, and the number of interviews actually performed should be presented much more clearly. We believe that we have done so when we revised our manuscript in response to comment #3 of Reviewer 2.

#4

Method

Design and setting: please clarify the study design which in the title is stated as a grounded theory design. In the current version the RCT and REVERT study is mention, but no information regarding the choice of a GT design and reasons for motivating this study design.

Response: In response to comment #2 of Reviewer 2, we have revised the manuscript to present a stronger rationale for choosing a grounded theory design. Please see the revised text mentioned above.

#5

Recruitment of participants: Which principles guided the selection of participants, a strategic sampling procedure, other (GPs as well as patients)? The GPs who recruited patients, what kind of information did they receive? Did the information provided to the GPs and patients include that your interest was "unnecessary"? In such case, consider the concept's impact on the participants' willingness to speak free during the interviews.

Response: We agree with the reviewer that the guidelines for recruitment deserve more explanation. Therefore, we have provided a more thorough description of the recruitment process and underlying principles in the Methods section:

Old text:

At the end of the one-year intervention period, a subset of GPs was invited for an interview. To secure an adequate case mix regarding practice type and socioeconomic status of the practice area, only 1 general practitioner per REVERT practice was invited for an interview. Half of all invited GPs were working in a practice that had received patient information on vitamin testing. The GPs were recruited by mail and telephone.

Patients were recruited through the participating GPs; GPs were asked to invite patients during consultations in which vitamin testing was a topic of conversation. When patients consented to be interviewed on this topic, GPs provided the patients' name and telephone number to the researchers, who contacted the patients.

New text (p. 4 line 37-40, p.5 line 2-9):

At the end of the one-year intervention period, we have invited all participating general practices for an interview by telephone or face to face by one of the researchers. To secure an adequate case mix regarding practice type and socioeconomic status of the practice area, only 1 general practitioner per REVERT practice was invited for an interview.

Patients were recruited through the participating GPs; GPs were asked to invite patients during consultations in which vitamin testing was a topic of conversation. The GPs asked them if they were willing to be interviewed about vitamin testing. When patients consented to be interviewed on this topic, GPs provided the patients' name and telephone number to the researchers, who contacted the patients. We aimed to recruit a mixed sample in terms of age, gender, ethnicity and educational level, because large variation as to demographic characteristics helps to recruit a sample with the widest range of possible experiences, opinions and preferences. This is necessary for a full exploration of this issue.

Response: Regarding the reviewer's second remark, patients did have a consultation prior to the interview, in which some of the patients were denied a vitamin test. This was, however, part of the research; patients could speak freely to the interviewers, who were not associated with the GP practice. The interviewers did not observe any evidence of patients being hesitant to speak freely.

#6

Data collection:

How was the grounded theory approach applied in the data collection process?

Response: We agree with the reviewer that the choice for the grounded theory design can be explained more clearly. In response to comment #2 of Reviewer 2, we have revised the manuscript to present a stronger rationale for choosing a grounded theory design. Please see the revised text mentioned above.

#7

Had the two master's student any prior experience or training in qualitative interviewing?

Response: Again, the comments of Reviewers 2 and 3 partly overlap; the new text can be found under our response to comment #3 of Reviewer.

#8

Were the interview protocols developed by the authors or have the protocols and questions been used before?

Response: The interview protocol were developed by the authors and questions have not been used before. This is already described in the full-text: "The interviews were semi-structured, and the content was developed collaboratively in a multidisciplinary team of researchers, GPs, and a psychologist (BM) using previous research about analysing de-implementation projects."

#9

Were the questions validated?

Response: the questions were not validated. However, this was an exploratory study, with a theory-based topic list, which is common practice in this type of qualitative studies. Please also note that in the Methods section (in both the original and the revised manuscript), it is explained that "During data collection, interim meetings were held with the interviewers (HH, RB) and psychologist (BM) to discuss data and monitor progress towards saturation." These discussions revealed that the topic list was functional; the questions incited GPs and patients to tell about their personal preferences, beliefs and experiences.

#10

Barriers and facilitators are stated in the title and aim, were there specific questions about Bs and Fs?

Response: Yes, there were several questions about Bs and Fs in the interviews. Specific barriers and facilitators were explored based on the framework by Grol and Wensing, as explained in the Methods section. Due to the exploratory set-up of the interviews, interviewees could talk about any additional Bs and Fs they perceived.

#11

Please give a detailed explanation of the development of the questions and protocols including for example pilot-testing.

Response: A detailed explanation of the development of the questions is described in the full-text, see comment #8 and #9 of Reviewer 3.

#12

The authors refer to Grol, Wensing and Bosch, it's a bit unclear how this reference was used and for what.

Response: In response to comment #3 of Reviewer 2, we have revised the manuscript to make clear that the framework of Grol and Wensing was used to create four broad head topics in the topic list.

#13

The mean time for the patient-interviews was 15 minutes. Were all topics covered in these interviews?

Response: Yes all topics were covered in these 15 minutes.

#14

You expected to interview 40 participants (20 GPs and 20 patients) in order to reach saturation. In the analysis you mention that saturation was reached after coding about 14 interviews. These 14 include both GPs and patients. How many interviews were conducted? Did you exclude interviews from the analysis based on the principle of saturation? These issues need some clarification. The method section should be provided with a flow chart describing the selection (how many GPs/patients received an invitation, numbers that consented to participation, numbers and reasons for non-participation, numbers interviewed, and numbers analyzed).

Response: We fully agree with the reviewer that the description about the number of interviews that were aimed for, and the number of interviews actually performed should be presented much more clearly. We believe that we have done so when we revised our manuscript in response to comment #3 of Reviewer 2. We believe that including a flow-chart will not further clarify these numbers.

#15

Data analysis: Please specify how the applied grounded theory method was used in the analysis. Clearly describe the two different approaches (deductive and inductive) used and how these different analyses were performed. NVivo is a software for structuring data, not analyzing. You coded 14 interviews from GPs and patients, this is only 25% of what you calculated as sufficient in the data collection (p 5, line 7). The concept of saturation is questionable. How did you use this concept? When (and why) was saturation reached? Did the authors use for example triangulation or research group reflexivity?

Response: In response to comment #3 of Reviewer 2, we have revised the manuscript to give a more detailed description the use of the grounded theory and the concept of saturation. As stated, this also includes a more complete and clear description of the number of interviewees. To specify, we have interviewed 19 patients, thus well above our aim of 12.

VERSION 2 – REVIEW

REVIEWER	Gorkem Sezgin Macquarie University, Australia
REVIEW RETURNED	23-Jul-2019

GENERAL COMMENTS	I would recommend the authors to review the wording of the paragraph starting Page 4 Line 14. In particular, the sentence starting line 16 feels redundant considering the following sentence.
-------------------------	--

REVIEWER	Elisabeth Björk Brämberg Karolinska Institutet, Sweden
REVIEW RETURNED	05-Aug-2019

GENERAL COMMENTS	Dear editor, the authors have done a nice job revising the manuscript, and my recommendation is therefore Accept.
-------------------------	---

VERSION 2 – AUTHOR RESPONSE

Reviewer 2

#1

I would recommend the authors to review the wording of the paragraph starting Page 4 Line 14. In particular, the sentence starting line 16 feels redundant considering the following sentence.

Response: We agree with the reviewer that the sentence starting line 16 feels redundant considering the following sentence. We therefore combined the two sentence resulting in one new sentence. Furthermore we shortened the sentence starting with "The aim is...", which, in our opinion, results in better understanding of the sentence.

Old text:

Design and setting

SRQR reporting guidelines were used for this qualitative study.¹⁴ This qualitative study used a grounded theory design,¹⁵ because this design is explicitly suited for examining how meanings in people's perceptions are related to their actions. Applied to our study, using grounded theory allowed us to study how meanings attached to vitamin testing interrelate to choices and because this design is explicitly suited for examining how meanings in people's perceptions are related to their actions. Applied to our study, using grounded theory allowed us to study how meanings attached to vitamin testing interrelate to choices and actions regarding vitamin testing, for both GPs and patients. The aim is, ultimately, to develop new theoretical concepts, grounded in qualitative data, which represent barriers and facilitators for vitamin testing, currently not reported in the literature. These new theoretical concepts may be further developed and tested in future research.

New text (p.... line ...)

Design and setting

SRQR reporting guidelines were used for this qualitative study.¹⁴ This qualitative study used a grounded theory design,¹⁵ because this design is explicitly suited for examining how meanings in people's perceptions are related to their actions. Applied to our study, using grounded theory allowed us to study how meanings attached to vitamin testing interrelate to choices and actions regarding vitamin testing, for both GPs and patients. The aim is, ultimately, to develop new theoretical concepts, grounded in qualitative data, which represent barriers and facilitators for vitamin testing. These new theoretical concepts may be further developed and tested in future research.

Reviewer 3

#1

Dear editor, the authors have done a nice job revising the manuscript, and my recommendation is therefore Accept.

Response: We would like to thank the reviewer for reviewing our manuscript.

VERSION 3 - REVIEW

REVIEWER	Gorkem Sezgin Australian Institute of Health Innovation, Macquarie University, Sydney, Australia
REVIEW RETURNED	15-Aug-2019

GENERAL COMMENTS	Congratulations to the authors on their study. I believe their research will be a valuable contribution to the journal, and therefore suggest the manuscript to be accepted.
-------------------------	--