BMJ Open publishes all reviews undertaken for accepted manuscripts. Reviewers are asked to complete a checklist review form ([http://bmjopen.bmj.com/site/about/resources/checklist.pdf](http://bmjopen.bmj.com/site/about/resources/checklist.pdf)) and are provided with free text boxes to elaborate on their assessment. These free text comments are reproduced below.

## ARTICLE DETAILS

<table>
<thead>
<tr>
<th>TITLE (PROVISIONAL)</th>
<th>PRINT AND ONLINE NEWSPAPER COVERAGE OF THE LINK BETWEEN HPV AND ORAL CANCER IN THE UK: A MIXED METHODS STUDY</th>
</tr>
</thead>
<tbody>
<tr>
<td>AUTHORS</td>
<td>Dodd, Rachael; Marlow, Laura; Forster, Alice; Waller, Jo</td>
</tr>
</tbody>
</table>

## VERSION 1 - REVIEW

<table>
<thead>
<tr>
<th>REVIEWER</th>
<th>Dannielle E. Kelley</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>University of North Carolina at Chapel Hill, USA</td>
</tr>
<tr>
<td>REVIEW RETURNED</td>
<td>28-Oct-2015</td>
</tr>
</tbody>
</table>

## GENERAL COMMENTS

First, I would like to compliment the authors on their research – understanding how the general public is exposed to important health issues through mass media is a key step towards the development of health communication and interventions to promote healthy behavior. I like that the authors recognized the variation in scope and reach across newspapers, and discussed the findings according to newspaper type. I think this paper would be a useful contribution to the literature, and I hope the authors follow up with the research mentioned in the discussion.

I do have a few concerns, most of which regard the organization and reporting of the data. I will list my concerns in the order they appear in the paper:

1. In the introduction, third paragraph, the authors cite an online survey administered in the UK, US, and Australia. However, the remainder of that paragraph focuses on data from the US alone. This creates a bit of a disconnect for the reader as the study focuses on HPV related content in UK newspapers. Data from the UK would be more relevant and useful here.

2. You have eight themes, but only report one Cohen’s Kappa. Reporting the reliability statistic for each theme is crucial in content analysis as it enables replication and the possibility for standardized variables. Even if the Kappas were not sufficient for certain themes, that information should be provided in the text and the authors may offer an explanation for why some variables may not have been as reliable – without this, there is ambiguity in the findings and the field cannot advance. Please see the article below for reference:


A good example of how to present this information can be found in

-On a similar note, a description of the themes and individual codes within the themes should be provided.

3. What was the justification for only including articles with 100 words about the link between oral cancer and HPV? This seems like a difficult distinction to make because there are a variety of ways to talk about oral cancer and HPV, making this a difficult criteria to define. Also, considering the varying article lengths, 100 words may mean something different across newspaper types.

4. Regarding the discussion of peaks on page 6, the peak in 2013 is attributed to mouth cancer awareness month. To my knowledge, this is something that has been in place since at least 2009, so why is the peak attributed to this particular month in 2013? Further, are you certain the other peaks are attributable to the event mentioned? If so, I think this needs to be made clear for the reader.

5. How were the quotes in text and in the tables selected? Were these the most representative quotes of the eight themes? On page 10, paragraph 3, the authors discuss ONE newspaper that “suggested Douglas was criticized for his public honesty” followed by one quote about Zeta-Jones. What is the significance of this in terms of your results? If these isolated quotes were mentioned for each of the eight themes in a uniform fashion, it may be less of an issue, but here it almost seems this would fit better in the discussion.

6. Please be consistent with citing quotes in the text. For some quotes the authors provide the newspaper title and the date (which I think is the most useful format for the reader). Others, the authors provide a quotation number and reference a table. This format is frustrating and disrupts the flow of the paper. Also, not all quotes are cited; for example, page 10: “A few of the articles suggested that HPV transmitted through oral sex ‘could overtake tobacco and alcohol as the main risk factor in the next decade.’” But do not offer a citation – this is an important quote, would also be helpful to know the frequency in which this appeared in your sample.

7. I understand why the authors separated the tables by theme, but this information could be organized in a way that is much more useful. I would like to see the table as ONE comprehensive table organized by date – that way, we can see the progression of content through the years and in relation to news events. This will provide a much clearer picture of how HPV was discussed around these events. You can denote the theme by simply adding another column. Currently, Tables 1, 3, and 4, are limited by the organization.

8. Table 2 is a key table and ALL eight themes should be included. Why just include health themes? You have great data here, make it easier to see.

9. My final comments may not be one the authors can address in this paper as I am not going to ask for reanalysis of the data, but it seems odd that there was no mention of HPV and cervical cancer in
I understand the purpose was to look at the reported connection between HPV and head and neck cancers, however, in my perusal of such articles here in the US, the connection between HPV and cervical cancer or the prevalence among females, is often mentioned, although not the subject of the article. It seems that knowing how and if HPV in males and females was discussed in the same article would be important information for informing public health campaigns and message design to strengthen public knowledge about the implications of HPV for males and females. Finally, did the authors code for how often was the same story line was presented across issues and newspapers? This would also be useful in understanding which types of stories gain traction in the media.

**REVIEWER**
Jessie M. Quintero Johnson, PhD
University of Massachusetts Boston
Boston, MA USA

**REVIEW RETURNED**
03-Nov-2015

**GENERAL COMMENTS**
The purpose of this study, to content analyze changes in the frequency of stories about the link between HPV and oral cancer in print news, is both interesting and meaningful. As the authors suggest, the presentation of health information in mass media is likely to influence public awareness about HPV. I believe the manuscript in its present form will need the following revisions in order to strengthen both the readability and the significance of the findings:

1) Introduction

a. First and foremost, it might be helpful to provide some kind of theoretical framework to guide both the rationale and the method used in the present study. The rationale for this content analytic study (as with all others) rests on the presumption that health information in mass media influences health perceptions, attitudes, and behaviors. It seems important to support this premise (and therefore this study) with a brief discussion of a well-tested theory that explains the relationship between media exposure and health outcomes.

b. The organizational structure of the introduction is somewhat disjointed, and therefore, difficult to follow. It might make more sense to order the introduction in a way that more clearly delineates the purpose of this study: to content analyze the presentation of health information in news media about a relatively misunderstood/unknown health topic over time.

c. The components of the introduction are all necessary to provide the readers with a compelling rationale for the present study. However, the jumps between a discussion about media coverage and celebrity influence to the discussion about public awareness of the link between HPV and oral cancer to the discussion of how mass media are important conduits of health information are difficult to follow.

d. Some of this information seems superfluous (e.g., discussion of the case of Jade Goody and Angelina Jollie’s mastectomy, the
discussion about public awareness of the signs and risk factors for head and neck cancers and the fact that the majority of oral cancers are diagnosed at an advanced stage) because it does not clearly relate to the primary purpose of the study (to explore how news coverage of the link between HPV and oral cancer has changed over time). I encourage the authors to make the connections between the literature reviewed in the introduction and the purpose of the study clearer and more explicit.

2) Method

a. First, it is unclear how the authors have differentiated “qualitative” and “quantitative” analysis in the method section. I understand the authors used the words “quantitative analysis” to refer to their calculations of the frequency of articles in the sample. The authors stated that they used the “Framework Analysis” method to analyze the content of the articles. It will be helpful for the authors to provide a more detailed explanation of this analytic approach, particularly the relationship between the four emergent themes and the application of those themes to the news content (e.g., were the identified themes used as an a priori coding scheme to classify the news content or was news content grouped into the four most relevant categories, based on the author’s and NVivo’s interpretation?).

b. Some of the confusion surrounding the use emergent themes and the use of an a priori coding scheme is likely a function of the description of the reliability analysis. I have some serious concerns about the reliability of the coding as it is presented in this manuscript. As the establishment of intercoder reliability is the cornerstone of the content analytic procedure, it seems paramount that the authors address the following issues:

i. First, I am not certain there is enough information about the criteria used for categorizing the news articles as broadsheet, middle market, or tabloid newspapers. It is unclear why the authors chose not to conduct reliability analysis on these categorization decisions, particularly because this categorization scheme is the basis upon which a number of quantitative comparisons are conducted. It seems important to provide more information about these categories (e.g., sample news sources, the use of some kind of objective external classification scheme, why these categorizations are important, etc.).

ii. Second, it is unclear which of the themes the authors coded (i.e., the four “main themes” discussed listed on page 7? the mysterious “sub-themes” mentioned on page 5?). Clarification is warranted. I recommend providing the readers with a list and description of all the themes that emerged and/or were applied to the content in the method section (and provide the reliability statistics for each theme/sub-theme).

iii. More importantly, the reliability coefficient provided for the coding of the main themes and the sub-themes of the articles (Cohen’s kappa = .68) indicates that there was considerable error in the application of the coding scheme – a serious, and potentially fatal, limitation in the analysis. Although there is no definitive criterion for a “cut-off” reliability coefficient, a number of researchers have suggested that reliability coefficients below .70 are problematic (i.e., indicative of intercoder “unreliability” that is likely due to conceptual and/or coder error beyond chance) (see Neuendorf’s 2002 The
Content Analysis Guidebook for a review of studies related to reliability testing in content analysis). Moreover, it is unclear whether the kappa coefficient provided is a summary coefficient that combines the reliability coefficients across all the themes/sub-themes. Clarification is warranted.

3) Results

a. The organization of the results section needs some revision. First, it might be helpful to provide a clear trend analysis section that details changes in the frequency of news articles about HPV and oral cancer over time. It might also be interesting to provide a trend analysis for the frequency of the four main themes over time. I suspect that a visual representation of the trends in frequencies will be easier to interpret than the current format on page 6 (which details changes in content over time, but not in chronological order).

b. The criteria for analyzing the headlines are unclear. Did the authors analyze the headlines using the four identified themes? If not, then how did they determine the focus of their analysis of these headlines? The authors reported that 33 headlines emphasized the risk of oral sex and HPV but provided sample headlines (e.g., “Your life: Oral sex cancer risk” and “Early alert for throat cancer”) that failed to mention HPV at all. Clarification is warranted here – it will be helpful to provide a more precise description of the analytic procedure used to examine the headlines.

c. When reporting the frequency with which certain categories/themes occurred, it will be helpful to provide both the percentages and the total (n/N) for each.

4) Discussion:

a. Some of the discussion focuses on the differences in coverage of the link between HPV and oral cancer across publication type. It might be helpful, here and/or in the introduction, to provide a stronger rationale for analyzing these differences between publication types.

b. I was looking for a clearer “implications” section in the discussion. The authors provide some implications throughout this section but it would be helpful to group these implications together in one new section/sub-section.

5) Abstract:

a. The abstract details could use more concision and precision, both in the “objectives” and the “conclusions” sections. For example, the first two sentences in the objectives section can be combined in a way that will enhance the flow of the abstract and emphasize the significance of the purpose of the present study (e.g., “The role of human papillomavirus (HPV) in some oral cancers has been reported in the news press, though little is known about the content of these articles.”) In the conclusions section, the repetitive word “this” appears in the last sentence makes the meaning of that sentence unclear (i.e., the first “this” refers to the frequent linking of HPV and oral sex and the second “this” refers to oral sex itself).
Reviewer: 1

Dannielle E. Kelley
University of North Carolina at Chapel Hill, USA

Please leave your comments for the authors below. First, I would like to compliment the authors on their research – understanding how the general public is exposed to important health issues through mass media is a key step towards the development of health communication and interventions to promote healthy behavior. I like that the authors recognized the variation in scope and reach across newspapers, and discussed the findings according to newspaper type. I think this paper would be a useful contribution to the literature, and I hope the authors follow up with the research mentioned in the discussion.

Thank you for these positive comments.

I do have a few concerns, most of which regard the organization and reporting of the data. I will list my concerns in the order they appear in the paper:

1. In the introduction, third paragraph, the authors cite an online survey administered in the UK, US, and Australia. However, the remainder of that paragraph focuses on data from the US alone. This creates a bit of a disconnect for the reader as the study focuses on HPV related content in UK newspapers. Data from the UK would be more relevant and useful here.

Thank you for your comments. Unfortunately, there have been no studies conducted yet in the UK about the awareness of HPV in oral cancer, therefore there are no data for the UK to be presented. We have updated the introduction to say this (page 3).

2. You have eight themes, but only report one Cohen's Kappa. Cohen's Kappa across the eight themes should be reported. Reporting the reliability statistic for each theme is crucial in content analysis as it enables replication and the possibility for standardized variables. Even if the Kappas were not sufficient for certain themes, that information should be provided in the text and the authors may offer an explanation for why some variables may not have been as reliable – without this, there is ambiguity in the findings and the field cannot advance. Please see the article below for reference:


A good example of how to present this information can be found in table 1 of this article:


-On a similar note, a description of the themes and individual codes within the themes should be provided.

Thank you for your suggestion. We have now included a table for the four main themes and seven sub-themes and calculated a reliability statistic for each of these (Table 1; page 9). There were some inconsistencies in the reliability statistics among the themes found, which could potentially be due to the small numbers coded in the 20% sample that was second-coded and some themes having multiple levels.
3. What was the justification for only including articles with 100 words about the link between oral cancer and HPV? This seems like a difficult distinction to make because there are a variety of ways to talk about oral cancer and HPV, making this a difficult criteria to define. Also, considering the varying article lengths, 100 words may mean something different across newspaper types.

We included all the articles in the quantitative analysis, so the frequency of articles can be seen across time. However, we felt that 100 words was a reasonable cut off to ensure the articles had sufficient content to be included in the content analysis. We also considered it important that the articles included in the content analysis were mainly focused on the link between HPV and oral cancer were not covering another story that included a sentence to say HPV was linked with oral cancer (which was the case in some). This is explained in the manuscript (page 7).

4. Regarding the discussion of peaks on page 6, the peak in 2013 is attributed to mouth cancer awareness month. To my knowledge, this is something that has been in place since at least 2009, so why is the peak attributed to this particular month in 2013? Further, are you certain the other peaks are attributable to the event mentioned? If so, I think this needs to be made clear for the reader.

We cannot explain why there were no peaks in Mouth Cancer Awareness Month in previous years. Perhaps it was greater in 2013 following Michael Douglas' disclosure. We have checked the peaks against the articles and have amended the numbers to reflect those attributable to the events mentioned. This has now been made clear in the manuscript on page 7.

5. How were the quotes in text and in the tables selected? Were these the most representative quotes of the eight themes? On page 10, paragraph 3, the authors discuss ONE newspaper that “suggested Douglas was criticized for his public honesty” followed by one quote about Zeta-Jones. What is the significance of this in terms of your results? If these isolated quotes were mentioned for each of the eight themes in a uniform fashion, it may be less of an issue, but here it almost seems this would fit better in the discussion.

The quotes in the text and the tables were selected as being most representative of the themes. The inclusion of the quotes with regards to Michael Douglas were included to demonstrate that some of the newspapers recognised that there was scepticism around what Michael Douglas had said. We also considered it important to include quotes from the newspapers that included information about Zeta-Jones to recognise that these disclosures can have an impact on others, especially partners. We have however, taken on board your comments and removed these - this is mentioned in the discussion (page 18).

6. Please be consistent with citing quotes in the text. For some quotes the authors provide the newspaper title and the date (which I think is the most useful format for the reader). Others, the authors provide a quotation number and reference a table. This format is frustrating and disrupts the flow of the paper. Also, not all quotes are cited; for example, page 10: “A few of the articles suggested that HPV transmitted through oral sex ‘could overtake tobacco and alcohol as the main risk factor in the next decade.’” But do not offer a citation – this is an important quote, would also be helpful to know the frequency in which this appeared in your sample.

We have formatted the quotes as the reviewer suggested. It was not possible to include all quotes in the text due to length restrictions, so some of these are supplied in the tables, but the newspaper and publication date have been added following the quotes. The number of citations of HPV overtake tobacco and alcohol as the main risk factor has also been included on page 13.

7. I understand why the authors separated the tables by theme, but this information could be
organized in a way that is much more useful. I would like to see the table as ONE comprehensive table organized by date – that way, we can see the progression of content through the years and in relation to news events. This will provide a much clearer picture of how HPV was discussed around these events. You can denote the theme by simply adding another column. Currently, Tables 1, 3, and 4, are limited by the organization.

Thank you for this suggestion. As the quotes do not necessarily reflect the first instance of the theme appearing in the newspapers, we feel that ordering them by date wouldn't be useful to the reader, as it would also affect the chronology of the quote numbers. We have now put all the quotes into one comprehensive table, ordered by where they appear in the text, representative by theme (Table 2; page 10/11).

8. Table 2 is a key table and ALL eight themes should be included. Why just include health themes? You have great data here, make it easier to see.

Thank you for your suggestion, we have now included all the themes in the table (Table 3; page 15).

9. My final comments may not be one the authors can address in this paper as I am not going to ask for reanalysis of the data, but it seems odd that there was no mention of HPV and cervical cancer in the articles. I understand the purpose was to look at the reported connection between HPV and head and neck cancers, however, in my perusal of such articles here in the US, the connection between HPV and cervical cancer or the prevalence among females, is often mentioned, although not the subject of the article. It seems that knowing how and if HPV in males and females was discussed in the same article would be important information for informing public health campaigns and message design to strengthen public knowledge about the implications of HPV for males and females. Finally, did the authors code for how often the same story line was presented across issues and newspapers? This would also be useful in understanding which types of stories gain traction in the media.

There is mention of HPV and cervical cancer in the articles and this has been stated in the manuscript on page 14 under the theme ‘Health Information’ and also in Table 3 (previously table 2). We did not code for how often the same story line was presented across issues and newspapers, but this is a valid point and we have reflected on it in the discussion on page 19.

Reviewer: 2

Jessie M. Quintero Johnson, PhD
University of Massachusetts Boston

Please leave your comments for the authors below The purpose of this study, to content analyze changes in the frequency of stories about the link between HPV and oral cancer in print news, is both interesting and meaningful. As the authors suggest, the presentation of health information in mass media is likely to influence public awareness about HPV. I believe the manuscript in its present form will need the following revisions in order to strengthen both the readability and the significance of the findings:

1) Introduction

a. First and foremost, it might be helpful to provide some kind of theoretical framework to guide both the rationale and the method used in the present study. The rationale for this content analytic study
(as with all others) rests on the presumption that health information in mass media influences health perceptions, attitudes, and behaviors. It seems important to support this premise (and therefore this study) with a brief discussion of a well-tested theory that explains the relationship between media exposure and health outcomes.

Thank you for your suggestion. We have now added a few lines about media priming and evidence from a UK media campaign called Be Clear on Cancer, which showed increases in awareness of key symptoms across a number of cancers and also an increase in GP attendance (among other outcomes). This is on page 5.

b. The organizational structure of the introduction is somewhat disjointed, and therefore, difficult to follow. It might make more sense to order the introduction in a way that more clearly delineates the purpose of this study: to content analyze the presentation of health information in news media about a relatively misunderstood/unknown health topic over time.

We have now amended the structure of the introduction to reflect these helpful comments.

c. The components of the introduction are all necessary to provide the readers with a compelling rationale for the present study. However, the jumps between a discussion about media coverage and celebrity influence to the discussion about public awareness of the link between HPV and oral cancer to the discussion of how mass media are important conduits of health information are difficult to follow.

We have now amended the structure of the introduction to reflect these helpful comments.

d. Some of this information seems superfluous (e.g., discussion of the case of Jade Goody and Angelina Jollie’s mastectomy, the discussion about public awareness of the signs and risk factors for head and neck cancers and the fact that the majority of oral cancers are diagnosed at an advanced stage) because it does not clearly relate to the primary purpose of the study (to explore how news coverage of the link between HPV and oral cancer has changed over time). I encourage the authors to make the connections between the literature reviewed in the introduction and the purpose of the study clearer and more explicit.

We included the information about celebrity diagnoses to reflect that of Michael Douglas’ disclosure in the context of HPV and oral cancer. We hope the amended structure of the introduction makes this clearer.

2) Method

a. First, it is unclear how the authors have differentiated “qualitative” and “quantitative” analysis in the method section. I understand the authors used the words “quantitative analysis” to refer to their calculations of the frequency of articles in the sample. The authors stated that they used the “Framework Analysis” method to analyze the content of the articles. It will be helpful for the authors to provide a more detailed explanation of this analytic approach, particularly the relationship between the four emergent themes and the application of those themes to the news content (e.g., were the identified themes used as an a priori coding scheme to classify the news content or was news content grouped into the four most relevant categories, based on the author’s and NVivo’s interpretation?).

We have now included a little more information under the subheadings of "quantitative" and "qualitative" analysis which we hope clarifies this. The analytic approach of framework analysis is already described in the text and we do not believe this needs further detail. We have however, added a bit of further detail about the generation of codes from the content and then the generation of main
themes and sub-themes from these codes (page 6).

b. Some of the confusion surrounding the use emergent themes and the use of an a priori coding scheme is likely a function of the description of the reliability analysis. I have some serious concerns about the reliability of the coding as it is presented in this manuscript. As the establishment of intercoder reliability is the cornerstone of the content analytic procedure, it seems paramount that the authors address the following issues:

i. First, I am not certain there is enough information about the criteria used for categorizing the news articles as broadsheet, middle market, or tabloid newspapers. It is unclear why the authors chose not to conduct reliability analysis on these categorization decisions, particularly because this categorization scheme is the basis upon which a number of quantitative comparisons are conducted. It seems important to provide more information about these categories (e.g., sample news sources, the use of some kind of objective external classification scheme, why these categorizations are important, etc.).

The news articles were checked against Newsbrands at http://www.newsworks.org.uk/Market-Overview and regional newspapers were categorised by discussion after looking at each newspaper and further information for each newspaper. This has now been made clear in the manuscript and figures updated accordingly. We have also included some sample news sources under each category (page 6) and explained why these categorisations are important in the discussion (page 17).

ii. Second, it is unclear which of the themes the authors coded (i.e., the four “main themes” discussed listed on page 7? the mysterious “sub-themes” mentioned on page 5?). Clarification is warranted. I recommend providing the readers with a list and description of all the themes that emerged and/or were applied to the content in the method section (and provide the reliability statistics for each theme/sub-theme).

The coding has been clarified now in the manuscript and we have also now included a table with the main themes and sub-themes and calculated a reliability statistic for each of these (Table 1; page 9).

iii. More importantly, the reliability coefficient provided for the coding of the main themes and the sub-themes of the articles (Cohen’s kappa = .68) indicates that there was considerable error in the application of the coding scheme – a serious, and potentially fatal, limitation in the analysis. Although there is no definitive criterion for a “cut-off” reliability coefficient, a number of researchers have suggested that reliability coefficients below .70 are problematic (i.e., indicative of intercoder “unreliability” that is likely due to conceptual and/or coder error beyond chance) (see Neuendorf’s 2002 The Content Analysis Guidebook for a review of studies related to reliability testing in content analysis). Moreover, it is unclear whether the kappa coefficient provided is a summary coefficient that combines the reliability coefficients across all the themes/sub-themes. Clarification is warranted.

As some of the content was coded under more than one theme or sub-theme, some ordering inconsistencies occurred. After accounting for these and having done more detailed analysis with the themes and sub-themes, Cohen's Kappa is actually 0.71 across all themes and sub-themes. The reliability coefficients for each main theme and sub-theme is now given in table 1, page 9 along with a list of all the main themes and sub-themes. There were some inconsistencies in the reliability statistics among the themes found, which could potentially be due to the small numbers coded in the 20% sample that was second-coded and some themes having multiple levels.

3) Results

a. The organization of the results section needs some revision. First, It might be helpful to provide a
clear trend analysis section that details changes in the frequency of news articles about HPV and oral cancer over time. It might also be interesting to provide a trend analysis for the frequency of the four main themes over time. I suspect that a visual representation of the trends in frequencies will be easier to interpret than the current format on page 6 (which details changes in content over time, but not in chronological order).

Figure 1 in the manuscript represents changes in content over time in a chronological order. We conducted some additional analysis to look at the changes in the frequency of the four main themes over time (which we have included in the reviewers comments document), but this did not really tell us anything other than the main themes peaked at the same time as the number of articles did, so we have chosen to not include this in the paper.

b. The criteria for analyzing the headlines are unclear. Did the authors analyze the headlines using the four identified themes? If not, then how did they determine the focus of their analysis of these headlines? The authors reported that 33 headlines emphasized the risk of oral sex and HPV but provided sample headlines (e.g., “Your life: Oral sex cancer risk” and “Early alert for throat cancer”) that failed to mention HPV at all. Clarification is warranted here – it will be helpful to provide a more precise description of the analytic procedure used to examine the headlines.

The headlines were analysed descriptively and this has now been clarified in the manuscript. The headlines offered were to illustrate the words ‘risk’, ‘alert’ and ‘dangers’, not HPV and oral sex. This should also have read ‘the risk of oral sex and cancer’ and so has been amended in the text and a headline added to illustrate that point (page 8).

c. When reporting the frequency with which certain categories/themes occurred, it will be helpful to provide both the percentages and the total (n/N) for each.

Percentages have now been included.

4) Discussion:

a. Some of the discussion focuses on the differences in coverage of the link between HPV and oral cancer across publication type. It might be helpful, here and/or in the introduction, to provide a stronger rationale for analyzing these differences between publication types.

We agree with this and have discussed why differences between publication types are important in the discussion (page 17).

b. I was looking for a clearer “implications” section in the discussion. The authors provide some implications throughout this section but it would be helpful to group these implications together in one new section/sub-section.

Thank you for your suggestion. We have grouped more of these implications into the conclusion section at the end of the discussion on page 19/20.

5) Abstract:

a. The abstract details could use more concision and precision, both in the “objectives” and the “conclusions” sections. For example, the first two sentences in the objectives section can be combined in a way that will enhance the flow of the abstract and emphasize the significance of the purpose of the present study (e.g., “The role of human papillomavirus (HPV) in some oral cancers has been reported in the news press, though little is known about the content of these articles.”) In the
conclusions section, the repetitive word “this” appears in the last sentence makes the meaning of that sentence unclear (i.e., the first “this” refers to the frequent linking of HPV and oral sex and the second “this” refers to oral sex itself).

We have taken your changes on board and adapted the abstract to reflect this (page 2).

**VERSION 2 – REVIEW**

<table>
<thead>
<tr>
<th>REVIEWER</th>
<th>Dannielle Kelley</th>
</tr>
</thead>
<tbody>
<tr>
<td>UNIVERSITY OF NORTH</td>
<td>University of North Carolina at Chapel Hill</td>
</tr>
<tr>
<td>CAROLINA AT CHAPEL</td>
<td></td>
</tr>
<tr>
<td>HILL</td>
<td></td>
</tr>
<tr>
<td>REVIEW RETURNED</td>
<td>01-Dec-2015</td>
</tr>
</tbody>
</table>

| GENERAL COMMENTS | I appreciate the authors’ diligence in addressing most of my concerns. However, there are still a few issues that should be addressed before publication:

1. Paragraph 2 in the introduction (the US, UK, AUS survey) is still confusing. First, the 61% awareness is confusing - if the focus of this paper is media coverage in the UK, why not report the statistics for men (39.2%) and women (61.6%) in the UK? That information seems more compelling than the overall statistic across US/UK/AUS.

In the same paragraph, the discussion of physician awareness seems irrelevant and is still confusing. How does this relate to your analysis of mass media coverage? Finally, the last sentence in paragraph 2 says "providers have expressed concern about appropriate communication regarding HPV-related oral cancer." Is "communication" in this sentence referring to patient-provider communication or mass media communication? If the later, please specify, if the former, I suggest removing it because it seems irrelevant. Also, it seems some quotes were removed from the table since the first submission.

2. Thank you for consolidating the quote tables. However, there is still a lack of organization. As I mentioned in my previous review, the longitudinal nature of the data is a major strength of this study. The quotes should be organized by data, regardless of theme/subtheme. Mass media reflects the cultural narrative of the time in which it was published. If the table were organized by date, you would be able to see that narrative more clearly and it would be of more use to the reader as the evolution of the HPV conversation will be readily apparent. If the authors do not wish to do this, please at least organize by date within themes as the table is currently unorganized.

Minor edits:
Results: please be consistent when reporting frequencies. For example, page 7 "About one third (35/112; 31%)..." and within the same sentence: "...Douglas mentioned in a quarter (n=28/112; 28%)..." Related to this, "one third" followed by the percentage is redundant, please choose one or the other.

Why are only some of the quotes mentioned in text included in the table? If there is a reason for not including all quotes in the text in table 2, please explain as a footnote to table 2.
Reviewer: 1

Dannielle Kelley
University of North Carolina at Chapel Hill

Please leave your comments for the authors below I appreciate the authors' diligence in addressing most of my concerns. However, there are still a few issues that should be addressed before publication:

1. Paragraph 2 in the introduction (the US, UK, AUS survey) is still confusing. First, the 61% awareness is confusing - if the focus of this paper is media coverage in the UK, why not report the statistics for men (39.2%) and women (61.6%) in the UK? That information seems more compelling than the overall statistic across US/UK/AUS.

   Thank you for your suggestion. This has now been amended to reflect the figures for the UK.

   In the same paragraph, the discussion of physician awareness seems irrelevant and is still confusing. How does this relate to your analysis of mass media coverage? Finally, the last sentence in paragraph 2 says "providers have expressed concern about appropriate communication regarding HPV-related oral cancer." Is "communication" in this sentence referring to patient-provider communication or mass media communication? If the later, please specify, if the former, I suggest removing it because it seems irrelevant. Also, it seems some quotes were removed from the table since the first submission.

   The discussion of physician awareness is included in this paragraph as it demonstrates levels of knowledge of HPV as a risk factor for oral cancer in different populations. This has now been separated from the sentence that describes that there have been no studies from the UK that have investigated knowledge of HPV as a risk factor for oral cancer. We hope this is now less confusing.

   The final sentence did relate to patient-provider communication and therefore has been removed.

   In your previous review you suggested removing quotes about Catherine Zeta-Jones as these appeared in the text in isolation. Therefore, the two quotes removed from the table were those referred to in the text about Catherine Zeta-Jones. These were removed as now this has just been mentioned in the discussion.

2. Thank you for consolidating the quote tables. However, there is still a lack of organization. As I mentioned in my previous review, the longitudinal nature of the data is a major strength of this study. The quotes should be organized by data, regardless of theme/subtheme. Mass media reflects the cultural narrative of the time in which it was published. If the table were organized by date, you would be able to see that narrative more clearly and it would be of more use to the reader as the evolution of the HPV conversation will be readily apparent. If the authors do not wish to do this, please at least organize by date within themes as the table is currently unorganized.

   As the quotes do not necessarily reflect the first instance of the theme appearing in the newspapers, we feel that ordering them by date wouldn't be useful to the reader. We have however, now ordered each theme by date (Table 2; page 9/10).

Minor edits:
Results: please be consistent when reporting frequencies. For example, page 7 "About one third
"...and within the same sentence: "...Douglas mentioned in a quarter (n=28/112; 28%)..." Related to this, "one third" followed by the percentage is redundant, please choose one or the other.

Thank you for your comment, the percentage has now been removed (page 7) from these two examples and all other frequencies are reported with percentages.

Why are only some of the quotes mentioned in text included in the table? If there is a reason for not including all quotes in the text in table 2, please explain as a footnote to table 2.

The table only includes those quotes which are longer and could not be incorporated into a sentence. We did not feel that shorter quotes that are given in the text needed to also be given in the table. A footnote has now been added to table 2 to reflect this reasoning (page 10).

**VERSION 3 - REVIEW**

<table>
<thead>
<tr>
<th>REVIEWER</th>
<th>Dannielle Kelley</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>UNC Chapel Hill</td>
</tr>
<tr>
<td>REVIEW RETURNED</td>
<td>08-Feb-2016</td>
</tr>
</tbody>
</table>

| GENERAL COMMENTS | No further comment                        |
Print and online newspaper coverage of the link between HPV and oral cancer in the UK: a mixed-methods study

Rachael H Dodd, Laura A V Marlow, Alice S Forster and Jo Waller


Updated information and services can be found at: http://bmjopen.bmj.com/content/6/2/e008740

These include:

References
This article cites 36 articles, 1 of which you can access for free at: http://bmjopen.bmj.com/content/6/2/e008740#BIBL

Open Access
This is an Open Access article distributed in accordance with the terms of the Creative Commons Attribution (CC BY 4.0) license, which permits others to distribute, remix, adapt and build upon this work, for commercial use, provided the original work is properly cited. See: http://creativecommons.org/licenses/by/4.0/

Email alerting service
Receive free email alerts when new articles cite this article. Sign up in the box at the top right corner of the online article.

Topic Collections
Articles on similar topics can be found in the following collections

Communication (215)
Health services research (1524)
Qualitative research (719)

Notes

To request permissions go to: http://group.bmj.com/group/rights-licensing/permissions

To order reprints go to: http://journals.bmj.com/cgi/reprintform

To subscribe to BMJ go to: http://group.bmj.com/subscribe/