PEER REVIEW HISTORY

ARTICLE DETAILS

TITLE (PROVISIONAL)	Attitudes and perceptions of health professionals towards management of hypothyroidism in general practice: a qualitative interview study
AUTHORS	Dew, Rosie; King, Kathryn; Okosieme, Onyebuchi; Pearce, Simon; Donovan, Gemma; Taylor, Peter; Hickey, Janis; Dayan, Colin; Leese, Graham; Razvi, Salman; Wilkes, Scott

VERSION 1 – REVIEW

REVIEWER	Jim Parle University of Birmingham B15 2TT UK
REVIEW RETURNED	16-Oct-2017

GENERAL COMMENTS	Training of the researcher doesn't seem to be described? A bit more information on recruitment would help: how many people refused? Is there anything linking the responders? (I assume not) References: I think quoting a 1998 paper on congivitve dysfunction and hypothyroidism is a bit odd, especially as there is more recent evidence. This sentence is internally contradictory; I suggest putting 'Most' at the front!! "Interviews were conducted face-to-face at the participant's place of work or the University of Sunderland, however, due to participant availability, one GP was interviewed at home, and one GP interviewed by telephone." Similarly there are a few minor typos, words running into each other and so on; and this clause: "an age specific and symptom focus were appropriate may help" probably should read: "an age specific and symptom focus WHERE appropriate may help "

REVIEWER	Brooke Nickel University of Sydney, Australia
REVIEW RETURNED	25-Oct-2017

GENERAL COMMENTS	This paper explores the attitudes and perceptions of health professionals towards management of hypothyroidism across a few counties in northern England.

	Overall the paper is well written and the methodology is sound however, I have a few concerns including the generalisability of the data and the small sample of participants (particularly the 4 pharmacists and 3 different nurses). I suggest a major revision of this manuscript to be accepted. My comments and suggestions for the authors to consider are below. Main comments: 1. In the Abstract and the Methods you state that this is an in-depth interview study however the interview seem to be semi-structured and last approximately 45 minutes without any further feedback from participants. I am not sure I would classify this as in-depth but rather a "qualitative study using semi-structured interviews". Please change or justify. 2. Again, in the abstract and the Settings section of the Methods you name the counties in England where participants were invited from however it is not clear that these are counties and where they are in England. I only was able to get the information about the North of England from the Strengths and Limitations and Discussion sections of the manuscript and had to google that they were indeed counties. Please make this clearer so that people from outside the UK can put this information into context. 3. Were the transcripts coded by more than one researcher or was RD the only researcher who coded these transcripts? If so, please make this clear in the Analysis section of the Methods and include as a limitation in the Discussion (ie. that the data may be prone to the researcher's bias). 4. I suggest removing the reference to Table 1 and Figure 1 from the Methods and placing it at the beginning of the Results section along with a few opening sentences or a paragraph. At the moment, it feels like you just jump right into the themes. Perhaps you could include a breakdown of the participant characteristics as it is difficult to collate this easily from the current format of Table 1 and a description of how the themes and sub-themes are laid out. 5. The Results seems quite long and quote heavy. Please consider removing one or two of the quotes from some of the themes. Quotes are meant to support the theme discussed and not provide information on everything that was said. 6. I feel like some of the results and discussion points on pharmacists ie. "Pharmacists felt like they could provide limited advice…" are difficult to imply as the sample only included 4 pharmacists. Please consider removing this information and making the results and discussion more general (ie "overall health professionals…") or stating this as a major limitation in the Discussion. 7. I am a bit confused who the word "prescriber" refers to? Is it the GPs, the GPs and pharmacists, all of the health professionals? Please make this clear throughout. 8. Also, try and be consistent when saying "professionals" vs "health professionals" vs "prescribers" (if applicable) throughout the manuscript. My preference would be for "health professionals" and not just professionals. 9. As mentioned above please highlight the small sample of pharmacists and nurses in the Strengths and Weaknesses section of the Discussion. 10. Please consider re-wording the title for Figure 1. Themes and categories of what?

	Minor comments: 1. Is the word “of” missing in the Participants section of the Abstract after “interviewed comprising…”? Please also add a period to this sentence. 2. Consider changing the Discussion sub-heading “Summary” to “Principal findings” or “A statement of the principal findings”

REVIEWER	Don McLeod QIMR Berghofer Medical Research Institute, Australia
REVIEW RETURNED	13-Nov-2017

GENERAL COMMENTS	The authors use qualitative interviews, specifically a grounded theory approach, to assess the experiences, attitudes, and perceptions of health professionals in the North of England regarding hypothyroidism treatment. Nine GPs, four pharmacists, two practice nurses, and one nurse practitioner were interviewed during the study. The work clearly falls in the scope of BMJ Open. It addresses an important clinical issue: suboptimal replacement of hypothyroidism. A qualitative research design is appropriate to explore important issues, which could be then be used to formulate interventions. The writing is well organized and clear. The difficulty in reviewing any qualitative research with non-random sampling and literally “ad-hoc” data analysis (which then informs later study design) is that more faith than usual is required to believe that the research team has not influenced the reported results. Reviewers (and subsequent readers) also don’t have the same transparency about the appropriateness of study design decisions as a well-performed quantitative study. That said, many “themes” identified by the research team do ring true, and the difficulties elaborated above are not specific to this study. After the following comments/questions are addressed, I would be happy to see this work published. Comments/questions: Strengths and limitations box: (1) This is much more detailed than the relevant section of the Discussion (which doesn’t even mention the potential for researchers influencing study outcome). I think the later manuscript section needs to be expanded. And as a readability issue for the box, it would make more sense to group strengths and then limitations. (2) Were this study’s interviews and thematic development performed before those in the recently published work in BJGP Open (and any literature reviews/meta-analysis planning)? If not, it would be unlikely that the interviewer had no prior assumptions on the topic. Likewise, theme development occurred with other researchers (and the themes identified were then incorporated into later interviews); potential influence of study results is not solely related to the interviewer. For these reasons, I think dismissal of (inadvertent) potential influence by the study team on reported results is too strong.

	Method - Sampling: (3) What is striking about the sample described in Table 1 is that the GPs were generally experienced, as were the nurses, but the pharmacists were all relatively junior. Would the authors expect any differences in results with a different mix of professional experience (e.g., would more experienced community pharmacists have the same desire to view patient records, or may they become more skilled at assessing the issues without it?). (4) The non-random (purposive) sampling intended to mix rural and urban participants. Are there no rural pharmacists? Does this matter? (5) How could saturation of themes occur for nurse practitioners when only one was sampled? For this purpose, were they grouped with GPs because they practice independently, or with practice nurses? (6) Theoretical sampling occurred until no new themes emerge. After how many interviews with no new themes does this occur? One, two, or more? Results – use of guidelines (7) Will clinicians working outside of NHS North of Tyne understand what TRAMP guidelines are (in the quote)? Results - Poor Levothyroxine Adherence as the Main Reason for Suboptimal Treatment (8) I think this section conflates at least two critical issues. If suboptimal treatment is defined as a continuing high TSH level, there is ample evidence in the literature that poor adherence to therapy is the major (but not only) cause of this. Taking excessive levothyroxine because of perceived ongoing thyroid symptoms is a separate phenomenon. Discussion – Summary (9) I would suggest changing or removing the last sentence. In the BJGP Open paper (ref 29), only 5 patients had high “out of range” TSHs, for which this comment is most relevant (and while I think low “out of range” TSHs are a separate issue, no details are given in that paper about why patients had low TSHs, or the cause of their hypothyroidism – i.e., any thyroid cancer patients?). Given the sampling design for ref 29, I strongly suspect that the high “out of range” patients are not representative of all patients with high TSH levels. Discussion – Comparisons with existing literature (10) “Having an individualistic approach for each patient to tailor TSH levels within the reference range has been suggested to be the optimal approach for management of hypothyroid patients.” This should be referenced and more nuance introduced. While it is not a terrible suggestion (because patients may have their own individual TSH set-point), and can be thought of as a practical approach to patients with “normal” TSH levels and ongoing symptoms, it is a very controversial statement. The TRAMP guidelines available on the internet (that these clinicians say they are using) do not mention this approach.

	There is also evidence that having different TSH concentrations within the normal range doesn't improve symptoms (Walsh JCEM 2006; and recent oral presentation of a parallel design RCT by Samuels at the 2017 American Thyroid Association meeting), and I am not aware of any empiric evidence in favour of this approach. (11) "Lack of agreement with national guidelines has been identified by Lugtenberg et al., (2009) to be the main barrier to guideline adherence by GPs…" This sentence is confusing and its relevance is unclear to this study. My understanding of the Dutch paper is that it was personal GP disagreement with individual guideline recommendations (i.e., because they may not apply to their individual patients), not lack of agreement between varying guidelines. Did any of the clinicians self-identify this issue in the in-depth interviews? If not, is it likely an important consideration here? (12) "However, in our qualitative interview study with hypothyroid patients, good levothyroxine adherence was reported by nearly all patients, even those who had TSH levels outside of the reference range…" Again, this statement assumes that the very small number of hypothyroid patients with ongoing high TSH levels who volunteered to be interviewed in the previous study are representative, which is highly unlikely. Discussion - Strengths and limitations: (13) See comment 1. (14) No participants gave feedback or corrections to the interview transcripts. This is expected; the interviewees are all busy people and the transcripts should be accurate. What would actually interest me for qualitative research like this is if the participants agreed with the research team's classifications of their quotes into the various themes. Clearly that would have implications for resources on the study though. References: (15) There are a number of references that are incomplete.

REVIEWER	Dr Vijay Panicker Sir Charles Gairdner Hospital, Perth, Western Australia Australia
REVIEW RETURNED	13-Nov-2017

GENERAL COMMENTS	Dew and colleagues present findings from a well constructed qualitative study on the attitudes of health professionals towards managment of hypothyroidism, in an attempt to further explore the reasons many patients on thyroxine are not at target biochemical levels. The paper is well-written, the findings interesting, if not particularly surprising, and my only concerns are related to how much power the study can have to impact health policy. Major points 1. Sample size. Although not that familiar with this design of study I accept the premise that the sampling proceeded until there was data saturation and no new themes.

	However, it still appears a small sample: 9 GP's, 4 pharmacists and 3 nurses/nurse practitioners. In my experience there is significant variation in knowledge on particular conditions between GP's, depending on experience, training, personal familiarity and I imagine other health professionals to be the same. Are the authors confident, therefore, that this sample is truly reflective of health professionals both locally and nationally? Or can this be used as a pilot to design a more streamlined study which could include a greater sample size? 2. Recruitment and sampling. I am interested to know more about how participants were recruited: were they sampled randomly from a database or did they respond to an advertisement? The reason I feel this is important is that if the responders were volunteers from the clinical research network for instance, could you have selected out GP's who were more likley to follow guidelines and regularly review patients and therefore less likely to have higher numbers of thyroxine treated patients with abnormal TSH's? That is, given that you can't directly show that the GP's selected were those with the usual high rate of off-target thyroxine patients, I think it important to show that you haven't selected GP's who are more likely to follow guidelines and stay up to date with new information. 3. Summary & Implications I feel the summary could be more concise with clear statement of the main findings and implications of these for clinical practice/health policy.

VERSION 1 – AUTHOR RESPONSE

Reviewer: 1
Reviewer Name: Jim Parle
Institution and Country: University of Birmingham, B15 2TT, UK Competing Interests: none declared

1) Training of the researcher doesn't seem to be described?

Response: This can be found under the ‘contribution’ section near the end of the manuscript: “RD gained qualitative research skills through attendance at a qualitative interview training course at Newcastle University, and from working on previous research projects.”

2) A bit more information on recruitment would help: how many people refused? Is there anything linking the responders? (I assume not)

Response: We have not stated how many participants refused to take part as since participants were recruited through the NIHR Clinical Research Network: North East and North Cumbria and local professional networks via mail and email, and the researcher was not directly involved with recruitment, it is unknown how many participants did not want to take part. However, we have added the following statement under ‘Participants’ section in the methods section: “Only the participant and the researcher were present during the interviews, and the researcher had not met the participants previously. Most participants were not known to each other; however, three pharmacists had lecturing roles at the University of Sunderland and were known to each other.”

3) References: I think quoting a 1998 paper on cognitive dysfunction and hypothyroidism is a bit odd, especially as there is more recent evidence.

Response: This reference has been removed and replace with the following:

Juarez-Cedillo T, Basurto-Acevedo L, Vega-Garcia S, Martha ASR, Retana-Ugalde R, Juarez-Cedillo E, et al. Prevalence of thyroid dysfunction and its impact on cognition in older mexican adults: (SADEM study). J Endocrinol Invest 2017;40(9):945-52

4) This sentence is internally contradictory; I suggest putting 'Most' at the front!!
"Interviews were conducted face-to-face at the participant's place of work or the University of Sunderland, however, due to participant availability, one GP was interviewed at home, and one GP interviewed by telephone."

Response: 'Most' has been inserted at the beginning of this sentence found under 'Participants' in the methods section.

5) Similarly there are a few minor typos, words running into each other and so on; and this clause:
"an age specific and symptom focus were appropriate may help"

probably should read:
"an age specific and symptom focus WHERE appropriate may help "

The recommended change has been altered under 'Implications for future research and clinical practice' in the discussion section, and the manuscript has also been read through by the authors to check for any further typos and grammatical errors.

Reviewer: 2

Reviewer Name: Brooke Nickel
Institution and Country: University of Sydney, Australia Competing Interests: None declared.

This paper explores the attitudes and perceptions of health professionals towards management of hypothyroidism across a few counties in northern England. Overall the paper is well written and the methodology is sound however, I have a few concerns including the generalisability of the data and the small sample of participants (particularly the 4 pharmacists and 3 different nurses). I suggest a major revision of this manuscript to be accepted.

My comments and suggestions for the authors to consider are below.

Main comments:
1) In the Abstract and the Methods you state that this is an in-depth interview study however the interviews seem to be semi-structured and last approximately 45 minutes without any further feedback from participants. I am not sure I would classify this as in-depth but rather a "qualitative study using semi-structured interviews". Please change or justify.

Response: In-depth has been removed from the manuscript, and the study is now referred to as a semi-structured interview study in the abstract and methods sections.

2) Again, in the abstract and the Settings section of the Methods you name the counties in England where participants were invited from however it is not clear that these are counties and where they are in England. I only was able to get the information about the North of England from the Strengths and Limitations and Discussion sections of the manuscript and had to google that they were indeed counties. Please make this clearer so that people from outside the UK can put this information into context.

Response: Reference to the above locations as counties and their location within the North of England has been added in the abstract and methods sections.

3) Were the transcripts coded by more than one researcher or was RD the only researcher who coded these transcripts? If so, please make this clear in the Analysis section of the Methods and include as a limitation in the Discussion (ie. that the data may be prone to the researcher's bias).

Response: The transcripts were coded by one researcher. This has been added into the methods section, and the following sentence has also been included in the 'strengths and weaknesses' section of the discussion: "Open coding was performed by RD, however, negotiation of categories and themes were conducted by RD, KK and SW to help reduce this limitation."

4) I suggest removing the reference to Table 1 and Figure 1 from the Methods and placing it at the beginning of the Results section along with a few opening sentences or a paragraph. At the moment, it feels like you just jump right into the themes. Perhaps you could include a breakdown of the participant characteristics as it is difficult to collate this easily from the current format of Table 1 and a description of how the themes and sub-themes are laid out.

Response: Reference to Table 1 and Figure 1 have been removed from methods section and is now located at the beginning of the results section and the following sentence has also been inserted: "Participant characteristics are shown in Table 1. Participant age ranged from 25-60, and four participants were male and twelve were female, with six participants working in rural areas and ten working in urban areas. The three main themes that emerged from the data were perceived health professional control, perceived health professional responsibility and attitudes to thyroid control, which were underpinned to the TPB as shown in Figure 1."

5) The Results seems quite long and quote heavy. Please consider removing one or two of the quotes from some of the themes. Quotes are meant to support the theme discussed and not provide information on everything that was said.

Response: To address this issue, four quotes have been removed from the theme 'perceived health professional control', two quotes have been removed from the theme 'perceived health professional responsibility' and one quote has been removed from the theme 'attitudes to thyroid control'.

6) I feel like some of the results and discussion points on pharmacists ie. "Pharmacists felt like they could provide limited advice…" are difficult to imply as the sample only included 4 pharmacists. Please consider removing this information and making the results and discussion more general (ie "overall health professionals…") or stating this as a major limitation in the Discussion.

Response: Since we cannot summarise completely the findings from the pharmacists that took part in our study with the findings from the GPs and nurses, we have included a sentence in our strengths and weaknesses section under the discussion section that reads "Since our sample only included four pharmacists and three nurses the evidence presented from these groups may lack validity and generalisability outside of this small sample size." This sentence also addresses comment number 9 from this reviewer.

7) I am a bit confused who the word "prescriber" refers to? Is it the GPs, the GPs and pharmacists, all of the health professionals? Please make this clear throughout.

Response: The authors would like to apologise for this confusion. Prescriber has been changes to "GPs and nurses"

8) Also, try and be consistent when saying “professionals” vs “health professionals” vs “prescribers” (if applicable) throughout the manuscript. My preference would be for “health professionals” and not just professionals.

Response: All references to professionals have been changed to “health professionals” throughout the manuscript.

9) As mentioned above please highlight the small sample of pharmacists and nurses in the Strengths and Weaknesses section of the Discussion.

Response: A sentence has been added under ‘Strengths and weaknesses’ in the discussion section. Please see response to comment number 6.

10) Please consider re-wording the title for Figure 1. Themes and categories of what?

Response: The title of Figure 1 has been changed to “Concept diagram showing attitudes and perceptions of health professionals towards management of hypothyroidism in general practice” to give a more coherent explanation to the concept diagram.

Minor comments:
1) Is the word “of” missing in the Participants section of the Abstract after “interviewed comprising…”? Please also add a period to this sentence.

Response: This sentence now reads: “sixteen participants were interviewed between March and August 2016, comprising of nine general practitioners, four pharmacists, two practice nurses and one nurse practitioner”

2) Consider changing the Discussion sub-heading “Summary” to “Principal findings” or “A statement of the principal findings”

Response: ‘Summary’ in the discussion section has now been changed to “Principal Findings”

Reviewer: 3

Reviewer Name: Don McLeod
Institution and Country: QIMR Berghofer Medical Research Institute, Australia Competing Interests: None declared.

The authors use qualitative interviews, specifically a grounded theory approach, to assess the experiences, attitudes, and perceptions of health professionals in the North of England regarding hypothyroidism treatment. Nine GPs, four pharmacists, two practice nurses, and one nurse practitioner were interviewed during the study.

The work clearly falls in the scope of BMJ Open. It addresses an important clinical issue: suboptimal replacement of hypothyroidism. A qualitative research design is appropriate to explore important issues, which could be then be used to formulate interventions. The writing is well organized and clear.

The difficulty in reviewing any qualitative research with non-random sampling and literally “ad-hoc” data analysis (which then informs later study design) is that more faith than usual is required to believe that the research team has not influenced the reported results.

Reviewers (and subsequent readers) also don't have the same transparency about the appropriateness of study design decisions as a well-performed quantitative study. That said, many "themes" identified by the research team do ring true, and the difficulties elaborated above are not specific to this study. After the following comments/questions are addressed, I would be happy to see this work published.

Comments/questions:

1) Strengths and limitations box:
(1) This is much more detailed than the relevant section of the Discussion (which doesn't even mention the potential for researchers influencing study outcome). I think the later manuscript section needs to be expanded. And as a readability issue for the box, it would make more sense to group strengths and then limitations.

Response: The strengths and limitations in the 'strength and limitation box' have been grouped to start with strengths, and then followed by limitations. Additionally, to address comment 3 from reviewer 2, the 'strengths and weaknesses' section in the discussion now has been expanded to include the influence of researcher bias as a study limitation, and also issues with sampling.

(2) Were this study's interviews and thematic development performed before those in the recently published work in BJGP Open (and any literature reviews/meta-analysis planning)? If not, it would be unlikely that the interviewer had no prior assumptions on the topic. Likewise, theme development occurred with other researchers (and the themes identified were then incorporated into later interviews); potential influence of study results is not solely related to the interviewer. For these reasons, I think dismissal of (inadvertent) potential influence by the study team on reported results is too strong.

Response: To address this concern reference to the interviewer not having prior medical assumptions has been removed from the 'strength and limitations box' and the bullet point now reads: "A known potential limitation of qualitative research is the influence the researchers may have had in the process of data gathering and analysis, causing bias. However, to help reduce this limitation the researcher who conducted the interviews did not have medical training, had not met the participants prior to the interviews or discussed the research with them"
2) Method - Sampling:

(3) What is striking about the sample described in Table 1 is that the GPs were generally experienced, as were the nurses, but the pharmacists were all relatively junior. Would the authors expect any differences in results with a different mix of professional experience (e.g., would more experienced community pharmacists have the same desire to view patient records, or may they become more skilled at assessing the issues without it?).

Response: Under 'Strengths and weaknesses' section in the discussion the following sentence has been included to acknowledge this concern: "Since our sample only included four pharmacists and three nurses the evidence presented from these groups may lack validity and generalisability outside of this small sample size. Moreover, a different sample of participants with different levels of experience may have provided different results, particularly pharmacists with more years of experience, and also from rural areas."

(4) The non-random (purposive) sampling intended to mix rural and urban participants. Are there no rural pharmacists? Does this matter?

Response: This issue has been highlighted under ‘Strengths and weaknesses’ section in the discussion section “Since our sample only included four pharmacists and three nurses the evidence presented from these groups may lack validity and generalisability outside of this small sample size. Moreover, a different sample of participants with different levels of experience may have provided different results, particularly pharmacists with more years of experience, and also from rural areas.”

(5) How could saturation of themes occur for nurse practitioners when only one was sampled? For this purpose, were they grouped with GPs because they practice independently, or with practice nurses?

Response: The sentence “Data gathered from the interview with the nurse practitioner was similar to that of the data obtained from the interviews with the practice nurses.” has been included under the sampling section of the methods section.

(6) Theoretical sampling occurred until no new themes emerge. After how many interviews with no new themes does this occur? One, two, or more?

Response: To address this comment the following sentence has been adapted under the ‘sampling’ section of the methods section: “To test emerging themes, theoretical sampling thereafter proceeded until data saturation was achieved and no new themes were forthcoming32 in the following two interviews.”

3) Results – use of guidelines
(7) Will clinicians working outside of NHS North of Tyne understand what TRAMP guidelines are (in the quote)?

Response: This quote has now been expanded to include an explanation of TRAMP guidelines: “I think that's all in the TRAMP [Thyroid Regional Assessment and Management Plan, NHS North of Tyne and Gateshead Area Prescribing Committee, North of England] guidelines. I think I try to follow what I'm meant to do. (GP-4)”

Results - Poor Levothyroxine Adherence as the Main Reason for Suboptimal Treatment
(8) I think this section conflates at least two critical issues. If suboptimal treatment is defined as a continuing high TSH level, there is ample evidence in the literature that poor adherence to therapy is the major (but not only) cause of this. Taking excessive levothyroxine because of perceived ongoing thyroid symptoms is a separate phenomenon.

Response: To address this issue, the title has been change to “incorrect levothyroxine adherence as the main reason for suboptimal treatment”. Additionally, the section now reads: “Some health professionals in our study felt that avoiding weight gain encouraged levothyroxine adherence in patients. Health professionals believed that poor adherence was the main reason patients struggle to lower TSH levels. Additionally, the majority of health professionals said they had experienced a small proportion of patients who would take more levothyroxine than prescribed to alleviate tiredness or in an attempt to lose weight”

4) Discussion – Summary
(9) I would suggest changing or removing the last sentence. In the BJGP Open paper (ref 29), only 5 patients had high “out of range” TSHs, for which this comment is most relevant (and while I think low “out of range” TSHs are a separate issue, no details are given in that paper about why patients had low TSHs, or the cause of their hypothyroidism – i.e., any thyroid cancer patients?). Given the sampling design for ref 29, I strongly suspect that the high “out of range” patients are not representative of all patients with high TSH levels.

Response: This sentence has been removed as suggested.

Discussion – Comparisons with existing literature
(10) “Having an individualistic approach for each patient to tailor TSH levels within the reference range has been suggested to be the optimal approach for management of hypothyroid patients.”
This should be referenced and more nuance introduced. While it is not a terrible suggestion (because patients may have their own individual TSH set-point), and can be thought of as a practical approach to patients with “normal” TSH levels and ongoing symptoms, it is a very controversial statement. The TRAMP guidelines available on the internet (that these clinicians say they are using) do not mention this approach. There is also evidence that having different TSH concentrations within the normal range doesn’t improve symptoms (Walsh JCEM 2006; and recent oral presentation of a parallel design RCT by Samuels at the 2017 American Thyroid Association meeting), and I am not aware of any empiric evidence in favour of this approach.

The statement “Having an individualistic approach for each patient to tailor TSH levels within the reference range has been suggested to be the optimal approach for management of hypothyroid patients.” has been removed from the discussion section.

(11) “Lack of agreement with national guidelines has been identified by Lugtenberg et al., (2009) to be the main barrier to guideline adherence by GPs…”
This sentence is confusing and its relevance is unclear to this study. My understanding of the Dutch paper is that it was personal GP disagreement with individual guideline recommendations (i.e., because they may not apply to their individual patients), not lack of agreement between varying guidelines. Did any of the clinicians self-identify this issue in the in-depth interviews? If not, is it likely an important consideration here?

Response: To address this concern, this sentence has been removed from the discussion section.

(12) “However, in our qualitative interview study with hypothyroid patients, good levothyroxine adherence was reported by nearly all patients, even those who had TSH levels outside of the reference range…”
Again, this statement assumes that the very small number of hypothyroid patients with ongoing high TSH levels who volunteered to be interviewed in the previous study are representative, which is highly unlikely.

Response: To highlight the lack of generalisability of this reference the sentence has been changed to: “However, in our qualitative interview study with hypothyroid patients, although not generlisable outside of the patient population, good levothyroxine adherence was reported by nearly all patients, even those who had TSH levels outside of the reference range29,”

Discussion - Strengths and limitations:
(13) See comment 1.

Please see response to comment 1 above

(14) No participants gave feedback or corrections to the interview transcripts. This is expected; the interviewees are all busy people and the transcripts should be accurate. What would actually interest me for qualitative research like this is if the participants agreed with the research team’s classifications of their quotes into the various themes. Clearly that would have implications for resources on the study though.

Response: Unfortunately, the study was completed in August 2016 and as mentioned by the reviewer the researchers do not have the resources to perform any follow up with the participants, although this indeed this would be very insightful.

5) References:
(15) There are a number of references that are incomplete.

Response: All references have been thoroughly checked, and those that were incomplete have been updated.

Reviewer: 4

Reviewer Name: Dr Vijay Panicker
Institution and Country: Sir Charles Gairdner Hospital, Perth, Western Australia, Australia Competing Interests: None declared

Dew and colleagues present findings from a well-constructed qualitative study on the attitudes of health professionals towards management of hypothyroidism, in an attempt to further explore the reasons many patients on thyroxine are not at target biochemical levels. The paper is well-written, the findings interesting, if not particularly surprising, and my only concerns are related to how much power the study can have to impact health policy.

Major points
1) Sample size.
Although not that familiar with this design of study I accept the premise that the sampling proceeded until there was data saturation and no new themes. However, it still appears a small sample: 9 GP's, 4 pharmacists and 3 nurses/nurse practitioners. In my experience there is significant variation in knowledge on particular conditions between GP's, depending on experience, training, personal familiarity and I imagine other health professionals to be the same. Are the authors confident, therefore, that this sample is truly reflective of health professionals both locally and nationally? Or can this be used as a pilot to design a more streamlined study which could include a greater sample size?

Response: Since our article is reporting qualitative findings, they are not truly reflective of local or national health professionals, but they are transferrable within the patient population. The following sentence found in the ‘strength and weaknesses’ section in the discussion section helps address this issue: “Since our sample only included four pharmacists and three nurses the evidence presented from these groups may lack validity and generalisability outside of this small sample size. Moreover, a different sample of participants with different levels of experience may have provided different results, particularly pharmacists with more years of experience, and also from rural areas. Additionally, within our GP sample it was unknown whether those that took part in our study were more likely to follow guidelines and regularly review patients, thus being less likely to have higher numbers of thyroxine treated patients with abnormal TSH than those who did not take part in our study. However, although this study was conducted with a sample of sixteen participants in the North of England, and may not be generally applicable, the findings are transferrable within this professional population, and this study provides novel insights into the attitudes, experiences and behaviour of health professionals involved in the management of hypothyroidism.”

2) Recruitment and sampling.
I am interested to know more about how participants were recruited: were they sampled randomly from a database or did they respond to an advertisement?

The reason I feel this is important is that if the responders were volunteers from the clinical research network for instance, could you have selected out GP's who were more likley to follow guidelines and regularly review patients and therefore less likely to have higher numbers of thyroxine treated patients with abnormal TSH's? That is, given that you can't directly show that the GP's selected were those with the usual high rate of off-target thyroxine patients, I think it important to show that you haven't selected GP's who are more likely to follow guidelines and stay up to date with new information.

Response: Participants responded to advertisements sent out through the clinical research network. This is now highlighted under 'participants' in the methods section: "Participants were recruited through the NIHR Clinical Research Network: North East and North Cumbria and local professional networks via mail and email, and participants were given the option to contact the researcher if they wanted to take part."
Unfortunately it is unknown whether the GPs that took part in our study were more likely to follow guidelines than others. This has been highlighted in the 'Strengths and weaknesses' section of the discussion section: "Additionally, within our GP sample it was unknown whether those that took part in our study were more likely to follow guidelines and regularly review patients, thus being less likely to have higher numbers of thyroxine treated patients with abnormal TSH than those who did not take part in our study."

3) Summary & Implications
I feel the summary could be more concise with clear statement of the main findings and implications of these for clinical practice/health policy.

Response: To address this comment and make the 'principal findings' and 'implication for future research and practice' sections more concise, reference to the main themes perceived health professional control, perceived health professional responsibility and attitudes to thyroid control have been included where appropriate.

The authors can confirm that they have proof read the manuscript. The manuscript has been checked for grammatical errors.

We look forward to hearing from you.

VERSION 2 – REVIEW

REVIEWER	Jim Parle University of Birmingham, UK
REVIEW RETURNED	20-Dec-2017

GENERAL COMMENTS	issues appear to al have been addressed Except I cannot see the Juarez reference in the paper and (with apologies if you think this is inappropriate): why not quote our cross-sectional study showing no link between SCH and cog function (Lesley M Roberts. PhD*, Helen Pattison. PhD**, Andrea Roalfe, MSc*, Jayne Franklyn. MD, PhD***, Sue Wilson PhD*, F D Richard Hobbs, FRCP*, James V Parle. MD*. Is subclinical thyroid dysfunction in the elderly associated with depression or cognitive dysfunction? Annals of Internal Medicine, 2006; 145 (8): 573-581) or our RCT showing no significant impact on cognition of Rx with T4 in SCH (J Parle, L Roberts, S Wilson, H Pattison, A Roalfe, M S Haque, C Heath, M Sheppard, J Franklyn, FDR Hobbs.

	A Randomized Controlled Trial of the effect of thyroxine replacement on cognitive function in community-living elderly subjects with sub-clinical hypothyroidism: the Birmingham Elderly Thyroid Study. J. Clinical Endocrinology and Metabolism, 2010 95: 3623-3632; doi:10.1210/jc.2009-2571). It does feel as if you have chosen papers to back up an assertion of cog dysfuncion rather than accept the uncertainty around the issue!

REVIEWER	Brooke Nickel University of Sydney, Australia
REVIEW RETURNED	14-Dec-2017

GENERAL COMMENTS	The authors have adequately addressed the concerns raised in the prior review and the manuscript is improved. I have no further comments to add.

REVIEWER	Don McLeod QIMR Berghofer Medical Research Institute, Australia
REVIEW RETURNED	02-Jan-2018

GENERAL COMMENTS	I have no further comments or questions.

REVIEWER	Vijay Panicker Sir Charles Gairdner Hospital, Perth, Australia
REVIEW RETURNED	02-Jan-2018

GENERAL COMMENTS	I am happy with the response to my comments and that publication is now appropriate