

PEER REVIEW HISTORY

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

ARTICLE DETAILS

TITLE (PROVISIONAL)	A survey on female genital mutilation/cutting in Jeddah, Saudi Arabia
AUTHORS	Rouzi, Abdulrahim; Berg, Rigmor C; Alamoudi, Rana; Alzaban, Faten; Sehlo, Mohammad

VERSION 1 - REVIEW

REVIEWER	Ana Lúcia Teixeira Interdisciplinary Centre of Social Sciences (CICS.NOVA), Faculty of Social Sciences and Humanities (NOVA FCSH), Portugal
REVIEW RETURNED	30-Jun-2018

GENERAL COMMENTS	<p>The article aims to fill a knowledge gap regarding the existence of FGM in Saudi Arabia, through a cross-sectional study in a specific city, Jeddah.</p> <p>Globally, the paper reveals adequate knowledge on the topic under analysis. A reasonable literature review was made, identifying the main gaps that support the research. However, in order to be considered ready for publication in a journal of this quality, I would suggest some improvements, namely:</p> <p>Methods Could the authors elaborate on the age criterion for eligibility – 18 to 75 years old? My first comment is that, usually, FGM prevalence studies use the age group from 15 to 49 years old. Knowing this is not a prevalence study, the same criterion could be used in order to have some comparability. My second comment refers to the importance of going beyond reproductive age (as done in several studies in Europe). That being said, the authors option for surveying women up to 75 years old is valuable, but why stop at 75 years old? The third comment goes for the lower threshold. For survey purposes, 18 years old is commonly the lower threshold since it is, for most countries, the age of majority. Is this the case in Saudi Arabia? If not, why not lower the threshold, again for comparability issues?</p> <p>Results The authors should provide information on the value of the test statistics used (as well as the degrees of freedom) and not only the probability value. This should be described fully.</p>
-------------------------	---------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------

The sentence “Most of the women in this sample had heard of FGM/C (89·6%), but 2·3% were unsure” (p. 4) should be moved to the end of the first paragraph of the results’ section. This is the case since it is a descriptive sentence and relates to all women surveyed. Its current placement compromises the flow of the argument.

Discussion

Regarding the study limitations, the authors state that the sample “is likely to be representative of the population in Jeddah” (p. 5). How do the authors come to this conclusion? The sample is likely to be representative of which characteristics? Age, nationality, employment status, earnings, etc.? If this aspect is not clear, one cannot state that the study design is appropriate to answer the research question – more information should be provided regarding the representativeness of the sample towards the target population.

The authors address summarily the concern with the recall bias as the only drawback from self-reporting studies. Studies such as Klouman, Manongi, & Klepp (2005), Snow, Slinger, Okonofua, Oronsaye, & Wacker (2002), and Morison, Scherf, Ekpo, Paine, West, Coleman, & Walraven, (2001), show inconsistencies between self-reported and clinically determined FGM status to different extents. However, a more recent study shows that while there was complete agreement between reporting having undergone FGM (or not) and what was found by clinical inspection, the reliability of reported form of FGM was low (Elmusharaf, Elhadi, & Almroth, 2006). As such, the shortcomings of relying in self-report surveys should be addressed more thoroughly since they may impact on the interpretation of variables such as the type of FGM.

Another limitation the authors recognize is the lack of information about the origin of the Saudi women. This is very important since FGM is strongly rooted in tradition and culture. Hence, place of birth should be used (and not only nationality) to better understand why it is practiced in Saudi Arabia. The authors refer the closeness to Yemen as a possible explanation for high percentages of FGM among Saudi women but another hypothesis could be considered, such as these women being second generation migrants, born to mothers from FGM practicing countries.

References used in this review

Klouman, E. , Manongi, R. and Klepp, K. (2005), Self-reported and observed female genital cutting in rural Tanzania: associated demographic factors, HIV and sexually transmitted infections. *Tropical Medicine & International Health*, 10: 105-115. doi:10.1111/j.1365-3156.2004.01350.x

Snow, R. C., Slinger, T. E., Okonofua, F. E., Oronsaye, F. and Wacker, J. (2002), Female genital cutting in southern urban and peri-urban Nigeria: self-reported validity, social determinants and secular decline. *Tropical Medicine & International Health*, 7: 91-100. doi:10.1046/j.1365-3156.2002.00829.x

Morison, L. , Scherf, C. , Ekpo, G. , Paine, K. , West, B. , Coleman, R. and Walraven, G. (2001), The long-term reproductive health consequences of female genital cutting in rural Gambia: a community-based survey. *Tropical Medicine & International Health*, 6: 643-653. doi:10.1046/j.1365-3156.2001.00749.x

	Elmusharaf, S., Elhadi, N., and Almroth, L. (2006), Reliability of self reported form of female genital mutilation and WHO classification: cross sectional study. <i>BMJ</i> , 333: 1-5. doi: https://doi.org/10.1136/bmj.38873.649074.55
--	-----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------

REVIEWER	Livia Elisa Ortensi University of Milan - Bicocca, Italy
REVIEW RETURNED	03-Jul-2018

GENERAL COMMENTS	<p>I think this study has a broad potential but is at the same time at a very preliminary phase. So I suggest that author advance their research and submit a revised version of the paper.</p> <p>The most significant contribution to the literature is showing that Saudi Arabia is at the same time a country with immigrant women with FGM and, most importantly, native women with FGM. Saudia Arabia is therefore also a practising country.</p> <p>I suggest authors keep the essential distinction between Saudi women and migrants along all the study, trying to underline the differences between Saudi cut and uncut women and migrants. Native and migrants population may differ a lot in their characteristics.</p> <p>In their current form, the study mix outcomes of migrants and natives cut women. In this way, authors give indeed some results about the characteristics of cut women in Jeddah, but we have a mix of immigrants and natives. As immigrants are from countries covered by DHS surveys, and mostly from Yemen, the sample of cut women has some similarities with outcomes from Yemen DHS. In table 1 I think is most useful to show row % instead of column %. In this way, you would indicate that the prevalence of FGM among Saudi women is 13% among naturalised Saudi is 30%, and among foreign women is 33%.</p> <p>A revised Table 2 could be similar to the current one, but keeping cut native Saudi, uncut native Saudi, other cut women and other uncut women as categories.</p> <p>It's also not clear how Saudi women differ from foreign or naturalised women regarding socioeconomic characteristics and how this impact on their FGM status.</p> <p>In trying to understand if and how the sample can be biased it would be useful to compare women socioeconomic characteristics with those of women residing in the city of Jeddah (e.g. using census data?)</p> <p>It would also be interesting to know what type is prevalent among Saudi women and if they have a different set of related complications</p> <p>Minor observations:</p> <p>Page 3 lines 17-19: I suggest the authors underline that cut women from Oman, Saudi Arabia, and the United Arab Emirates are not included in the UNICEF estimation that is cited.</p> <p>Page 4 lines 27-28 the Kruskal-Wallis is a test on the median not on the mean.</p> <p>A curiosity about Eligibility criteria: the selection of women who can read and speak Arabic may impact on the observed prevalence among migrants and/or natives?</p>
-------------------------	-------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------

VERSION 1 – AUTHOR RESPONSE

Reviewer #1, Comment #1 “Could the authors elaborate on the age criterion for eligibility – 18 to 75 years old? My first comment is that, usually, FGM prevalence studies use the age group from 15 to 49 years old. Knowing this is not a prevalence study, the same criterion could be used in order to have some comparability.”

Response: We chose to be as inclusive as possible, and chose 18 years as the lower end because it is the age used conventionally in Saudi Arabia as the age of majority. We also chose the upper age limit of 75 years as it is the life expectancy in Saudi Arabia (<https://www.cia.gov/library/publications/the-world-factbook/fields/2102.html> and <https://www.worldlifeexpectancy.com/saudi-arabia-life-expectancy>).

Reviewer #1, Comment #2 "My second comment refers to the importance of going beyond reproductive age (as done in several studies in Europe). That being said, the author's option for surveying women up to 75 years old is valuable, but why stop at 75 years old?"

Response: We chose the age limit of 75 years as it is the life expectancy in Saudi Arabia (see above) and very few of our patients are above the age of 75.

Reviewer #1, Comment #3 “The third comment goes for the lower threshold. For survey purposes, 18 years old is commonly the lower threshold since it is, for most countries, the age of majority. Is this the case in Saudi Arabia? If not, why not lower the threshold, again for comparability issues?”

Response: The reviewer is correct. As stated above we chose 18 years as the lower threshold because it is the conventional age of majority in Saudi Arabia (the age of majority in medical and government institutions is 18 years, but the religious authorities disagree).

Reviewer #1, Comment #4 “The authors should provide information on the value of the test statistics used (as well as the degrees of freedom) and not only the probability value. This should be described fully.”

Response: We have added the value of the test for significant difference between women with FGM/C and women with no FGM/C to table 1 (last column).

Reviewer #1, Comment #5 “The sentence “Most of the women in this sample had heard of FGM/C (89.6%), but 2.3% were unsure” (p. 4) should be moved to the end of the first paragraph of the results’ section. This is the case since it is a descriptive sentence and relates to all women surveyed. Its current placement compromises the flow of the argument.”

Response: Done.

Reviewer #1, Comment #6 “Regarding the study limitations, the authors state that the sample “is likely to be representative of the population in Jeddah” (p. 5). How do the authors come to this conclusion? The sample is likely to be representative of which characteristics? Age, nationality, employment status, earnings, etc.? If this aspect is not clear, one cannot state that the study design is appropriate to answer the research question – more information should be provided regarding the representativeness of the sample towards the target population.”

Response: We revised this statement to “the sample is likely to be representative of the population in Jeddah only since the population of Jeddah may be different than other cities in Saudi Arabia.” We mentioned on the same page “it is the principal gateway to Islam’s two holiest shrines in Mecca and Medina. Muslims are obliged to visit Mecca to perform religious duties at least once during their lifetime, if financially feasible. Some may elect to immigrate and live in the Hejaz region.”

Reviewer #1, Comment #6 “The authors address summarily the concern with the recall bias as the only drawback from self-reporting studies. Studies such as Klouman, Manongi, & Klepp (2005), Snow, Slinger, Okonofua, Oronsaye, & Wacker (2002), and Morison, Scherf, Ekpo, Paine, West, Coleman, & Walraven, (2001), show inconsistencies between self-reported and clinically determined FGM status to different extents. However, a more recent study shows that while there was complete agreement between reporting having undergone FGM (or not) and what was found by clinical inspection, the reliability of reported form of FGM was low (Elmusharaf, Elhadi, & Almroth, 2006). As such, the shortcomings of relying in self-report surveys should be addressed more thoroughly since they may impact on the interpretation of variables such as the type of FGM.”

Response: Thank you, we revised to “Studies have shown inconsistencies between self-reported and clinically determined FGM status to different extents.^{22,23} However, a more recent study shows that while there was complete agreement between reporting having undergone FGM (or not) and what was found by clinical inspection, the reliability of reported form of FGM was low.²⁴”

Reviewer #1, Comment #7 "Another limitation the authors recognize is the lack of information about the origin of the Saudi women. This is very important since FGM is strongly rooted in tradition and culture. Hence, place of birth should be used (and not only nationality) to better understand why it is practiced in Saudi Arabia. The authors refer the closeness to Yemen as a possible explanation for high percentages of FGM among Saudi women but another hypothesis could be considered, such as these women being second-generation migrants, born to mothers from FGM practicing countries."

Response: We agree. Therefore, we added, "Another possibility could be considered, such as these women being second-generation migrants, born to mothers from FGM practicing countries." Regrettably, we did not ask about participants' place of birth and therefore cannot report on this.

Reviewer #2, Comment #1 “I suggest authors keep the essential distinction between Saudi women and migrants along all the study, trying to underline the differences between Saudi cut and uncut women and migrants. Native and migrants population may differ a lot in their characteristics.

In their current form, the study mix outcomes of migrants and natives cut women. In this way, authors give indeed some results about the characteristics of cut women in Jeddah, but we have a mix of immigrants and natives. As immigrants are from countries covered by DHS surveys, and mostly from Yemen, the sample of cut women has some similarities with outcomes from Yemen DHS.

In table 1 I think is most useful to show row % instead of column %. In this way, you would indicate that the prevalence of FGM among Saudi women is 13% among naturalised Saudi is 30%, and among foreign women is 33%.

A revised Table 2 could be similar to the current one, but keeping cut native Saudi, uncut native Saudi, other cut women and other uncut women as categories.

It's also not clear how Saudi women differ from foreign or naturalised women regarding socioeconomic characteristics and how this impact on their FGM status.

In trying to understand if and how the sample can be biased it would be useful to compare women socioeconomic characteristics with those of women residing in the city of Jeddah (e.g., using census data?)

It would also be interesting to know what type is prevalent among Saudi women and if they have a different set of related complications.”

Response: We appreciate all these observations. We added them as one comment because we think it is about the same thing (compare cut and uncut Saudis, naturalized-Saudis, and non-Saudi with regards to the type of FGM/C, socioeconomic status,...etc.). While we appreciate the comments, we think this is not the scope of the manuscript. It needs another manuscript to document the differences as requested.

Reviewer #2, Comment #2 “Page 3 lines 17-19: I suggest the authors underline that cut women from Oman, Saudi Arabia, and the United Arab Emirates are not included in the UNICEF estimation that is cited.

Response: We wrote "UNICEF'S report includes Oman, Saudi Arabia, and the United Arab Emirates as countries where FGM/C exists, but, "the evidence comes from (sometimes outdated) small-scale studies or anecdotal accounts."

Reviewer #2, Comment #3 “page 4 lines 27-28 the Kruskal-Wallis is a test on the median not on the mean.”

Response: We think there may be a mistake here. The Kruskal-Wallis is a rank-based nonparametric test that can be used to check if there are statistically significant differences between two or more groups of an independent variable on a continuous or ordinal dependent variable. Thus, it can be used to test the difference between groups on mean age.

Reviewer #2, Comment #4 “A curiosity about Eligibility criteria: the selection of women who can read and speak Arabic may impact on the observed prevalence among migrants and/or natives?”

Response: It could impact both. However, there are no good data on literacy among women in Saudi Arabia, so we can only speculate.

VERSION 2 – REVIEW

REVIEWER	Ana Lúcia Teixeira Interdisciplinary Centre of Social Sciences (CICS.NOVA), Faculty of Social Sciences and Humanities (NOVA FCSH), Portugal
REVIEW RETURNED	14-Oct-2018

GENERAL COMMENTS	<p>The authors have answered to all my questions and comments but there are still some issues that should be addressed.</p> <p>Regarding the comments identified as #1 and #2, the authors have adequately justified their methodological option. Nevertheless, that reasoning should be presented on the article: why did they chose those specific upper and lower thresholds.</p> <p>In what concerns comment #4, the revised version of the article now includes a new version of table 1, providing information on the value of the test statistics used (as well as the degrees of freedom) on the last column. However, some issues remain:</p> <p>(1) the column heading refers to “test difference” – this label is correct when you are using a t-test, but when you are performing a chi-square test you are testing independence, not difference. Hence, the column heading should be changed to something as “test statistics”;</p> <p>(2) which statistic was used to test the mean age difference between women with and without FGM? I would go for a t-test (or a Mann–</p>
-------------------------	----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------

Whitney test if needed). But the reported test statistic is an “r”. What does did refer to?;
 (3) the note at the end of the table states the following: “*= statistically significant differences between women with FGM/C and women with no FGM/C were found for age, Saudi nationality vs not, married vs not, income <5000 Saudi Riyal vs >5000. ns= not statistically significant.”
 The significance should be reported for each test, not for the entire column; additionally, no test statistic is marked as “ns”, so there is no point in including that information on the note below the table;
 (4) the chi-square statistics refer to each variable as a whole; hence the test statistic and p-value should appear on the variable row (not on the category row) or, at least, the cells should be merged in order to include all variable’s categories [see attached file for example]

Regarding comment #6, the authors have not fully address my remark. The authors state that “It consists of Saudi and non-Saudi women and is likely to be representative of the population in Jeddah, Saudi Arabia only since the population of Jeddah may be different than other cities in Saudi Arabia”. A representative sample is a group that closely matches the characteristics of its population as a whole. In order to say that our sample is representative (or likely to be representative), we have to define to which characteristics we are referring to. There are many ways to evaluate representativeness (eg gender, age, socioeconomic status, profession, education). This sample is not representative of the population in Jeddah in terms of gender, for instance. Hence, the authors should state to which features the sample is representative of or not use the term “representative” at all.

Still on comment #6, in which refers to the study limitations, the authors have simply copied my comments and pasted them to their article – cf. page 5 of the marked copy manuscript and my submitted review (bmjopen-2018-024684). This must be altered.

Example for comment #4 (4)

This is what we have now:

Nationality				
Saudi	87 (49.7)	572 (75.7)	683 (70.9)	X ² =81.23, df=1 p<0.001
Naturalized Saudi	23 (13.1)	54 (7.1)	79 (8.2)	
Non-Saudi	65 (37.2)	130 (17.5)	201 (20.9)	

This is what we should have:

Nationality				X ² =81.23, df=1 p<0.001
Saudi	87 (49.7)	572 (75.7)	683 (70.9)	
Naturalized Saudi	23 (13.1)	54 (7.1)	79 (8.2)	
Non-Saudi	65 (37.2)	130 (17.5)	201 (20.9)	

Or:

Nationality				
-------------	--	--	--	--

	Saudi	87 (49.7)	572 (75.7)	683 (70.9)	X ² =81.23, df=1 p<0.001
	Naturalized Saudi	23 (13.1)	54 (7.1)	79 (8.2)	
	Non-Saudi	65 (37.2)	130 (17.5)	201 (20.9)	

REVIEWER	Livia Elisa Ortensi University of Milan Bicocca, Italy
REVIEW RETURNED	23-Oct-2018

GENERAL COMMENTS	<p>Dear authors, I appreciate the improvements over your paper. I still have doubts about the use of the Kruskal Wallis test to assess significant differences regarding mean age at cutting between national groups. I suppose that you followed the SPSS standard here https://statistics.laerd.com/spss-tutorials/kruskal-wallis-h-test-using-spss-statistics.php including suggestions on how to report data.</p> <p>The way you report results is ok, "Saudi women had the lowest mean rank of 12.79, and Egyptian women had the highest with 34.5" however, as is said in that same page, "Remember, the distribution of your data will determine whether you can report differences with respect to medians".</p> <p>Here you can find some more hints https://stats.stackexchange.com/questions/33759/do-we-need-to-report-the-median-or-the-mean-when-using-a-kruskal-wallis-test http://www.biostathandbook.com/kruskalwallis.html</p> <p>Quoting from the latter page "The null hypothesis of the Kruskal–Wallis test is not that the means are the same. It is therefore incorrect to say something like "The mean concentration of fructose is higher in pears than in apples (Kruskal–Wallis test, P=0.02)," although you will see data summarized with means and then compared with Kruskal–Wallis tests in many publications."</p> <p>I suggest authors change this sentence "a Kruskal Wallis test showed significant differences in the mean age FGM/C was performed in each nationality group (p = .002) " accordingly. Of course, you have significant differences in age at cutting, but with the Kruskal Wallis test, you are not testing differences between means.</p> <p>One way to say it could be: For women who had had FGM/C later than one week after birth, a Kruskal Wallis test showed significant differences related to age at cutting between nationality groups (p = .002). Saudi women had the lowest mean rank of 12.79, and Egyptian women had the highest with 34.5.</p> <p>Then median age could be reported instead of the means (this is a suggestion according to also to the references I linked).</p>
-------------------------	---------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------

VERSION 2 – AUTHOR RESPONSE

Reviewer #1, Comment #1 "Regarding the comments identified as #1 and #2, the authors have adequately justified their methodological option. Nevertheless, that reasoning should be presented on the article: why did they chose those specific upper and lower thresholds."

Response: We added “We chose the lower age limit as 18 years because it is the age used conventionally in Saudi Arabia as the age of majority and the upper age limit as 75 years as very few of our patients are above the age of 75 years” in the results section.

Reviewer #1, Comment #2 “In what concerns comment #4, the revised version of the article now includes a new version of table 1, providing information on the value of the test statistics used (as well as the degrees of freedom) on the last column. However, some issues remain:

(1) the column heading refers to “test difference” – this label is correct when you are using a t-test, but when you are performing a chi-square test you are testing independence, not difference. Hence, the column heading should be changed to something as “test statistics.”

Response: Thank you. We changed the column heading to ‘test statistics’

Reviewer #1, Comment #3 “(2) which statistic was used to test the mean age difference between women with and without FGM? I would go for a t-test (or a Mann–Whitney test if needed). But the reported test statistic is an “r”. What does it refer to?”

Response: To check whether the difference between the two groups’ average age was statistically different, we calculated Pearson correlation coefficient, denoted r.

Reviewer #1, Comment #4 “(3) the note at the end of the table states the following: “*= statistically significant differences between women with FGM/C and women with no FGM/C were found for age, Saudi nationality vs not, married vs not, income <5000 Saudi Riyal vs >5000. ns= not statistically significant.” The significance should be reported for each test, not for the entire column; additionally, no test statistic is marked as “ns”, so there is no point in including that information on the note below the table”

Response: Thank you for this observation. We have changed as suggested.

Reviewer #1, Comment #5 “(4) the chi-square statistics refer to each variable as a whole; hence the test statistic and p-value should appear on the variable row (not on the category row) or, at least, the cells should be merged in order to include all variable’s categories [see attached file for example]”

Response: Thank you. We agree and this will be taken care of before publication.

Reviewer #1, Comment #6 “Regarding comment #6, the authors have not fully address my remark. The authors state that “It consists of Saudi and non-Saudi women and is likely to be representative of the population in Jeddah, Saudi Arabia only since the population of Jeddah may be different than other cities in Saudi Arabia”. A representative sample is a group that closely matches the characteristics of its population as a whole. In order to say that our sample is representative (or likely to be representative), we have to define to which characteristics we are referring to. There are many ways to evaluate representativeness (eg gender, age, socioeconomic status, profession, education). This sample is not representative of the population in Jeddah in terms of gender, for instance. Hence, the authors should state to which features the sample is representative of or not use the term “representative” at all.”

Response: In the revised version we explained that “the sample is likely to be representative of the population in Jeddah only since the population of Jeddah may be different than other cities in Saudi Arabia.” We mentioned on the same page “it is the principal gateway to Islam’s two holiest shrines in Mecca and Medina. Muslims are obliged to visit Mecca to perform religious duties at least once during their lifetime, if financially feasible. Some may elect to immigrate and live in the Hejaz region.” We acknowledged that our sample is hospital-based “The hospital-based, convenience sample is non-random.” We appreciate the comments of the reviewer and therefore we changed it to “the sample may be representative of the population in Jeddah only.”

Reviewer #1, Comment #7 “Still on comment #6, in which refers to the study limitations, the authors have simply copied my comments and pasted them to their article – cf. page 5 of the marked copy manuscript and my submitted review (bmjopen-2018-024684). This must be altered.”

Response: We changed it to “The data are based on self-report and may be susceptible to recall bias and low reliability. Studies have shown inconsistencies between self-reported and clinically determined FGM status to different extents.22-24”

Reviewer #2, Comment #1 “Dear authors, I appreciate the improvements over your paper.”

Response: Thank you.

Reviewer #2, Comment #2 “I still have doubts about the use of the Kruskal Wallis test to assess significant differences regarding mean age at cutting between national groups. I suppose that you followed the SPSS standard here <https://statistics.laerd.com/spss-tutorials/kruskal-wallis-h-test-using-spss-statistics.php> including suggestions on how to report data.....”

Response:

We appreciate the opportunity to respond to the thoughtful and salient questions raised by the reviewers.

VERSION 3 - REVIEW

REVIEWER	Ana Lúcia Teixeira Interdisciplinary Centre of Social Sciences (CICS.NOVA), School of Social Sciences and Humanities (NOVA FCSH), Portugal
REVIEW RETURNED	18-Feb-2019

GENERAL COMMENTS	All comments were answered and the manuscript has greatly improved. I wish the authors much success in all their future work and publications.
-------------------------	------------------------------------------------------------------------------------------------------------------------------------------------

REVIEWER	Livia Elisa Ortensi University of Bologna, Italy
REVIEW RETURNED	05-Mar-2019

GENERAL COMMENTS	I have only minor observations on this revised version that authors can easily implement A small comment on the result section: "58.6% had some university education or had completed a university degree, and 42.0% were current students". I think you are mixing the comment on education with the following point (occupation) but dealing with this in the same sentence is misleading, the reader may understand that the rest (i.e. not holding some university degree) are all students. As for the two points were age is analysed 1) use of the Kruskal Wallis test,
-------------------------	------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------

	<p>2) use of r instead of a t-test or a Mann–Whitney test to analyse age between the two subpopulations (table 1 line 1) https://www.quora.com/Whats-the-difference-between-t-test-and-correlation (I'm puzzled by this choice) I wholeheartedly agree with reviewer 1's advice "I would go for a t-test (or a Mann–Whitney test if needed)".</p>
--	-----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------

VERSION 3 – AUTHOR RESPONSE

"A small comment on the result section:

"58.6% had some university education or had completed a university degree, and 42.0% were current students". I think you are mixing the comment on education with the following point (occupation) but dealing with this in the same sentence is misleading, the reader may understand that the rest (i.e. not holding some university degree) are all students"

Response: We changed to "Slightly less than half (42.0%) of the women were current students, while about a third (28.5%) were employed, full- or part-time."

"As for the two points were age is analysed

1) use of the Kruskal Wallis test,

2) use of r instead of a t-test or a Mann–Whitney test to analyse age between the two subpopulations (table 1 line 1)

<https://www.quora.com/Whats-the-difference-between-t-test-and-correlation> (I'm puzzled by this choice)

I wholeheartedly agree with reviewer 1's advice "I would go for a t-test (or a Mann–Whitney test if needed)".

Response: We changed to "We conducted X2 analyses, and t-test to compare the demographic characteristics of women with and without FGM/C, using the Statistical Package for the Social Sciences (SPSS Inc., Chicago, IL), version 24.0." Accordingly, this was added in Table1.