~~Comments  
This paper describes an ecological study of the reach of parkrun in terms of ethnic diversity in parkrun. The paper addresses one of the outstanding issues in the evidence base on this physical activity phenomenon and uses a novel (to parkrun research) method to investigate it. The paper is essentially well-written, has interesting and useful findings and has a sound methodological approach. The question of the effect of clustering within the definition of the participation measure is my only substantive concern, and just requires explanation (see below). Otherwise I have a number of minor comments which should be considered before accepting this article for indexing.~~  
Introduction  
  
Par 1, line 1: I don’t think “collection” is the right word – it implies that they are in one place when parkrun’s main asset is that it is disseminated. A small wording revision should address this.

**Agreed - use Network!**

Par 4, last sentence: word missing: density would have…

“We thus hypothesised that at the community level, areas with higher ethnic density have lower levels of participation in parkrun.”

**I think we are correct.**

It might be helpful to have a line about physical activity rates among ethnically diverse populations here – sport is not the only form of physical activity and therefore you need to demonstrate that total activity, which is what matters for health, is also lower than for other groups/communities.  
   
Methods  
   
Under data sources, could the authors please indicate whether “finishers” were unique or just a total count ignoring repeat participation (this is mentioned only in the limitations but should be earlier). The authors should describe what the potential implications are of this for the analysis and interpretation – they do mention that they do not expect it change the results in the limitations but do not provide the rationale for such a conclusion. It also begs the question why they did not use unique persons because it would be possible to do this with parkrun data. In other words they should explain why they chose to operationalise participation this way.

**Yes, need to indicate this is a total count ignoring repeat participation. Will add this in to the updated version.**  
   
Percentage working age – is there any reason why the authors chose this particular variable for age and how does it relate to the objectives of the analysis?

**We know from previous work that parkrun is highest in the age group from (16-65). Because the geographic areas are relatively small (around 2000) some areas may have a disproportionate number of older people who may be less likely to engage in parkrun. This may bias the results if these areas are, for example, also more rural or have a higher percentage white-British persons.**  
   
LSOAs - Could the authors describe why this particular level of spatial classification was used – parkruns draw on varying areas depending on population density but also the proximity to other parkruns. Could they also state whether there were any cases where two or more SOAs were equidistant and if so how were they allocated.

**The individual parkruns are not the unit of analysis, the LSOA are. Therefore, LSOA are not allocated.**  
   
Data analysis – the authors should describe what assumptions for poisson regression were tested (over-dispersion for example). They should also describe how age was operationalised.

   
Results  
   
The authors talk about ethnic diversity but do not give the reader much idea about what ethnicities this covers in these areas. The authors should describe this somewhere (intro, methods, results) to give the non-UK reader some further context.

We did not look at Percentage of population non-White-British  
   
Par 1: third last line: Remove “the parkrun distance” as it is not relevant as such to the point being made.

**This can be easily removed.**  
   
Par 2: The result for age has been reversed in the results from how it was described in the methods which actually makes it more difficult to understand. Is there any reason why you talk about % non-working age rather than % working age? At the very least it should be consistent between methods and results.  
  
Par 3: you make reference to major areas and urban minor areas but have not defined this anywhere. Either here or in the methods would be suitable.  
   
Par on Model 3: As I read the table, despite the attenuation of the effects for IMD and Ethnic density they remained significant in the model – should be explicitly stated in the text.  
  
Limitations  
   
The authors should also note that this research was conducted in one country and the associations may be different in other countries with different geo-demographic patterns and parkrun density.

**True - need to add a comment on this.**

~~Is the work clearly and accurately presented and does it cite the current literature?~~

~~Yes~~

~~Is the study design appropriate and is the work technically sound?~~

~~Yes~~

~~Are sufficient details of methods and analysis provided to allow replication by others?~~

~~Partly~~

~~If applicable, is the statistical analysis and its interpretation appropriate?~~

~~Yes~~

~~Are all the source data underlying the results available to ensure full reproducibility?~~

~~Yes~~

~~Are the conclusions drawn adequately supported by the results?~~

~~Yes~~

~~Competing Interests~~

~~I am actively engaged in parkrun research and from time to time submit approval requests to the parkrun Global Research Board of which a number of the authors are a member.~~

~~Reviewer Expertise~~

~~Social epidemiology and evaluation in the prevention of lifestyle-related chronic disease.~~

~~I confirm that I have read this submission and believe that I have an appropriate level of expertise to confirm that it is of an acceptable scientific standard.~~