

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)	Does the ethnic density effect extend to obesity? A cross-sectional study of 415,166 adults in East London
AUTHORS	Sutaria, Shailen; Mathur, Rohini; Hull, Sally

VERSION 1 – REVIEW

REVIEWER	Marilyn Tseng Cal Poly San Luis Obispo, USA
REVIEW RETURNED	02-Aug-2018

GENERAL COMMENTS	<p>The manuscript describes the association between ethnic density and obesity among different ethnic groups in three areas of East London. The topic is worthwhile; the association of ethnic density with mental health outcomes is better established, but its association with physical health outcomes is unclear, and many of the physical health outcomes of interest are obesity-related. A strength of the study is its large sample of >400,000 adults. Issues to be addressed include:</p> <ol style="list-style-type: none">1. The introduction adequately presents the research question but could be tightened and sharpened substantially. For example, information in paragraph 3 is somewhat redundant; the literature on ethnic density presented in the introduction is quite thin and could be expanded to include the existing studies on ethnic density and obesity, diet, and obesity-related conditions as further justification for this work; the rationale for suggesting that 'results in the UK may differ' could be strengthened – would the suggested mechanisms for an ethnic density-obesity association not be the same as what would occur in the US, i.e., acculturation, stress, etc?. A better rationale for the work might also resolve the issue of what would be expected for White British participants, as acculturation and ethnic discrimination would not be the relevant mechanisms here.2. The implications of the work should be considered more deeply. While I support the authors' interest in the issue, the authors' suggested implication (lines 45-50) seems superficial. They might for example address the possibility that others might interpret a protective effect as a justification for residential ethnic segregation.3. The methods are insufficiently described.<ol style="list-style-type: none">a. In addition to the Study setting, the authors should also include a paragraph on the Study population. In any case, the Methods should begin with a thorough description of the population, including the years of data collection; how the three boroughs were selected; number of GP practices, how these practices were selected for inclusion in the study; how representative the practices and their patient populations are of London overall.
-------------------------	--

	<p>b. The authors might discuss the fact that just under half of the potentially eligible population did not have a valid BMI measure. Might this have affected representativeness of the sample?</p> <p>c. Explanation should be provided for how the authors chose the ethnic groups of interest. Not only the last paragraph of the results do the authors mention that the analyses were restricted to five ethnic groups, so excluding Chinese and Black Caribbean participants, although Table 1 includes these participants.</p> <p>d. The authors might consider analyzing BMI as a continuous rather than a categorical outcome, especially given the likelihood of different overweight/obesity cutoffs for different ethnic groups. This might also address the inclusion of overweight as an outcome in the sensitivity analyses; if a graded outcome is of interest, then either a continuous or ordinal polychotomous outcome might be more appropriate.</p> <p>e. Further justification or precedent should be provided for using MSOAs as the appropriate size to characterize area ethnic density.</p> <p>f. Adjustment for socioeconomic deprivation is important, and the authors did their best given the lack of information at the individual level. However, they might also consider adjusting for area-level socioeconomic deprivation, at the same level that they characterized ethnic density.</p> <p>g. The authors focus on obesity but then also include analyses with overweight/obesity as an outcome.</p> <p>4. As part of a description of the study population, the authors should also provide some information in the text on the extent/range of ethnic density experienced by participants in their study. They should also discuss this in comparison with other studies in UK and US for the sake of comparison. Figure 2 might show some aspect of this but is hard to see.</p> <p>Other minor comments:</p> <ol style="list-style-type: none"> 1. Many of the references seem dated, at least 10 years old. The authors might look into the more recent literature, even if it requires looking at US studies. 2. There are grammatical errors (mainly in the form of run-on sentences) throughout. 3. The Figures are of poor resolution and hard to see. Figure 1 might be replaced by a table. 4. The authors state some advantages over US studies including size of the geographical area; they should provide some more context on what sizes were used in those US studies. 5. The difference between Black African and Black Caribbean is an interesting point raised in the discussion and might be discussed a bit more to provide readers with somewhat more social/historical context for different immigration/settlement patterns, and their current socioeconomic position/opportunities.
REVIEWER	Ana Basto-Abreu National Institute of Public health, Mexico
REVIEW RETURNED	06-Aug-2018
GENERAL COMMENTS	<p>In this paper, the authors estimated the association between ethnic status and obesity in East London and the role of ethnic density in the obesity prevalence by sex and ethnic group. The topic is relevant to understand the patterns of obesity among ethnic groups. The authors found that the own ethnic density does not reduce obesity prevalence and, in some groups, might even increase it. In general terms the study was conducted under an</p>

	<p>appropriate methodology, but several methodological aspects must be addressed and improve it. Minor and major issues are described below.</p> <p>ABSTRACT</p> <ul style="list-style-type: none"> - The authors may not assess “risk” with the transversal study. Please avoid the word, as “risk” assumes causality. - Primary and secondary outcome measure. As far as I understood, the 10% increase in own-group ethnic density was not compared to white British. Please clarify. <p>INTRODUCTION</p> <ul style="list-style-type: none"> - “This” in line 25, page 5 is difficult to understand what does it refers to. - Line 45, page 5 – “In the United Kingdom (UK), such associations have been found in the areas of mental health and more recently in smoking behavior”. What did these studies find for the ethnic density effect (protective or a risk factor)? <p>METHODS</p> <ul style="list-style-type: none"> - Please spell GP at first use. - Line 9 page 8: please state the 16 ethnicity groups. Why in results you divided in 10 ethnicity groups? Please clarify - Line 27 page 8: please add space between 0.5 and its units. - Used MSOA with 2000 to 6000 households. How did you estimate the 0.5 km² and how does it reflect the walking environment? Please add some more explanations. - Maybe adding the figure 2 in the methods section and explain how density was estimated could help to better understand this indicator. - What do you consider a valid BMI? In results, over 340,000 individuals had invalid BMI measures? Please clarify it in methods. Socioeconomic deprivation. Could you please give more details how you constructed this indicator? Because any information was given, I needed to google the indicator to understand what does it measure. My first intuition was that more indicators were needed to adjust in the model. Did you include all the following dimensions? How did you measure them at individual level? <ul style="list-style-type: none"> o Income o Employment o Health deprivation and Disability o Education Skills and Training o Barriers to Housing and Services o Crime o Living Environment. - Line 52 page 8 “These each”? <p>Data analysis</p> <ul style="list-style-type: none"> - The second model requires more detail. Why authors decided a 10% increase in the ethnic density and not keep the variable continuous? This should be explained. <p>RESULTS</p> <ul style="list-style-type: none"> - Table 1. “Other” in the white subgroup include which ethnic population? Tables should be readable by their own. Maybe adding a table endnote will help understanding. - Table 3 should be similar to table 2, and keep male and female as head of table to facilitate compare comparability between tables.
--	---

	<ul style="list-style-type: none"> - Figure 2 represents methodologic aspects of the study. Why did the authors report it as results? - Page 11, line 48. How can the authors state that the same results are observed for overweight individuals, when adding overweight and obese individuals? - Page 11, line 52. "A 10% increase in Indian ethnic density was associated with a 7% (95%CI 0.88-0.99) decrease in odds of obesity" But in the table we observe an OR=0.94, which would represent a decrease of 6%, not 7%, please clarify. - Page 12, line 4. "Sensitivity analyses using different thresholds for ethnic density, ..." I could not find results varying this threshold. What happen if you would keep the own ethnic density variable as continuous? <p>DISCUSSION</p> <ul style="list-style-type: none"> - The main findings should be collapsed in one single paragraph and state just the main results (you do not need to refer tables in this part of discussion). - What is HSE, please spell it. - Page 14, line 7 "Which may more accurately reflect...". Please reformulate the sentence. - Some paragraphs lack strength and are not able to conclude an idea. Example paragraph between lines13 and 20. <p>Limitations</p> <ul style="list-style-type: none"> - How did you measure the residual confounding for the IMD, or its ability to measure the individual deprivation?
--	---

VERSION 1 – AUTHOR RESPONSE

Reviewer 1 comments:

The introduction adequately presents the research question but could be tightened and sharpened substantially. For example, information in paragraph 3 is somewhat redundant; the literature on ethnic density presented in the introduction is quite thin and could be expanded to include the existing studies on ethnic density and obesity, diet, and obesity-related conditions as further justification for this work.

We have expanded the section on ethnic density to discuss a greater range of the existing studies on ethnic density obesity and eating habits, along with an expanded section on acculturation.

The methods are insufficiently described.

a. In addition to the Study setting, the authors should also include a paragraph on the Study population. In any case, the Methods should begin with a thorough description of the population, including the years of data collection; how the three boroughs were selected; number of GP practices, how these practices were selected for inclusion in the study; how representative the practices and

their patient populations are of London overall.

We have expanded this paragraph to include more details of the study setting. The boroughs were selected because of the high proportion of ethnic minority populations, making it an ideal location for ethnic density studies. The numbers of practices are now included. In the discussion section we point out that this population is not representative of the general population, but that the results are of relevance to many inner urban, multiethnic populations.

b. The authors might discuss the fact that just under half of the potentially eligible population did not have a valid BMI measure. Might this have affected representativeness of the sample?

We have now included a paragraph in the limitations section discussing this limitation, and the other potential limitations of using routine clinical data for studies such as this.

c. Explanation should be provided for how the authors chose the ethnic groups of interest. Not only the last paragraph of the results do the authors mention that the analyses were restricted to five ethnic groups, so excluding Chinese and Black Caribbean participants, although Table 1 includes these participants.

We have added a further explanation in the methods. All the major ethnic groups were included in the study for the analysis of objective 1. *Odds of obesity for different ethnic groups compared to White British– ethnic density section*. These results are shown in Figure 1.

However we could only use a subset of ethnic groups to examine the influence of ethnic density on obesity, as for Black Caribbean and Chinese groups the ethnic density did not vary sufficiently over the geographical area for analysis. This point is included in the methods section.

d. The authors might consider analyzing BMI as a continuous rather than a categorical outcome, especially given the likelihood of different overweight/obesity cutoffs for different ethnic groups. This might also address the inclusion of overweight as an outcome in the sensitivity analyses; if a graded outcome is of interest, then either a continuous or ordinal polychotomous outcome might be more appropriate.

Many thanks for this suggestion. We choose to treat weight status as a categorical dichotomous outcome due to previous research conducted by the group using similarly developed methodology to examine ethnic density and smoking. We considered using the same methodology would be appropriate and by treating obesity as a dichotomous outcome, interpretation would be easier.

e. Further justification or precedent should be provided for using MSOAs as the appropriate size to characterize area ethnic density.

A more detailed explanation has been provided in the methods section.

f. Adjustment for socioeconomic deprivation is important, and the authors did their best given the lack of information at the individual level. However, they might also consider adjusting for area-level socioeconomic deprivation, at the same level that they characterized ethnic density.

We were unable to do this as we lacked measures of deprivation over this level.

g. The authors focus on obesity but then also include analyses with overweight/obesity as an outcome.

Yes, obesity was our primary outcome of interest. However, as part of our sensitivity analysis we explored different weight cut offs – thereby including those both overweight and obese (BMI over 25).

4. As part of a description of the study population, the authors should also provide some information in the text on the extent/range of ethnic density experienced by participants in their study. They should also discuss this in comparison with other studies in UK and US for the sake of comparison. Figure 2 might show some aspect of this but is hard to see.

We have kept figure 2 (now Appendix figure 1) as a useful visual representation of the ethnic diversity in the study area compared to another London area. However we have also added a summary of the own group ethnic density experienced by participants (average, min, max) in table 1.

Other minor comments:

Many of the references seem dated, at least 10 years old. The authors might look into the more recent literature, even if it requires looking at US studies.

We have kept older UK references, but now added some newer more recent references from the literature.

There are grammatical errors (mainly in the form of run-on sentences) throughout.

Thank you, we have shortened sentences and checked for grammatical errors throughout.

The Figures are of poor resolution and hard to see. Figure 1 might be replaced by a table.

We have kept figure 1 and added a tabulated version of the data (Appendix table 1) for editorial teams discretion as to which to use. We will upload higher resolution images.

The difference between Black African and Black Caribbean is an interesting point raised in the discussion and might be discussed a bit more to provide readers with somewhat more social/historical context for different immigration/settlement patterns, and their current socioeconomic position/opportunities.

We have added some information on the differences in migratory history in the discussion, but a detailed exploration of this is beyond the scope of this work.

Reviewer: 2 comments

In general terms the study was conducted under an appropriate methodology, but several methodological aspects must be addressed and improve it. Minor and major issues are described below.

ABSTRACT

~

The authors may not assess “risk” with the transversal study. Please avoid the word, as “risk” assumes causality.

This has been changed, and replaced with the term ‘odds’.

Primary and secondary outcome measure. As far as I understood, the 10% increase in own-group ethnic density was not compared to white British. Please clarify.

Yes. Increase in ethnic density and odds of obesity was examined for each ethnic group and not between ethnic groups. This was our primary outcome measure. We have clarified this in the abstract and throughout the report.

INTRODUCTION

“This” in line 25, page 5 is difficult to understand what does it refers to.

This has been changed, and much of the introduction has been expanded to improve clarity.

Line 45, page 5 – “In the United Kingdom (UK), such associations have been found in the areas of mental health and more recently in smoking behavior”. What did these studies find for the ethnic density effect (protective or a risk factor)?

We have expanded this section to provide details of ethnic density effects reported.

METHODS

Line 9 page 8: please state the 16 ethnicity groups. Why in results you divided in 10 ethnicity groups? Please clarify

The 16 categories of the UK census has a reference added.
In the study area only a subset of ethnic groups had sufficient variation in ethnic density across the study area to enable us to use them in this study. We have added further explanation on these points in the methods section.

Used MSOA with 2000 to 6000 households. How did you estimate the 0.5 km² and how does it reflect the walking environment? Please add some more explanations.

We have re-written this section to improve clarity

Maybe adding the figure 2 in the methods section and explain how density was estimated could help to better understand this indicator.

We have rewritten this part of the methods section with a more detailed explanation of how ethnic density was calculated.

What do you consider a valid BMI? In results, over 340,000 individuals had invalid BMI measures? Please clarify it in methods.

We have clarified this and removed the term 'valid'. We included individuals who had a BMI recorded in the last 3 years and were not pregnant at the time of weight measure – as stated in methods, subsection – obesity.

Socioeconomic deprivation.

Could you please give more details how you constructed this indicator? Because any information was given, I needed to google the indicator to understand what does it measure. My first intuition was that more indicators were needed to adjust in the model. Did you include all the following dimensions? How did you measure them at individual level?

- o Income

- o Employment
- o Health deprivation and Disability
- o Education Skills and Training
- o Barriers to Housing and Services
- o Crime
- o Living Environment.

IMD as an indicator of socioeconomic deprivation is commonly used in UK studies. The indicator is based on census data and linked to a small geographic area. The IMD score for individuals is thus based on an area based value, and does not measure the indicators at individual level. We have now expanded this section and referenced further reading.

It is well known that the IMD indicator is susceptible to the ecological fallacy – in that not all individuals are representative of the area statistics they live in. This limitation is mentioned in the discussion

Data analysis

The second model requires more detail. Why authors decided a 10% increase in the ethnic density and not keep the variable continuous? This should be explained.

We used 10% as the threshold based on the effect seen in previous studies. We have rewritten this section to clarify its choice. Ethnic density in the model was a continuous variable, we used the lincom function to compare coefficients at different ranges of ethnic density and produce odds ratio, which we felt would be easier to interpret.

RESULTS

Table 1. “Other” in the white subgroup include which ethnic population? Tables should be readable by their own. Maybe adding a table endnote will help understanding.

We have added an endnote to the table, and made it clear in the methods which groups are included in “Other”.

Table 3 should be similar to table 2, and keep male and female as head of table to facilitate compare comparability between tables.

Many thanks for this suggestion, table 3 has been changed to be more consistent with table 2.

Page 11, line 52. "A 10% increase in Indian ethnic density was associated with a 7% (95%CI 0.88-0.99) decrease in odds of obesity" But in the table we observe an OR=0.94, which would represent a decrease of 6%, not 7%, please clarify.

This has been corrected.

Page 12, line 4. "Sensitivity analyses using different thresholds for ethnic density, ..." I could not find results varying this threshold. What happen if you would keep the own ethnic density variable as continuous?

This sentence has been corrected. Ethnic density was kept as a continuous variable in the analysis, but to aid interpretation, and based on previous studies, we transformed co-efficient and gave odds ratios based on linear combinations at 10% increase in thresholds of ethnic density.

DISCUSSION

The main findings should be collapsed in one single paragraph and state just the main results (you do not need to refer tables in this part of discussion).

-
This has been changed

Some paragraphs lack strength and are not able to conclude an idea. Example paragraph between lines 13 and 20.

Thank you. We have rewritten and expanded much of the discussion to improve clarity.

Limitations

How did you measure the residual confounding for the IMD, or its ability to measure the individual deprivation?

We are unable to measure residual confounding and IMD accuracy at measuring individual deprivation, we have discussed this in our limitations. However IMD is widely used in UK based studies as a proxy for individual deprivation when individual measures are unavailable.

VERSION 2 – REVIEW

REVIEWER	Marilyn Tseng California Polytechnic University, San Luis Obispo
REVIEW RETURNED	03-Oct-2018

GENERAL COMMENTS	<p>The authors responded to many of the comments on the original version, but several of the concerns remain.</p> <ol style="list-style-type: none"> 1. The authors acknowledge as a limitation that just under half of the potentially eligible population did not have a BMI measure but should address this further. Did missingness vary across ethnic groups? By residential ethnic density? The extent to which the findings might be due to selection bias should be more thoughtfully and thoroughly evaluated if data are available to evaluate this. 2. Socioeconomic deprivation at the level at which ethnic density was measured is a potentially important confounder, and another potential explanation for the findings. The authors should address this so readers have a sense for the extent to which this is truly a problem – for example, how well might LSOA-level deprivation capture MSOA-level? They should also explain that deprivation information was not available at the MSOA level. 3. I am not convinced by the authors' argument that interpretation of obesity as a dichotomous variable is easier than interpretation of BMI as a continuous variable, especially given the necessity for different cutpoints to define obesity in different ethnic groups. At the very least, it could be treated as an ordinal variable. <p>Other minor comments:</p> <ol style="list-style-type: none"> 1. Page 9 lines 1-10 seems redundant with page 8 lines 49-51. 2. Define LSOA when first introduced in the manuscript. 3. The point about insufficient variability in ethnic density for Black Caribbean and Chinese populations is not in the Methods section, contrary to the authors' rebuttal statement. 4. The presentation of results based on overweight as an outcome should more clearly be presented as sensitivity analyses, and these analyses should also be justified. 5. Information on 'other' Whites should be in text as well as table. The authors stated that they did this in the rebuttal, but it is not present in the Methods section of the revision. 6. The correction of 7% to 6% decrease for Indian ethnic density should be throughout the manuscript. 7. The authors state in their rebuttal that, 'In the discussion section we point out that this population is not representative of the general population, but that the results are of relevance to many inner urban, multiethnic populations.' But I do not see this in the submitted revision. 8. In the discussion the authors' statement that, 'Our data suggest that environmental factors such as food, exercise and cultural norms' is an overstatement. Your data suggest that there are other factors besides ethnic density but provide no information as to what those factors are.
-------------------------	---

REVIEWER	Ana Basto National Institute of Public health, Mexico
REVIEW RETURNED	18-Oct-2018

GENERAL COMMENTS	The present study has been greatly improved, comparing with the previous version. While the authors answered all my points, I do
-------------------------	--

	<p>not consider these answers “point by point”. Answers are too ambiguous, such as “We have re-written this section to improve clarity”, which makes difficult to follow what was changed in the manuscript. Thus, I cannot tell exactly if the authors addressed all my points.</p> <p>Reviewing the new manuscript, I believe it needs to be improved before publication. It is well written, but the structure is quite disordered. For example, abstract misses the statistical analysis. The introduction is excessively large and lacks a direction. In the second paragraph of the introduction, the authors spend 8 lines talking about the association between mental health and smoking with ethnic density while the objective of the paper is assessing obesity rates and their association with ethnic density. It drives the reader’s attention to other problems that will not be developed in the manuscript. Maybe it will be needed during discussion; for introduction it could be useful to say that ethnic density has been associated with lower odds of mental disorders and smoking and drive the attention to the obesity problem.</p> <p>The topic and the analysis are very interesting and valuable. I would recommend authors to write a more detailed answer to reviewers to facilitate our job to assess changes in the manuscript.</p>
--	--

VERSION 2 – AUTHOR RESPONSE

Reviewer 1:

The authors responded to many of the comments on the original version, but several of the concerns remain.

1. The authors acknowledge as a limitation that just under half of the potentially eligible population did not have a BMI measure but should address this further. Did missingness vary across ethnic groups? By residential ethnic density? The extent to which the findings might be due to selection bias should be more thoughtfully and thoroughly evaluated if data are available to evaluate this.

We agree that no-random absence of data (common in studies using the electronic health record) may be a source of bias. To address this further we have added the following paragraph in the strengths and limitations section.

“We found differences in the proportions of individuals with a BMI recorded in the last 3 years by ethnic group. The highest completeness was among Black Caribbean, with 80.9% having a BMI taken in the last 3 years, and lowest among Bangladeshis with 61.6%. These differences may reflect variation by ethnic groups in their use of primary care services.⁴² This may lead to differential recording as a source of bias. It is difficult to determine the direction of bias. Those individuals with no recent measure of BMI were younger and free of chronic diseases.”

Socioeconomic deprivation at the level at which ethnic density was measured is a potentially important confounder, and another potential explanation for the findings. The authors should address this so readers have a sense for the extent to which this is truly a problem – for example, how well

might LSOA-level deprivation capture MSOA-level? They should also explain that deprivation information was not available at the MSOA level.

We are unable to measure deprivation at the MSOA level, only the LSOA level which we use as a proxy measure for individual deprivation and incorporate into our analysis. Deprivation at LSOA level is likely to capture deprivation at MSOA level well, although we cannot quantify this. However in view of the small geographical areas MSOA operates on (1 MSOA = approximately 4-5 coterminous LSOA combined) MSOA deprivation is likely reflected by LSOA deprivation.

In the discussion, 'strength and limitations' we now acknowledge this with the following paragraph:

"We were also unable to measure deprivation at the level at which ethnic density was recorded. Our proxy measure of individual deprivation, IMD, is measured at the LSOA level and likely to have captured deprivation at the level ethnic density was recorded. However, it is possible it did not fully capture deprivation thereby potentially masking any ethnic density effect. "

3. I am not convinced by the authors' argument that interpretation of obesity as a dichotomous variable is easier than interpretation of BMI as a continuous variable, especially given the necessity for different cutpoints to define obesity in different ethnic groups. At the very least, it could be treated as an ordinal variable.

The analysis could have been conducted keeping BMI as a continuous variable. However we chose to categorise BMI, based on our understanding that the resulting analysis would be easier to convey. Furthermore we wanted our statistical analysis to mirror previous analysis done on ethnic density and smoking by our group, which treated smoking as binary (smoker/non smoker).

Other minor comments:

1. Page 9 lines 1-10 seems redundant with page 8 lines 49-51.

Page 8 lines 49-51:

Ethnic density was calculated as the percentage of people from each ethnic group living within each Middle Super Output Area, obtained from the 2011 census.³² We used census data rather than primary care data for the calculation of ethnic density as the census data includes the entire population, including individuals who may not be registered with a GP practice.

Page 9 lines 1-10

We chose Middle-layer Super Output Areas (MSOA), with average population size of 7,790, for the geographic area to calculate ethnic density as this has been used in other studies in England to identify the effect of ethnic density on health outcomes.¹⁷ The study area of east London has a population density of 12,600 individuals per square kilometer, and we estimate an average MSOA covers 0.6 km² in this area, representing the walking environment for individuals.

This has been re-written as:

Ethnic density was calculated as the percentage of people from each ethnic group living within each Middle Super Output Area (MSOA), obtained from the 2011 census. MSOA has been used previously as the geographical area to calculate ethnic density and consists of an average population size of 7,790. In the study area of east London, with a population density of 12,600 individuals per square kilometer, we estimate an average MSOA covers 0.6 km² in this area.

2. Define LSOA when first introduced in the manuscript.

Corrected. First introduced in “Methods”, subheading – “Socio-economic deprivation”.

3. The point about insufficient variability in ethnic density for Black Caribbean and Chinese populations is not in the Methods section, contrary to the authors’ rebuttal statement.

We have amended this section in the methods to read:

“We found that ethnic density varied sufficiently across our study area to be used for analysis among White British, Black African, Indian, Pakistani and Bangladeshi groups only; there was insufficient variability in ethnic density among Black Caribbean and Chinese populations for further analysis in our study (See figure 2).”

4. The presentation of results based on overweight as an outcome should more clearly be presented as sensitivity analyses, and these analyses should also be justified.

The results are in appendix table 2, which incorporates all sensitivity analysis. We have not changed this and feel the results are well placed there with all the other sensitivity analysis grouped together.

We have made reference to this table in the text and changed the text to make it clear this was a sensitivity analysis.

5. Information on ‘other’ Whites should be in text as well as table. The authors stated that they did this in the rebuttal, but it is not present in the Methods section of the revision.

We thought the reviewer was referring to the “mixed/other” ethnic group. Which is clarified in the text and the table with a footnote on table 1 stating

“Mixed/other (=White and Black Caribbean, White and Black African, White and Asian, Other mixed, Other Black, Other Asian, Other ethnic group)”

The “White other” ethnic grouping has no further subdivision, it would be individuals who self report their ethnicity as White other, (and therefore not White British or White Irish) so would include other European nationals.

6. The correction of 7% to 6% decrease for Indian ethnic density should be throughout the manuscript.

Corrected in all places (discussion, main findings)

7. The authors state in their rebuttal that, ‘In the discussion section we point out that this population is not representative of the general population, but that the results are of relevance to many inner urban, multiethnic populations.’ But I do not see this in the submitted revision.

We do not include this as a statement.

Our population is not representative of the UK population with East London having a higher proportion of ethnic minorities compared to UK. The area is described in the methods, ‘study setting’ section, which was not present in the original manuscript, but added following reviewers comments. We note that the ethnic mix allows this study to be done, which would not be possible in other areas. This study is more explorative in nature, and hypothesis generating rather than results being generalizable to other areas.

8. In the discussion the authors’ statement that, ‘Our data suggest that environmental factors such as food, exercise and cultural norms’ is an overstatement. Your data suggest that there are other factors besides ethnic density but provide no information as to what those factors are.

Many thanks we have amended this overstatement, to more accurately state:

Our data suggest ethnic density does not play protective role in preventing obesity; it may be that environmental factors such as food, exercise and cultural norms play a greater role in

determining obesity than the ethnic composition of the area.

Reviewer 2:

The present study has been greatly improved, comparing with the previous version.

Thank you, we made major changes to the manuscript following the first review.

Reviewing the new manuscript, I believe it needs to be improved before publication. It is well written, but the structure is quite disordered. For example, abstract misses the statistical analysis.

We now mention the statistical analysis in the abstract.

The Introduction is excessively large and lacks direction

This was expanded following reviewers initial concerns. Following further feedback we have now reduced the introduction and narrowed its focus.

In the second paragraph of the introduction, the authors spend 8 lines talking about the association between mental health and smoking with ethnic density while the objective of the paper is assessing obesity rates and their association with ethnic density. It drives the reader's attention to other problems that will not be developed in the manuscript.

As above, this has been summarized more succinctly as:

“Most research on ethnic density has been conducted in the area of mental health where increasing own ethnic density has a protective effect on a range of mental health outcomes.^{9,15 16 17-19}

²⁰ Beyond mental health, in the United Kingdom (UK) increasing ethnic density has been found protective against smoking, where increasing own-group ethnic density was associated with reduced odds of smoking.”

The Editor has requested that we elaborate on our responses to the original comments from Reviewer 2. We have done this below, and included updated responses following our second revision.

ABSTRACT The authors may not assess “risk” with the transversal study. Please avoid the word, as “risk” assumes causality.

This has been changed, and replaced with the term ‘odds’.

Primary and secondary outcome measure. As far as I understood, the 10% increase in own-group ethnic density was not compared to white British. Please clarify.

You are correct. Increase in own group ethnic density and odds of obesity was examined for each ethnic group and not between ethnic groups.

INTRODUCTION

Line 45, page 5 – “In the United Kingdom (UK), such associations have been found in the areas of mental health and more recently in smoking behavior”. What did these studies find for the ethnic density effect (protective or a risk factor)?

In the original manuscript we state:

In the United Kingdom (UK), such associations have been found in the areas of mental health and more recently in smoking behaviour.¹¹⁻¹³

In the submitted revised manuscript we expanded this section to summarise the previous research on ethnic density more fully, adding the following the paragraph:

“Most research on ethnic density has been conducted in the area of mental health where increasing own ethnic density has a protective effect on a range of mental health outcomes.^{9,15 16 17-19} A recent systematic review and meta-analysis observed a 18% reduction in relative odds of psychotic experiences and 12% reduction in relative odds of suicidal ideation for each 10% increase in own-group ethnic density.²⁰ Beyond mental health, in the United Kingdom (UK) increasing ethnic density has been found protective against smoking, where a 10% increase in own-group ethnic density was associated with a 2–43% reduction in the odds of being a current smoker varying according to ethnic group.^{17”}

However in view of Reviewer 2 concern that this may be off topic and the introduction is too large, we have summarised this to:

“Most research on ethnic density has been conducted in the area of mental health where increasing own ethnic density has a protective effect on a range of mental health outcomes.^{9,15 16 17-19 20} Beyond mental health, increasing ethnic density has been found protective against smoking, where increasing own-group ethnic density was associated with reduced odds of smoking.”

METHODS

Please spell GP at first use.

Corrected.

Line 9 page 8: please state the 16 ethnicity groups. Why in results you divided in 10 ethnicity groups? Please clarify

In the original manuscript we stated:

“Reported ethnicity was collapsed to the 16 groups as defined in the 2011 UK Census.²⁷”

In the revised manuscript we state:

“Reported ethnicity was collapsed into the 16 groups of the 2011 UK Census,²⁹ these were further collapsed into 9 groups: White British (White British, White Irish), White Other, Black African, Black Caribbean, Indian, Pakistani, Bangladeshi, Chinese, Mixed/other (White and Black Caribbean, White and Black African, White and Asian, Other mixed, Other Black, Other Asian, Other ethnic group).”

Used MSOA with 2000 to 6000 households. How did you estimate the 0.5 km² and how does it reflect the walking environment? Please add some more explanations.

We have added more explanation on our choice of ethnic density measured at the MSOA level and the method in which 0.6km² (previously was 0.5km²) was calculated. In methods, section ‘Ethnicity and ethnic density’, the following paragraph has been included:

“MSOA has been used previously as the geographical area to calculate ethnic density and consists of an average population size of 7,790. In the study area of east London, with a population density of 12,600 individuals per square kilometer, we estimate an average MSOA covers 0.6 km² in this area.”

We have now removed the term ‘walking environment’, this was a subjective impression of the likely area in which a resident would experience most of their daily experience, suggesting it would be reasonable to measure ethnic density at this level.

Maybe adding the figure 2 in the methods section and explain how density was estimated could help to better understand this indicator.

Figure 2 is now referred to in the methods section and we provide a more detailed explanation of how ethnic density was calculated including its level of geographical level of measurement and rationale for its choice in the methods, section ‘Ethnicity and ethnic density’.

What do you consider a valid BMI? In results, over 340,000 individuals had invalid BMI measures? Please clarify it in methods.

We have removed the term ‘valid’ and state in the methods, subsection “Obesity” how participants were included based on BMI:

“Participants were included if they had at least one BMI measure in the previous three years (2014-2017), the most recent BMI was used. Women who were pregnant at the time of BMI measurement were excluded.”

Socioeconomic deprivation.

Could you please give more details how you constructed this indicator?. Did you include all the following dimensions? How did you measure them at individual level?

- o Income
- o Employment
- o Health deprivation and Disability
- o Education Skills and Training
- o Barriers to Housing and Services
- o Crime
- o Living Environment.

IMD is an indicator of socioeconomic deprivation commonly used in UK studies and we provide a reference for this. The indicator is based on census data and linked to a small geographic area. The IMD score for individuals is thus based on an area based value, and does not measure the indicators at individual level.

In the methods, subsection 'Socio-economic deprivation' we have inserted the following paragraph:

"A proxy measure of individual socio-economic deprivation (Index of multiple deprivation, IMD) was obtained for each patient based on their LSOA (Lower Super Output Area) of residence. IMD is a widely used measure of relative deprivation in England, combining information on seven domains of deprivation (income, employment, education, health and disability, housing and living environment) from census data.³³

Data analysis

The second model requires more detail. Why authors decided a 10% increase in the ethnic density and not keep the variable continuous? This should be explained.

We used 10% as the threshold based on the effect seen in previous research. Ethnic density in the model was a continuous variable, we used the lincom function to compare coefficients at different ranges of ethnic density and produce odds ratio. This methodology and analysis is consistent with the groups previous published research on ethnic density and smoking.

We now acknowledge this is in the manuscript with the following sentence and references in the methods, subsection 'Data-analysis':

"In common with previous studies on ethnic density effects, we choose 10% as the threshold interval above which an association with a change in the odds of being obese was sought".^{17 19}

RESULTS

Table 3 should be similar to table 2, and keep male and female as head of table to facilitate compare comparability between tables.

Table 3 has been changed as suggested.

Figure 2 represents methodologic aspects of the study. Why did the authors report it as results?

We have removed this from the results section and placed in the methods as suggested.

Page 11, line 52. "A 10% increase in Indian ethnic density was associated with a 7% (95%CI 0.88-0.99) decrease in odds of obesity" But in the table we observe an OR=0.94, which would represent a decrease of 6%, not 7%, please clarify.

This has been corrected throughout

Page 12, line 4. "Sensitivity analyses using different thresholds for ethnic density, ..." I could not find results varying this threshold. What happen if you would keep the own ethnic density variable as continuous?

The sensitivity analysis referred to changing the method in which we calculated ethnic density. We have made this clearer by rewriting this in the methods, subsection 'Data-analysis' and include the following line:

"We also repeated the analysis using ethnic density calculated from primary care records rather than census data."

In the analysis ethnic density was kept as a continuous variable. However in reporting we transformed the co-efficient and gave odds ratio based on linear combinations at 10% increase in thresholds of ethnic density, in keeping with the groups previous work on ethnic density.

DISCUSSION

The main findings should be collapsed in one single paragraph and state just the main results.

This has been changed

What is HSE, please spell it.

Health Survey for England – corrected in manuscript

Limitations

How did you measure the residual confounding for the IMD, or its ability to measure the individual deprivation?

We are unable to measure residual confounding and IMD accuracy at measuring individual deprivation, we have discussed this in our limitations. However IMD is widely used in UK based studies as a proxy for individual deprivation when individual measures are unavailable.

VERSION 3 – REVIEW

REVIEWER	Marilyn Tseng California Polytechnic State University, San Luis Obispo
REVIEW RETURNED	30-Jan-2019

GENERAL COMMENTS	An important question, efficiently addressed using existing data.
-------------------------	---

REVIEWER	Ana Basto National Institute of Public Health, Mexico.
REVIEW RETURNED	29-Jan-2019

GENERAL COMMENTS	<p>3RD ROUND</p> <p>Thank you for providing clearer answers to my previous comments. The draft is much improved. I think it is a good study with a large sample, well executed and providing insights about the prevalence of obesity among ethnic groups and how ethnic density may change the obesity prevalence.</p> <p>Abstract</p> <p>According to your results and discussion, the objective of this article should be two (please state them in the order that you present in the article to facilitate reading and comprehension) – as you stated in the introduction.</p> <p>In abstract you still have the 7% decrease in odds of obesity instead of 7%. “Among Indian females, a 10% increase in Indian ethnic density was associated with a 7% decrease in odds of obesity (95% CI 0.88-0.99).”</p> <p>Introduction</p> <p>I am still not convinced about the introduction. It is shorter which helps, but it is somehow difficult to read and lacks a direction. First paragraph has excessive amount of ideas. Also, I found the use “this may be”, “These in turn” (2 times) does not facilitate comprehension.</p> <p>Methods</p> <p>Please explain why diabetes is independently associated with obesity. Diabetes does not modify the risk of obesity?</p> <p>Results</p> <p>I think table 1 in appendix repeats information.</p> <p>Discussion</p> <p>3rd paragraph- It is only the age distribution that reduces the obesity prevalence in comparison with the overall country? Make sure not to make overstatements.</p> <p>The author mention: “It is difficult to determine the direction of bias. Those individuals with no recent measure of BMI were younger and free of chronic diseases”. Age and chronic diseases are usually risk factors for obesity. If the direction of bias tends to reduce the obesity prevalence what would be consequences for the study. The authors added the proportion of missingness for BMI by ethnic group, which helps, but it would be more informative to present the mean age by ethnic group between the participants (having BMI measures) and not participants, as a sensitivity analysis (in appendix).</p>
-------------------------	---

VERSION 3 – AUTHOR RESPONSE

In response to the reviewers’ comments we have made the following changes:

1. Abstract objective has been changed into a 2 part sentence: *To examine the prevalence of obesity by ethnic group, and to examine the association between ethnic density and obesity prevalence.*
2. The 6% decrease in odds of obesity among Indian female in the text has been corrected to 7%.
3. The first paragraph has been modified with shorter sentence structure.
4. Appendix Table 1 (the tabulated data used for Figure 2) has been removed.
5. In addition we agree that differences in age and other factors are likely to account for the differences in obesity prevalence amongst our cross-section compared to national data.
6. We have added the following text about the impact of missing data:

Those individuals with no recent measure of BMI were younger and free of chronic diseases and therefore less likely to be obese, resulting in an overestimate of obesity prevalence in those ethnic groups with higher proportions of missing BMI data. However, while this may impact our estimates for obesity prevalence between ethnic groups, it is unlikely to have effected the examination of ethnic density and obesity prevalence within ethnic groups.

Reviewers' comments:

Reviewer: 2

Reviewer Name: Ana Basto

Thank you for providing clearer answers to my previous comments. The draft is much improved. I think it is a good study with a large sample, well executed and providing insights about the prevalence of obesity among ethnic groups and how ethnic density may change the obesity prevalence.

Abstract

According to your results and discussion, the objective of this article should be two (please state them in the order that you present in the article to facilitate reading and comprehension) – as you stated in the introduction.

In abstract you still have the 7% decrease in odds of obesity instead of 7%. “Among Indian females, a 10% increase in Indian ethnic density was associated with a 7% decrease in odds of obesity (95% CI 0.88-0.99).”

Introduction

I am still not convinced about the introduction. It is shorter which helps, but it is somehow difficult to read and lacks a direction. First paragraph has excessive amount of ideas. Also, I found the use “this may be”, “These in turn” (2 times) does not facilitate comprehension.

Methods

Please explain why diabetes is independently associated with obesity. Diabetes does not modify the risk of obesity?

Results

I think table 1 in appendix repeats information.

Discussion

3rd paragraph- It is only the age distribution that reduces the obesity prevalence in comparison with the overall country? Make sure not to make overstatements.

The author mention: "It is difficult to determine the direction of bias. Those individuals with no recent measure of BMI were younger and free of chronic diseases". Age and chronic diseases are usually risk factors for obesity. If the direction of bias tends to reduce the obesity prevalence what would be consequences for the study.

The authors added the proportion of missingness for BMI by ethnic group, which helps, but it would be more informative to present the mean age by ethnic group between the participants (having BMI measures) and not participants, as a sensitivity analysis (in appendix).

Reviewer: 1

Reviewer Name: Marilyn Tseng

An important question, efficiently addressed using existing data.