PEER REVIEW HISTORY

ARTICLE DETAILS

TITLE (PROVISIONAL)	Acceptability to patients of screening disposable transnasal endoscopy: qualitative interview analysis
AUTHORS	McGoran, John; Bennett, Andrea; Cooper, Joanne; De Caestecker, John; Lovat, Laurence; Guha, Neil; Ragunath, Krish; Sami, Sarmed

VERSION 1 – REVIEW

REVIEWER	Simon Panter South Tyneside and Sunderland NHS Foundation Trust, UK Consultant for Triple Endoscopy Inc, USA
REVIEW RETURNED	06-Apr-2019

GENERAL COMMENTS	An important study addressing TNE acceptability. Although in a limited screening population it adds to the literature.

REVIEWER	Francesca Pesola King's College London, UK
REVIEW RETURNED	18-Apr-2019

GENERAL COMMENTS	The paper reports results from a qualitative study exploring the expectations, experiences and tolerability of high-risk patients undergoing both a transnasal endoscopy (TNE) TNE and endoscopy on the same day. The analyses are carried out a sub-study of a larger clinical trial, which compares endoscopy with TNE (referred to as EG scan in the protocol). Overall, the manuscript could benefit with some additional information and clarifications. The authors need to stress how the study adds to research in the area and the potential clinical implications. More quotes to support the 4 themes may be useful for the reader. The introduction could summarise the expected clinical benefit of using TNE for screening. it would be good if the authors could clarify whether TNE would be used for screening in the general population or high risk groups as per sample involved (e.g. patients with BO). Participants and recruitment. The authors should make it clearer how the subsample for the qualitative study was selected. Is there a chance of a selection bias? Data collection. It would be useful to have more information on how guided the interview was. Were there specific questions used? Were they asked to all participants? How were questions defined/identified? Results. The majority of the sample (17/23) were men. This should be listed as a limitation and discussed in the discussion. Inclusivity in one's healthcare. The findings are not necessary generalisable as patients are taking part in research they have agreed to take part to. They may feel differently if TNE were a

	standard procedure like endoscopies and they had less decisional power. This above also applies to the sense of altruism and reciprocity. Conclusions. The authors state the themes align with the quantitative work but do not describe how and whether the qualitative findings shed any additional light.

REVIEWER	Kate M Guthrie The Miriam Hospital USA
REVIEW RETURNED	18-Jun-2019

GENERAL COMMENTS	While the study's objective is admirable and could be of great benefit to improving patients experience in health care, the manuscript does not adequately present methods nor results that a reader can use to evaluate the credibility of the conclusions. Methods are not clearly delineated; in particular, a description of the data analysis plan and procedures is entirely lacking. It is unclear from the manuscript how clarity in analysis comparisons between devices/procedures could be implemented with credibility. In addition, it is not clear why the interviews were stopped at 20, leaving less than confident interview numbers in each "group." It is not clear how saturation was defined or audited throughout: it may be a lack of clarity in terms but I am unclear how it is possible to saturate on themes DURING the process...? On what information was saturation deemed achieved? The term "tolerability" is seen once, but that is different than acceptability; in addition, acceptability is not defined. A presentation of the interview questions would be helpful to better evaluate the study, as would a listing of nodes/codes. How was the data reduced? The authors note that "checks were made" (re rigour) but what exactly does that mean and how was it done? What was the theoretical framework employed in the study? The range of data comprising themes was not noted. Inconsistency between use of structured versus semi-structured: which was it? What was the "iterative process"? Only 6 participants were female: this was not discussed, nor is there any presentation of results by sex. Further, there is no mention of other sociodemographics that may be relevant. The lack of ordering of procedures by patient is not explained - it does not seem possible to make the comparisons as a result of not ordering. Rationale for "purposive sampling" also not presented. Conclusions, therefore, are less than credible. "Reflexivity" should be explained in the context of the current study - and the fact that this study is apparently part of a larger study; as reflexivity is a standard practice in qualitative work, the important piece is what elements were monitored, how, and why. Themes "identified" are not all applicable to what will be "standard of care" (SoC) should the TNE move forward as a screening procedure. The "inclusivity" theme and the "altruism" theme are true of research, but not SoC: Therefore, are they really relevant to the research question at hand? Minimally, this should be better presented and discussed. Quotes are not always illustrative evidence of the point the authors are trying to make. It seems from the manuscript that procedures for TNE differed from those for C-OGD with respect to all the aspects of

	the patient experience as part of the research. This lack of similar attention to patient needs could inextricably altered the results: first, the authors don't acknowledge this, and second, it makes it impossible to truly evaluate whether what results are presented are a function of the actual procedures or the way the protocol was run. The sedation (13 of 23) vs not is also not fully explained, articulated or considered with respect to its impact on results. what is actually being evaluated if a patient has experienced a procedure while sedated? on another note, it was not explained how previous procedures may have impacted patient evaluations. Discussion does not address key points in the design or results that should be addressed: there are entire elements of the patient experience not accounted for, and while the authors note that there are quantitative acceptability outcomes, as well, they are not presented or in any way summarized to allow the reader to come to their own conclusions. Overall, there is not enough objectively presented data or summarizations and illustrative quotes to justify the conclusions.

VERSION 1 – AUTHOR RESPONSE

Reviewer: 1
Reviewer Name: Simon Panter
Institution and Country: South Tyneside and Sunderland NHS Foundation Trust, UK
Please state any competing interests or state ‘None declared’: Consultant for Triple Endoscopy Inc, USA

Please leave your comments for the authors below
An important study addressing TNE acceptability. Although in a limited screening population it adds to the literature.

The high-risk screening populations for BO and OVs are indeed relatively small. There is some international consensus that alternative methods to conventional OGD are required.

Reviewer: 2
Reviewer Name: Francesca Pesola
Institution and Country: King's College London, UK
Please state any competing interests or state ‘None declared’: N/A

Please leave your comments for the authors below

The paper reports results from a qualitative study exploring the expectations, experiences and tolerability of high-risk patients undergoing both a transnasal endoscopy (TNE) TNE and endoscopy on the same day. The analyses are carried out a sub-study of a larger clinical trial, which compares endoscopy with TNE (referred to as EG scan in the protocol).

Overall, the manuscript could benefit with some additional information and clarifications.

The authors need to stress how the study adds to research in the area and the potential clinical implications.

We have added the following sentence to the introduction to answer how this adds to research:

‘A deeper analysis of this could add to the clinical community’s understanding of patients’ sense-making process and enhance human factors in screening or surveillance such as recruitment and retention.’

And the clinical implications:

‘TNE as a potential screening and surveillance tool.’,

The added parts highlight the clinical benefit of qualitative analysis.

More quotes to support the 4 themes may be useful for the reader.

We have added more quotes which are highlighted.

The introduction could summarise the expected clinical benefit of using TNE for screening.

As well as newly added segments, please note the second and third paragraphs in Introduction. The design of the equipment allows this examination to take place in different locations and scenarios.

it would be good if the authors could clarify whether TNE would be used for screening in the general population or high risk groups as per sample involved (e.g. patients with BO).

We have added clarifications to paragraph 3 of the introduction. The research here supports the claim that it is a promising technology for screening of these conditions in the future but we are not proposing its application in a screening programme at this point in time. Rather we show here the degree to which acceptability can be examined.

Participants and recruitment. The authors should make it clearer how the subsample for the qualitative study was selected. Is there a chance of a selection bias?

Please see the section for additions providing greater clarification. Purposeful sampling was employed, and although selection bias is a concept generally reserved for quantitative research, care was taken in line with COREQ criteria that patient choices were made to explore the research aims, with no conflicts of interest noted on the researchers’ behalf on the issue.

Data collection. It would be useful to have more information on how guided the interview was. Were there specific questions used? Were they asked to all participants? How were questions defined/identified?

Thank you. We have expanded on this in the relevant section.

Results. The majority of the sample (17/23) were men. This should be listed as a limitation and discussed in the discussion.

The gender and other demographics of participants do not adversely impact on the thematic analysis, particularly as the target population is predominantly male. We have expanded on this in the first paragraph of the Results section.

Inclusivity in one's healthcare. The findings are not necessarily generalizable as patients are taking part in research, they have agreed to take part to. They may feel differently if TNE were a standard procedure like endoscopies and they had less decisional power. This above also applies to the sense of altruism and reciprocity.

We acknowledge this and have elaborated on it in the Discussion and the 'Limitations' section.

Conclusions.
The authors state the themes align with the quantitative work but do not describe how and whether the qualitative findings shed any additional light.

We have expanded on this in the Conclusions section.

Reviewer: 3
Reviewer Name: Kate M Guthrie
Institution and Country: The Miriam Hospital, USA
Please state any competing interests or state 'None declared': None declared

Please leave your comments for the authors below

While the study's objective is admirable and could be of great benefit to improving patients experience in health care, the manuscript does not adequately present methods nor results that a reader can use to evaluate the credibility of the conclusions.

Methods are not clearly delineated; in particular, a description of the data analysis plan and procedures is entirely lacking.

We have expanded on this by rewriting the data collection segment and adding further commentary to the data analysis segment.

It is unclear from the manuscript how clarity in analysis comparisons between devices/procedures could be implemented with credibility. In addition, it is not clear why the interviews were stopped at 20, leaving less than confident interview numbers in each "group."

The interviews were stopped after the 23rd interview. Based on similar studies that have been referenced we expected to require at least 20 participants. Once the cohort of 23 were interviewed, human factors through team discussion and analysis using NVivo concluded that data saturation was reached. We have added to the Participants and Recruitment section. As regards the comparisons between devices we have expanded on this at the beginning of the Discussion section.

It is not clear how saturation was defined or audited throughout: it may be a lack of clarity in terms but I am unclear how it is possible to saturate on themes DURING the process...?
On what information was saturation deemed achieved?

As mentioned above, the interviews were stopped after the 23rd interview and numbers required estimated based on similar preceding research. Additions have been made to the methods section, in particular the Participants and Recruitment section.

The term "tolerability" is seen once, but that is different than acceptability; in addition, acceptability is not defined.

Tolerability refers to quantitative comfort scores, whereas acceptability is a broader term that deepens understanding of willingness to undergo the procedure. We have clarified this in paragraph 4 of the Introduction.

A presentation of the interview questions would be helpful to better evaluate the study, as would a listing of nodes/codes.

We have added the interview schedule as an appendix. If necessary for inclusion, may we ask for clarification on the listing of nodes/codes please? In the analysis we had a large bank of information from which themes were derived and pertinent quotes inserted in the Results section.

How was the data reduced? The authors note that "checks were made" (re rigour) but what exactly does that mean and how was it done?

We have clarified the section on rigour. we have also expanded a bit on the data analysis section to explain how the themes were arrived at. All transcripts were added to the NVivo programme and nodes/codes gathered as standard. For this, we were aware of the need to present a substantial methodological approach while maintaining the interest of a reader who may not be overly familiar with qualitative methodology.

What was the theoretical framework employed in the study?

This was considered throughout the research. On reflection, the main influence on things was the community of inquiry concept. We have inserted the following paragraph into the Discussion.

“The community of inquiry theory is a concept that unites the themes arrived at in this study. The background is that TNE is shown to be accurate and safe in the delivery of BO and OV assessment but the problem remains that not enough is deeply understood about its acceptability to a population. To explore the challenges and advantages offered by TNE in the screening of BO and OVs, participants from different backgrounds and for different reasons underwent the same procedure. Their individual perspectives were all considered and conclusions that reflect this community’s account were arrived at, which will hopefully contribute to improved delivery of patient care. For the purposes of this study all of the accounts were communicated through the researchers but they are reflective of the conversations that may exist in the greater population.”

The range of data comprising themes was not noted.
Inconsistency between use of structured versus semi-structured: which was it? What was the "iterative process"?

Structured has been mentioned once, in reference to other means of collecting data. The approach taken for this research was through semi-structured interviews. We agree the mention of the iterative process seems out of place and confusing so have removed. (It referred to the iteration of the questions in the semi-structured interviews.)

Only 6 participants were female: this was not discussed, nor is there any presentation of results by sex. Further, there is no mention of other sociodemographics that may be relevant.

The gender and other demographics of participants was not thought to require separation of accounts, while of course their backgrounds and potential unique perspectives were taken into consideration when analysing. If enacted in a screening programme, the approach to patients will be uniform and we believe the diversity of the participants is therefore a strength. With regard to a female minority, the two conditions (BO and OV) are predominantly male problems and a 50/50 gender

balance in a quantitative research study would be erroneous. In a qualitative study we think this holds truth too.

Paragraph 1 of the Results section and paragraph 2 of the Discussion have been added to clarify this point.

The lack of ordering of procedures by patient is not explained - it does not seem possible to make the comparisons as a result of not ordering.

The ordering of procedures was that patients had TNE followed by C-OGD on the same day. However, the aim of the study is not to compare the two endoscopic procedures, rather to assess TNE's acceptability to a patient population that has experienced both types.

Rationale for "purposive sampling" also not presented. Conclusions, therefore, are less than credible.

We have expanded on this in the Methods section. Most patients were seen as eligible however and although the sampling was purposeful it did not have a significant bearing on patient selection.

"Reflexivity" should be explained in the context of the current study - and the fact that this study is apparently part of a larger study; as reflexivity is a standard practice in qualitative work, the important piece is what elements were monitored, how, and why.

Thank you. We have added to the 'Rigour' section.

Themes "identified" are not all applicable to what will be "standard of care" (SoC) should the TNE move forward as a screening procedure. The "inclusivity" theme and the "altruism" theme are true of research, but not SoC: Therefore, are they really relevant to the research question at hand? Minimally, this should be better presented and discussed.

The research was taken to gather an idea of the acceptability of TNE compared with C-OGD in any scenario. The use of it as a screening tool was therefore not a restrictive factor, rather a suggestion by the authors that it could be useful in this context. The themes of inclusivity and altruism we believe, go beyond that of the research subject to a patient community and we believe they relevant elements of a standard of care. This fits with the theoretical framework.

Further detail has been added to paragraph 3 of the Discussion.

Quotes are not always illustrative evidence of the point the authors are trying to make. It seems from the manuscript that procedures for TNE differed from those for C-OGD with respect to all the aspects of the patient experience as part of the research. This lack of similar attention to patient needs could inextricably altered the results: first, the authors don't acknowledge this, and second, it makes it impossible to truly evaluate whether what results are presented are a function of the actual procedures or the way the protocol was run.

We have added further quotes which I hope better illustrate the themes that arose. We would enjoy greater clarification on your point with regard to lack of attention to patient needs. It is acknowledged that the research setting is not 'real-world' but is in our opinion as clinicians an acceptable surrogate for gathering accounts on the patient experience. Moreover, the procedures for TNE and C-OGD and the protocol were run as they would occur in real life practice, so TNE was performed in an outpatient clinic environment using the appropriate techniques and C-OGD was performed in an endoscopy unit using techniques as per standard clinical practice. Therefore, we believe that the results presented

are likely to represent a function of the actual procedures. Finally, our main objective is to evaluate the patients experiences of clinic-based TNE rather than strictly compare it to C-OGD.

The sedation (13 of 23) vs not is also not fully explained, articulated or considered with respect to its impact on results. what is actually being evaluated if a patient has experienced a procedure while sedated?

Our focus is on exploring the patients' experiences of unsedated TNE, while they can use C-OGD (either sedated or unsedated) as a direct comparator. Sedation did have an impact on attitudes and is reflected in the results and discussion. We have added more quotes. Sedation was a patient choice and not an intervention made by this study in any way.

on another note, it was not explained how previous procedures may have impacted patient evaluations.

On that day the procedures were standard i.e. TNE then (sedated or unsedated C-OGD). The Results section (including added quotes) gives insights into patients' previous endoscopic procedures. We have clarified in the article the focus on the patient experience of TNE.

Discussion does not address key points in the design or results that should be addressed: there are entire elements of the patient experience not accounted for, and while the authors note that there are quantitative acceptability outcomes, as well, they are not presented or in any way summarized to allow the reader to come to their own conclusions. Overall, there is not enough objectively presented data or summarizations and illustrative quotes to justify the conclusions.

Within the confines of the word count we have added some more detail to the Discussion. The quantitative acceptability outcomes from the 'parent' study are already published and considered independent of this, as mentioned in the reflexivity portion. We would be willing to add further detail linking results to the Discussion if thought necessary in highlighting the conclusion.

VERSION 2 – REVIEW

REVIEWER	Francesca Pesola King's College London, UK
REVIEW RETURNED	29-Jul-2019

GENERAL COMMENTS	The authors have exhaustively addressed my queries.

REVIEWER	Kate M Guthrie Brown Medical School, USA
REVIEW RETURNED	14-Sep-2019

GENERAL COMMENTS	TNE review https://mc.manuscriptcentral.com/bmjopen?URL_MASK=c95ff59ab1314c109896d9da7c02c749 I thank the authors for their responses, which have improved the manuscript. I also learned a lot from your responses, so thank you for that. I do still have a few more requests and suggestions to help the manuscript both conform to conventional reports of qualitative data and to clarify a few remaining issues.

	The authors stated in their response that they had added the interview schedule as an appendix, but I am unable to find it. Again, knowing the questions asked (at least the major questions and prescribed probes) would allow a better evaluation of the study. SUGGESTION: a text box "figure" with a brief outline of the interview schedule. General suggestion: search for "at risk" and decide whether instances of "at risk" should be changed to "high risk" as per your previous revision response. I believe at least some of them should… Ordering effects: Given the study is done, the ordering effect (TNE first) cannot be undone, but it should be discussed. There are few patient experience studies that do not note ordering effects: this should be discussed. Regarding further explanation of the females in the sample: minimally, I would suggest that the analysts do due diligence and review the data, comparing presence and range of data and themes within sex, if nothing to be able to soundly conclude that there are no differences between males' and females' experiences. Paragraph 3 of discussion: consider adding wording that inclusivity and altruism easily function within research but may also apply to public health clinical settings. (As I consider this discussion, I wonder if some of my position isn't a function of the American health system I work in, and that the culture of other countries' systems impact this differently. Forgive me if I am speaking from a point of bias, but know that others will read it similarly, so it might be best to address.) Abstract/Participants: consider adding "(patient choice)" following "… without sedation" so that this element is clear in the abstract alone. Abstract/Conclusions: please include patient preference between TNE and sedated C-OGD, as well. [and in Results: please note whether the themes themselves differed in sedated vs non-sedated patients. I am wondering whether the range of experience differed, for instance, with respect to the comfort /convenience theme.] Abstract/Strengths… 3rd bullet: one does not perform thematic analyses on a group of patients, but on data. Please reword 3rd bullet, last sentence: delete; this is not knowable from the data. 4th bullet, 1st sentence: not necessary 5th bullet: fits with 3rd bullet. Add to 3rd bullet and then, "however, transferability the findings… MAY be limited." Introduction/paragraph 1: please add "predominantly" ahead of "…white, male, age 55 …" as not all patient sot be screened ARE white men. Also, consider adding the prevalence in women to (later) justify your sample. Introduction/paragraph 5: in the added sentence, consider rewording as follows: "A deeper analysis of [these experiences] could add..". Also, the term "sense-making" is not familiar to mean – I am unclear whether you mean decision-making or meaning-

	making, or something different. Please consider rewording. Methods/Study Design/1st paragraph: - insert “semi-structured” between face-to-face and interviews in the 1st sentence. - Likely just need either Figure 1 or Box 1, but not both. - last sentence: clarify to “… having undergone [both] endoscop[ic procedures] at that location”. A small but important detail of patient experience. Methods/Study Design/2nd paragraph: - assuming you meant that consent was sought [prior to] each interview; or otherwise reword to clarify that consent was obtained before any study procedures were initiated. Figure 1: a close-parenthesis is needed after “Hydrochloride 0.5%” Methods/Participants…: - the authors state, “The experiences of all 47 patients were CONSIDERED VALID FOR ANALYSIS of acceptability,” Please explain on what grounds this was determined. What made their experience “valid”? Frankly, I continue to wonder about those patients who chose sedated C-OGD: there is no acknowledgement of which experiences those particular patients could speak to and which they could not, given the sedation. This feels an important situation to contextualize. - the authors state, “consenting to involvement in an interview-based research study at a time following the day they underwent endoscopic examination.” I have 2 different issues with this sentence/procedure. First, please cite the range of time (in days): in other words, what was the range of days that passed between the procedure and the invitation to participate in this study? This is an important consideration with respect to recall of events and experiences, especially sensory experiences, which, unless extremely salient, are lost to recall within a very short amount of time. Second, this sentence brings to light the necessity to discuss self-selection bias as a potential limitation (in addition to the potential recall bias just discussed). There is a likelihood that those who would consent to this study had a preponderance of positive (versus negative) experiences with the procedure(s). These should be discussed in limitations. - final paragraph, added sentence: The description of the process for concluding saturation is still quite opaque. What does “it was posed” mean? What was the actual process? To have used thematic analysis to come to this conclusion would have required a significant amount of time between the 23rd and what could have been the 24th interview. Please explicate how saturation was determined. Methods/Data Collection: - how were interviewers trained? What credentials/experience in qualitative facilitation informed their appointment? - Suggest including the citation(s) for the larger study at the end of the added verbiage. - 2nd paragraph: seems like the information here would flow better if integrated with the 2nd sentence in above paragraph. - 3rd paragraph: please provide range for duration of interviews. - please describe how transcript accuracy to the audio was verified and corrected if necessary. Methods/Data Analysis:

	- Thematic analysis inherently includes the voices of all respondents. Why then is it that only RECURRENT words, phrases and sentiments noted in the initial review of the transcripts? Also, thematic analysis does not conventionally rely on, and actually proscribes, word counts: if the authors relied on high frequency utterances to identify themes, a thematic analysis is not possible. I hope that I am misunderstanding the authors' intentions with this description. The Data Analysis section, therefore, requires substantive revision. Methods/Rigour: - The authors state, "Our transcripts were analysed for deviant cases, which were noted but acknowledged not to represent the majority of participants' experiences." If the authors' intentions were to ONLY report the majority experiences, this should be duly acknowledged, so that readers would be aware and not expect a conventional results presentation (i.e., results which report on and characterize the range of experiences, i.e., from negative to positive, etc). - "Reflexivity" should be defined and contextualized within the context of this particular study and its procedures. - The authors state, " The researchers are aware of the requirements for a good screening tool…" Please describe how the potential for bias from this perspective – as well as the fact that 2 of the authors report conflicts of interest with the device developer/manufacturer – were mitigated. This is especially critical in qualitative work. Methods/Patient and Public Involvement: - For clarity, I suggest adding a phrase that reminds the reader that the patients were part of a larger trial. Results: - Given the reported purposive sampling strategy, I continue to hold that it is critical that Table 1 (or another table) present the equipoise of male:female and BO:OV:dyspnea patients in the analyzed sample. A footnote can explain the disproportionality of male:female patients as reflected in the site's patient population. A footnote – or incorporated into the table – should be sedation status. - 1st paragraph, last sentence: please describe how sedation impacted the ability of patients to discuss certain aspects of the procedure, hence giving the reader an understanding of the data that does NOT exist for those patient transcripts. - 2nd paragraph, 1st sentence: delete the words "dominant" and "careful", as there is no way for those terms to be understood by the reader. The phrase" using NVivo software" can also be deleted as the software is not responsible for identifying theme. - CRITICAL: for all quotes, please add descriptors of the patient to whom the quote is attributed, as is conventional in presentations of qualitative results. Suggestions might be: sex, age, BO/OV/dyspnea, sedation/not (as applicable to C-OGD). Thus, a quote might look like: "I could have said no. But I came looking for help." (male, 57, BO) - Inclusivity theme: the quotes in lines 7 and 8 are contrary to autonomy per se, while the quotes in lines 10 and 11 do not. Reflective of my request for the authors to articulate the range of experiences, these sentence provide the example: the challenge is in presenting the data as a range in the theme of inclusivity and autonomy. It may be resolved with a mere reordering of the sentences, as those beginning, "On attending,…" seem out of place.

	That is, the paragraph as a whole does not flow well to present the story of the data, in range, scope, or in the comparison of procedures. Please consider rewording. [From protocol: Mapping and interpretation. At this stage we will develop charts to illustrate and define the concepts, and to map the RANGE [my capitalization] and nature of experiences/expectations that participants describe, including any recommendations they make - Comfort theme: 2nd sentence: how, exactly, was the VAS "revisited" at interview? Did the interviewer have access to patient VAS data? Please explain and provide rationale. - Comfort theme, 4th sentence: "A sizeable number…" how many of those interviewed underwent C-OGD with sedation? (this could also be addressed in the Table requested earlier.) Discussion - Paragraph 1, line 2: replace "on this level" (which is vague) with "for screening of disease" or the equivalent. - Delete the first part of the 2nd sentence, and start it at "The aim…" then add to the end of the sentence something like, "but to consider the acceptability of TNE specifically for outpatient screening in high risk patients." Or the equivalent… - "The themes identified in this study SUPPORT…" ('encourage' seems to anthropomorphize the data) - 2nd paragraph: "… unites the themes IDENTIFIED in this study" is a more appropriate wording. - what do the authors mean by "participants from different backgrounds"? the only demographics noted are age and sex and these are merely listed not discussed in context with the data – though I would suggest they should be. - I continue to not have a sense of range of experience within each theme. And I still wonder about how the women might have different experiences – if only as a function of body dimensionality (in combination with size of endoscope. These gaps continue to minimize my enthusiasm for the paper and the results presented, especially if the device is to be used on a population level. -3rd paragraph, added text: some of these patients clearly, as per the quotes, did not feel they had a choice in having/not having the procedure. I am not sure I understand the point the authors are trying to make here… - reword: "The inability of the TNE to allow biopsies WAS A recurring concern" - at end of this paragraph the authors might go on to say that a thinner endoscope could be considered in redesigning the TNE used here: this is actually a reasonable implication. - I would also be interested in a discussion of the decision-making for using TNE for screening in, as yet, undiagnosed patients, versus patients with a diagnosis but in need of repeat screenings or follow-up. -the first 5+ lines of the 5th paragraph might be better placed in the introduction…? Conclusions: - Replace "reliable" with "credible" - reword: "… and highlight the strengths of qualitative methods to optimize clinical care." - reword: "…suggesting that unsedated TNE is POTENTIALLY acceptable, and MAY reflect…"

VERSION 2 – AUTHOR RESPONSE

I thank the authors for their responses, which have improved the manuscript. I also learned a lot from your responses, so thank you for that. I do still have a few more requests and suggestions to help the manuscript both conform to conventional reports of qualitative data and to clarify a few remaining issues.

The authors stated in their response that they had added the interview schedule as an appendix, but I am unable to find it. Again, knowing the questions asked (at least the major questions and prescribed probes) would allow a better evaluation of the study. SUGGESTION: a text box "figure" with a brief outline of the interview schedule.

This is now Box 1.

General suggestion: search for "at risk" and decide whether instances of "at risk" should be changed to "high risk" as per your previous revision response. I believe at least some of them should…

Thank you, we have done this.

Ordering effects: Given the study is done, the ordering effect (TNE first) cannot be undone, but it should be discussed. There are few patient experience studies that do not note ordering effects: this should be discussed.

This has been added to 'strengths and limitations'.

Regarding further explanation of the females in the sample: minimally, I would suggest that the analysts do due diligence and review the data, comparing presence and range of data and themes within sex, if nothing to be able to soundly conclude that there are no differences between males' and females' experiences.

We have expanded on the proportion of men and women, with further details added to the table in Results. Through this analysis and from searching of other literature determining quantitative measures, there is no evidence of substantial differences in experiences between men and women. (On an anecdotal level, I (JM) have only encountered a gender specific difference in experience when caring for a transnasal endoscopy patient.)

Paragraph 3 of discussion: consider adding wording that inclusivity and altruism easily function within research but may also apply to public health clinical settings. (As I consider this discussion, I wonder if some of my position isn't a function of the American health system I work in, and that the culture of other countries' systems impact this differently. Forgive me if I am speaking from a point of bias, but know that others will read it similarly, so it might be best to address.)
We have clarified this in the same paragraph. There may be something in the difference between the healthcare systems, with the interaction between patients and doctors perhaps being less transactional in the UK.

Abstract/Participants: consider adding "(patient choice)" following "… without sedation" so that this element is clear in the abstract alone.

We have enacted this.

Abstract/Conclusions: please include patient preference between TNE and sedated C-OGD, as well.

We have done this.

[and in Results: please note whether the themes themselves differed in sedated vs non-sedated patients. I am wondering whether the range of experience differed, for instance, with respect to the comfort /convenience theme.]

The themes did not differ. Further expansion is made on the experiences and acknowledgement of the memory impairment caused by sedation however comfort level includes features beyond the C-OGD itself and lots of data is available for TNE. This gives an insight without the need for direct comparison.

Abstract/Strengths…
3rd bullet: one does not perform thematic analyses on a group of patients, but on data. Please reword

This has been enacted. Thank you.

3rd bullet, last sentence: delete; this is not knowable from the data.

Done

4th bullet, 1st sentence: not necessary

Removed

5th bullet: fits with 3rd bullet. Add to 3rd bullet and then, “however, transferability the findings… MAY be limited.”

Changes have been made

Introduction/paragraph 1: please add “predominantly” ahead of “…white, male, age 55 …” as not all patient sot be screened ARE white men. Also, consider adding the prevalence in women to (later) justify your sample.

We have done this, thank you.

Introduction/paragraph 5: in the added sentence, consider rewording as follows: “A deeper analysis of [these experiences] could add..”. Also, the term “sense-making” is not familiar to mean – I am unclear whether you mean decision-making or meaning-making, or something different. Please consider rewording.

We have reworded this in the introduction.

Methods/Study Design/1st paragraph:
- insert “semi-structured” between face-to-face and interviews in the 1st sentence. Done
- Likely just need either Figure 1 or Box 1, but not both. We have kept the figure and removed the box
- last sentence: clarify to “… having undergone [both] endoscop[ic procedures] at that location”. A small but important detail of patient experience. Done

Methods/Study Design/2nd paragraph:
- assuming you meant that consent was sought [prior to] each interview; or otherwise reword to clarify that consent was obtained before any study procedures were initiated.

This has been clarified. Consent was of course obtained before interviews took place and was SOUGHT for interviews at the same time as that for the quantitative study, with rechecking after. This has been clarified, thank you.

Figure 1: a close-parenthesis is needed after “Hydrochloride 0.5%”
Methods/Participants…:
- the authors state, “The experiences of all 47 patients were CONSIDERED VALID FOR ANALYSIS of acceptability,” Please explain on what grounds this was determined. What made their experience “valid”? Frankly, I continue to wonder about those patients who chose sedated C-OGD: there is no acknowledgement of which experiences those particular patients could speak to and which they could not, given the sedation. This feels an important situation to contextualize.

For being considered valid for analysis we have added the phrase “as they reflected a suitably healthy and communicative cohort”. Further context is added to regarding sedation in the results and discussion when considering the more global interpretation of comfort, as discussed above. In addition we have added further detail, based on the cases that offered such information, of the motives for choosing sedation. We have mentioned that sedated and unsedated are considered equally valid as they reflect the population.

- the authors state, “consenting to involvement in an interview-based research study at a time following the day they underwent endoscopic examination.” I have 2 different issues with this sentence/procedure. First, please cite the range of time (in days): in other words, what was the range of days that passed between the procedure and the invitation to participate in this study? This is an important consideration with respect to recall of events and experiences, especially sensory experiences, which, unless extremely salient, are lost to recall within a very short amount of time.

This has been reworded to “Recruitment approaches for the quantitative and qualitative studies were done in parallel, with individuals approached for their consent to participate in an interview-based research study at the same time as for the quantitative study. They were then followed up within four weeks to ensure willingness to proceed with interview. Only the three individuals mentioned above withdrew consent or were withdrawn because of ill health.”

Second, this sentence brings to light the necessity to discuss self-selection bias as a potential limitation (in addition to the potential recall bias just discussed). There is a likelihood that those who would consent to this study had a preponderance of positive (versus negative) experiences with the procedure(s). These should be discussed in limitations.

Everyone was eligible and agreeable to interview except 3 who were sick or deceased. This has been clarified in the manuscript.

- final paragraph, added sentence: The description of the process for concluding saturation is still quiteopaque. What does “it was posed” mean? What was the actual process? To have used thematic analysisto come to this conclusion would have required a significant amount of time between the 23rd and whatcould have been the 24th interview. Please explicate how saturation was determined.

Thank you. Further information has been provided to explain the rationale and the process.

Methods/Data Collection:
- how were interviewers trained? What credentials/experience in qualitative facilitation informed their appointment?

Further information has been provided in the same paragraph.

- Suggest including the citation(s) for the larger study at the end of the added verbiage.

Done

- 2nd paragraph: seems like the information here would flow better if integrated with the 2nd sentence in
above paragraph. Done
- 3rd paragraph: please provide range for duration of interviews. Done
- please describe how transcript accuracy to the audio was verified and corrected if necessary. We have expanded on this

Methods/Data Analysis:
- Thematic analysis inherently includes the voices of all respondents. Why then is it that only RECURRENT words, phrases and sentiments noted in the initial review of the transcripts? Also, thematic analysis does not conventionally rely on, and actually proscribes, word counts: if the authors relied on high frequency utterances to identify themes, a thematic analysis is not possible. I hope that I am misunderstanding the authors' intentions with this description. The Data Analysis section, therefore, requires substantive revision.

To clarify, the themes were arrived at after careful analysis of the accounts of interview participants. Commonly occurring words, phrases (e.g. "gagging", "experience of (endoscopy)" AND SENTIMENTS were identified with the help of the software but they alone did not dictate the arrival at the 4 themes, and other accounts were read and considered. We hope the further clarification helps.

Methods/Rigour:
- The authors state, "Our transcripts were analysed for deviant cases, which were noted but acknowledged not to represent the majority of participants' experiences." If the authors' intentions were to ONLY report the majority experiences, this should be duly acknowledged, so that readers would be aware and not expect a conventional results presentation (i.e., results which report on and characterize the range of experiences, i.e., from negative to positive, etc).

Deviant cases were taken into consideration and the previous wording was unclear. Thank you, we have clarified.

- "Reflexivity" should be defined and contextualized within the context of this particular study and its procedures.

Done

- The authors state, " The researchers are aware of the requirements for a good screening tool…" Please describe how the potential for bias from this perspective – as well as the fact that 2 of the authors report conflicts of interest with the device developer/manufacturer – were mitigated. This is especially critical in qualitative work.

We have done this, expanding on the place of JM in the project.

Methods/Patient and Public Involvement:
- For clarity, I suggest adding a phrase that reminds the reader that the patients were part of a larger trial.

We have done this.

Results:
- Given the reporteds trategy, I continue to hold that it is critical that Table 1 (or another table) present the equipoise of male:female and BO:OV:dyspnea patients in the analyzed sample. A footnote can explain the disproportionality of male:female patients as reflected in the site's patient population. A footnote – or incorporated into the table – should be sedation status.

Done

- 1st paragraph, last sentence: please describe how sedation impacted the ability of patients to discuss
certain aspects of the procedure, hence giving the reader an understanding of the data that does NOT exist for those patient transcripts.

This has been clarified and as above, the limitations of the memory loss do not impede the arrival to 'comfort level and convenience' as a theme, especially with unsedated TNE and in consideration that it is reflective of the general population. The project does not seek to directly contrast C-OGD with TNE.

- 2nd paragraph, 1st sentence: delete the words "dominant" and "careful", as there is no way for those terms to be understood by the reader. The phrase" using NVivo software" can also be deleted as the software is not responsible for identifying theme.

Thank you, this has been clarified.

- CRITICAL: for all quotes, please add descriptors of the patient to whom the quote is attributed, as is conventional in presentations of qualitative results. Suggestions might be: sex, age, BO/OV/dyspnea, sedation/not (as applicable to C-OGD). Thus, a quote might look like:
"I could have said no. But I came looking for help." (male, 57, BO)

We have done this. Please note that the quotes are an illustrative example of the wider accounts of all 23.

- Inclusivity theme: the quotes in lines 7 and 8 are contrary to autonomy per se, while the quotes in lines 10 and 11 do not. Reflective of my request for the authors to articulate the range of experiences, these sentence provide the example: the challenge is in presenting the data as a range in the theme of
inclusivity and autonomy. It may be resolved with a mere reordering of the sentences, as those beginning, "On attending,…" seem out of place. That is, the paragraph as a whole does not flow well to present the story of the data, in range, scope, or in the comparison of procedures. Please consider rewording. [From protocol: Mapping and interpretation. At this stage we will develop charts to illustrate and define the concepts, and to map the RANGE [my capitalization] and nature of experiences/expectations that participants describe, including any recommendations they make

Thank you. We have clarified.

- Comfort theme: 2nd sentence: how, exactly, was the VAS "revisited" at interview? Did the interviewerhave access to patient VAS data? Please explain and provide rationale.

We have expanded on this in the first paragraph.

- Comfort theme, 4th sentence: “A sizeable number…” how many of those interviewed underwent COGD with sedation? (this could also be addressed in the Table requested earlier.)

The table has been adjusted.

Discussion
- Paragraph 1, line 2: replace “on this level” (which is vague) with “for screening of disease” or the equivalent.

Done

- Delete the first part of the 2nd sentence, and start it at “The aim…” then add to the end of the sentence something like, “but to consider the acceptability of TNE specifically for outpatient screening in high risk patients.” Or the equivalent…

Thank you, done.

- “The themes identified in this study SUPPORT…” (‘encourage’ seems to anthropomorphize the data)

Done

- 2nd paragraph: “… unites the themes IDENTIFIED in this study” is a more appropriate wording.

Done

- what do the authors mean by “participants from different backgrounds”? the only demographics noted are age and sex and these are merely listed not discussed in context with the data – though I would suggest they should be.

This has been refined. We do not have background information beyond gender and age so acknowledge that a description of ‘background’ is perhaps a stretch.

- I continue to not have a sense of range of experience within each theme. And I still wonder about how the women might have different experiences – if only as a function of body dimensionality (in combination with size of endoscope. These gaps continue to minimize my enthusiasm for the paper and the results presented, especially if the device is to be used on a population level.

We have tried to clarify the range of experiences based on the feedback you have given and appreciate this.

With regard to the experiences of women, we are unable to find in quantitative study literature for endoscopic procedures or parallel qualitative study literature for other medical procedures, any substantial difference between men’s and women’s experiences. We have found no difference between the accounts in our interview transcripts, likewise between those with different conditions and of different ages. It appears that the experience of endoscopy is mainly reliant on previous experience and individual traits. As practitioners in endoscopy we have observed no discernible difference in body dimensionality between the genders.

-3rd paragraph, added text: some of these patients clearly, as per the quotes, did not feel they had a choice in having/not having the procedure. I am not sure I understand the point the authors are trying to make here…

Thank you, we have clarified.

- reword: “The inability of the TNE to allow biopsies WAS A recurring concern”

Done, thanks.

- at end of this paragraph the authors might go on to say that a thinner endoscope could be considered
in redesigning the TNE used here: this is actually a reasonable implication.
- I would also be interested in a discussion of the decision-making for using TNE for screening in, as yet, undiagnosed patients, versus patients with a diagnosis but in need of repeat screenings or follow-up.

We have added a little into this.

-the first 5+ lines of the 5th paragraph might be better placed in the introduction…?
Conclusions:
- Replace “reliable” with “credible” Done
- reword: “… and highlight the strengths of qualitative methods to optimize clinical care.” Done
- reword: “…suggesting that unsedated TNE is POTENTIALLY acceptable, and MAY reflect…” Done