What determines community level parkrun participation?

Smith R[1],Schneider PP[1]

[1]ScHARR

Keywords: `parkrun`, `Participation`, `Physical Activity`, `Equity`, `Ethnicity`.

Intended Journal: Journal of Racial & Ethnic Health Disparities, Sports Medicine Open,Behavioural and Social Sciences, Journal Of Epi & Comm Health.

Abstract

There is substantial hetrogeneity in parkrun participation across different communities in England. Previous anaysis has shown that communities that live further from the nearest parkrun, and communities in deprived areas have lower participation rates (Schneider et al. 2019).

This paper is the first to make use of data at the LSOA level, provided by parkrun and available through the ONS, to investigate the determinants of parkrun participation.

We find that deprivation, distance to nearest event and ethnicity are strong predictors of parkrun participation rates. This creates an opportunity for parkrun to increase participation through engagement with these communities.

Introduction

There is strong evidence from survey data that ethnic minority individuals in England are less likely to be physcially active, and to engage in sport in general Sport England.

Previous analysis has been conducted investigating the factors associated with higher levels of partcipation Cleland et al., 2019. However, due to the biases associated with surveying only parkrun participants the findings have no external validity (i.e. estimating the relationship between education and number of runs once engaged with parkrun does not tell us about the determinants of initial engagement with parkrun). The authors themselves note the limitation that "The sample was limited to a nonrandom sample of parkrun participants in one State of Australia and may not be generalizable to other parkrun populations." (p.21).

Previous research found IMD very important and ethnic diversity low goyder et al. 2018 "community events of this nature are, whilst likely to provide health benefits to the population as a whole, also likely to be responsible for significant intervention generated inequalities in uptake of opportunities for physically active recreation."

Parkrun huge potential impact Stevinson et al. 2013. and Stevinson et al. 2019.

Methods

Data was obtained from multiple sources at the Lower layer Super Output Level, which are geographical areas containing around 1,000-3,000 people. There are 32844 LSOAs in England. parkrunUK provided data on the number of runs from each LSOA in England between the dates XXX and XXX. The ethnic breakdown of each LSOA was obtained from the Office of National Statistics (ONS) here. The deprivation scores & population sizes for each LSOA was obtained from the ONS. More information can be found here. The population density for each LSOA was obtained from the ONS here. The geographic weighted centroids for

each LSOA were obtained from the ONS here. The location of each parkrun event was obtained from the parkrunUK website here.

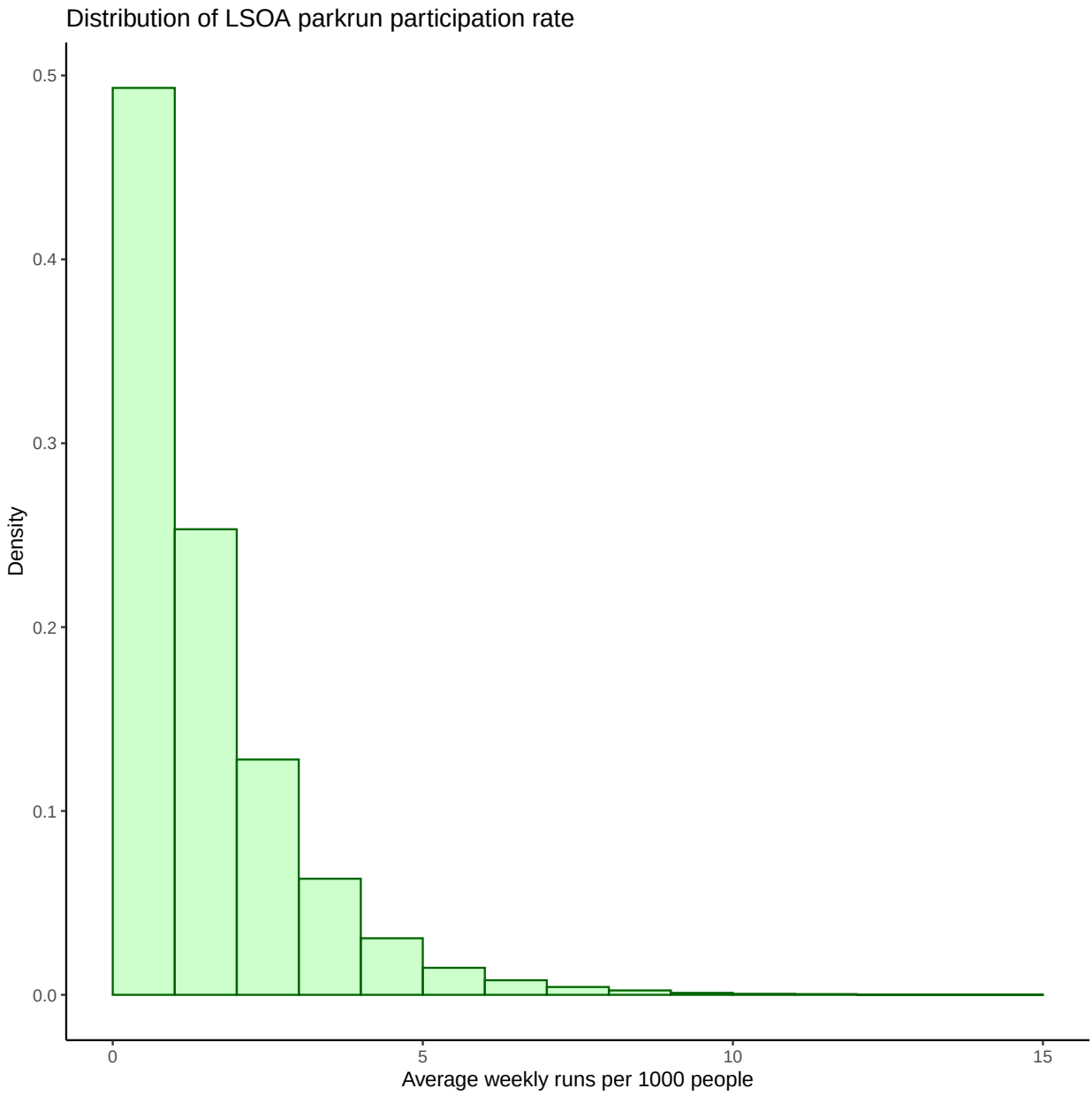
Variable	Description	Source
run_count	number of runs from each LSOA in England between XXX and XXX	parkrunUK
imd	IMD scores for each LSOA	https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/835115/IoD2019_Statistical_Release.pdf
total_pop	total number of individuals in each LSOA	https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/835115/IoD2019_Statistical_Release.pdf
pop_density	population density for each LSOA	https://www.ons.gov.uk/peoplepopulationandcommunity/populationandmigration/populationestimates/datasets/lowersuperoutputareapopulationdensity
perc_bme	percent of population non-white-british	https://www.ons.gov.uk/peoplepopulationandcommunity/populationandmigration/populationestimates/datasets/lowersuperoutputareamidyearpopulationestimates
mn_dstn	distance from LSOA centroid to nearest parkrun	derived
perc_non_working_age	derived from ONS data on age-groups in each LSOA	derived
run_rate	derived from run_count and LSAO populations	derived

After merging these datasets we had detailed data on 32844 LSOAs participation in parkrun and several characteristics of the LSOAs which we hypothesised may influence participation. Since previous work has found corelations between participation and deprivation, distance to nearest event, and population density we included all of these variables. We also extended the analysis to include the percent of the population of black and ethnic minority and the percent of the population who are working age. We included ethnicity as we hypothesised that areas with higher percentages of residents of ethnic minorities would have lower participation rates. We included the percent of the population that is working age as a control to limit for the effect of populations heavily skewed toward older people (e.g. care homes), or very young people (e.g orphanages/immediately around special needs schools)). Since participation in parkrunUK is dominated by those aged 20-60 we felt this was justified.

We used R Software Environment to produce simple summary statistics and viewed simple correlations between the variables described. We then moved on to estimate the partial correlation coefficients between the variables, before running a poisson regression model on scaled data to estimate the relative influence of several determinants on parkrun participation.

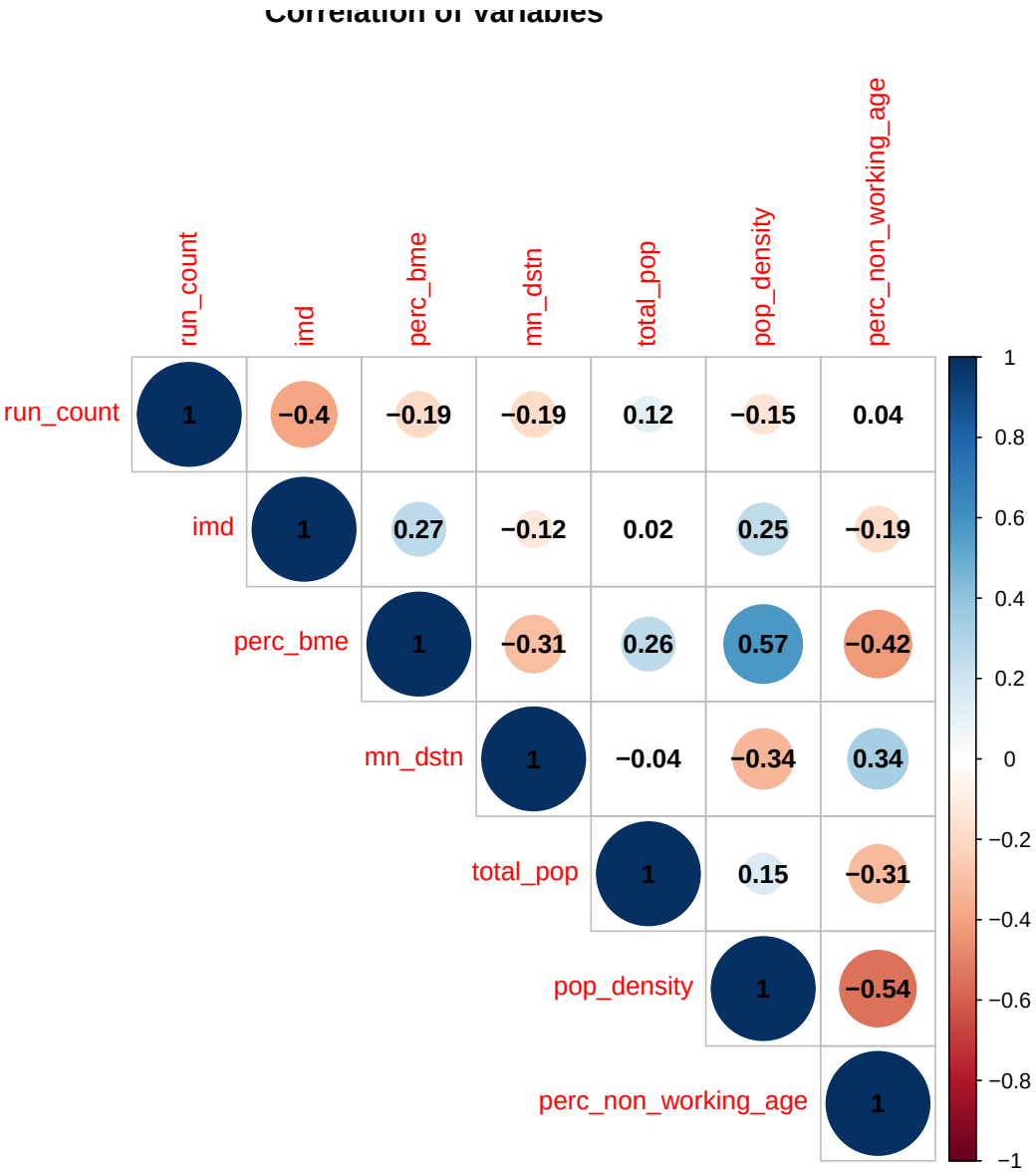
Results

Participation in parkrun varies between LSOAs. Around half of all communities (LSOA) average less than 1 run per week per 1000 people. Approximately a quarter average between 1 and 2 runs, and around an eighth between 2 and 3 runs.



Correlation Matrix

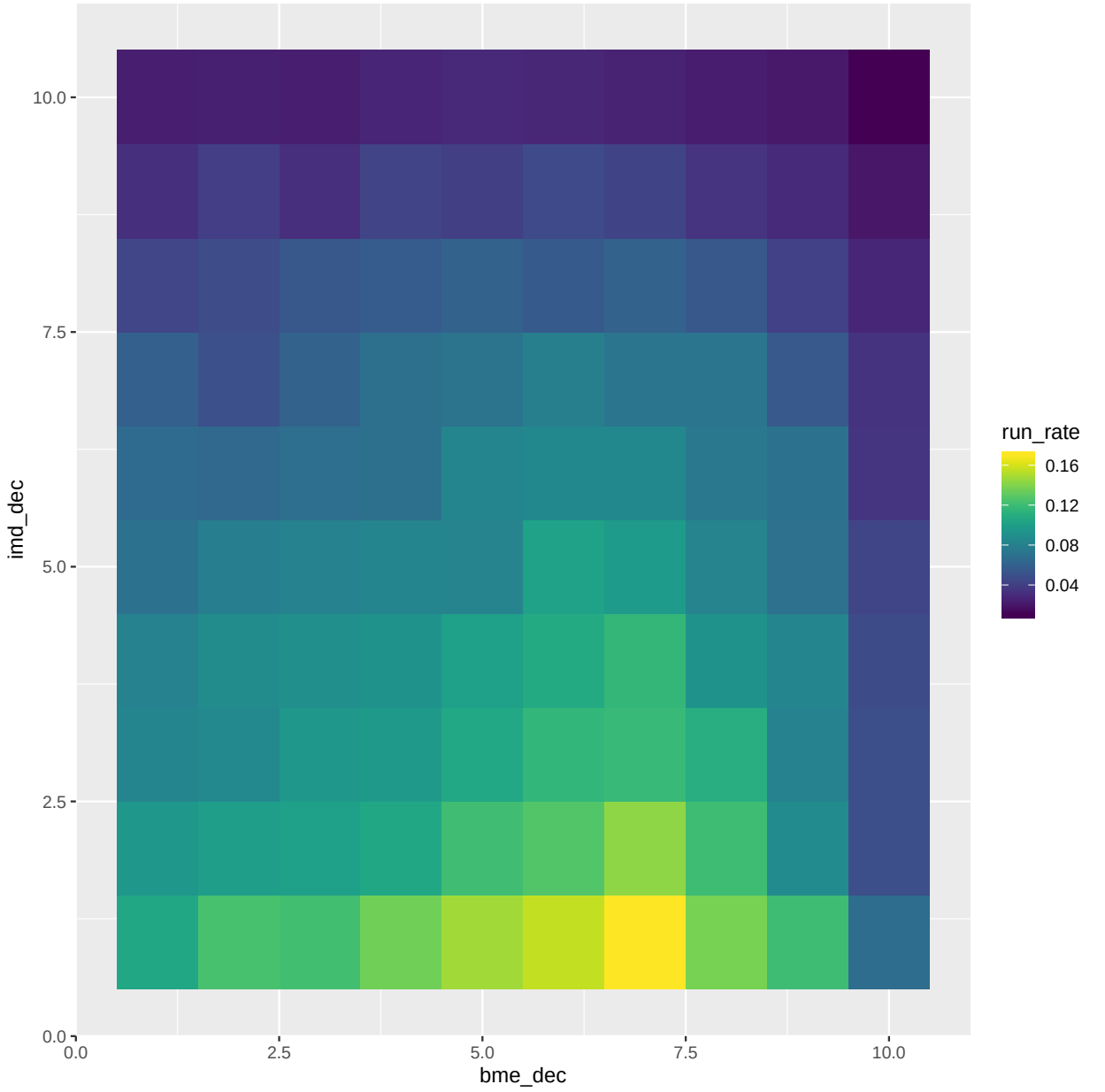
There is a negative correlation between particpation (run_count) and deprivation (imd), distance to nearest parkrun (mn_distance), population density (pop_density) and percent BME (perc_bme). Percent BME was strongly postively correlated with population density, negatively correlated with percent non-working age, and moderately positvely correlated with IMD suggesting that areas with more BME residents are more densely populated, more deprived and have fewer older people.



Colour plot

We are interested in particular in how participation in parkrun varies by the percent BME and IMD of the community. We split our LSOA into deciles based on the two variables and represent the mean participation rate for LSOAs which fall into the repsective deciles (e.g. 2,5 means decile 2 for BME and 5 for IMD) as a colour on the plot.

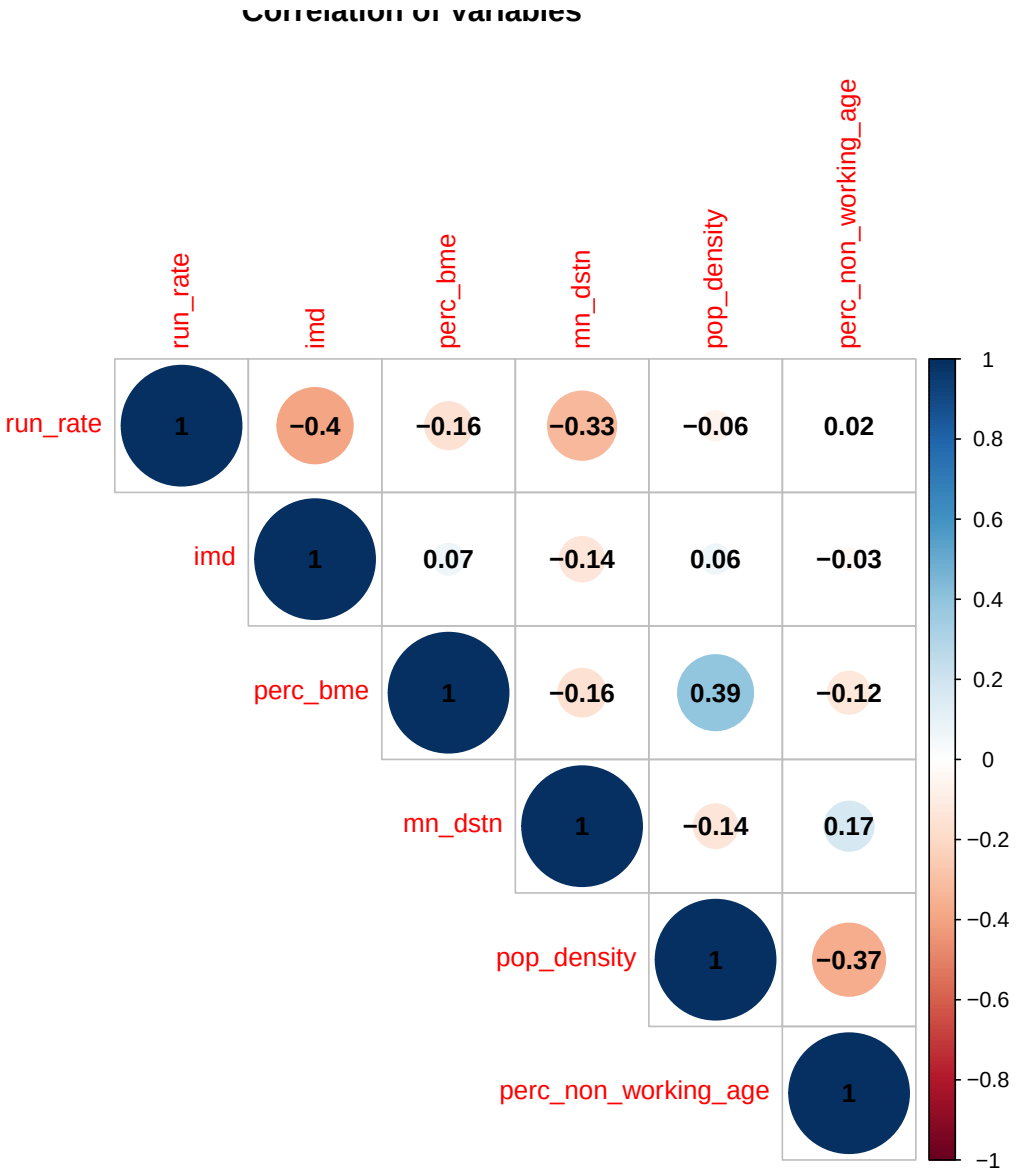
The plot shows that participation is lowest in areas with the highest deprivation, and that there is a complex relationship between ethnicity and participation, with higher participation in areas with moderately high percentage BME.



In order to control for confounders, we created a partial correlation matrix. This estimates the correlation between the variables in the analysis holding all other variables constant.

Looking at the top row, it is clear that participation rates and deprivation is strongly correlated, but less so now we are controlling for other variables. Distance is also strongly correlated, as our previous analysis published in XXX showed. The percent BME is also negatively correlated, suggesting that some of the relationship between participation and IMD is attributable to the ethnicity of residents.

Partial correlation matrix



Poissson Model

Poisson models are commonly used to estimate count data, where values are constrained by 0. In this case, because it is not possible for a community to have negative participation we use a Poisson regression.

The results of the poisson regression are not notoriously easy to interpret, but show that areas with a higher percent of residents who are non-White-British have lower participation rates, even when controlling for the effect of deprivation and distance to events. The effect is smaller than deprivation and distance, but still sizeable and significant.

```
Call:
glm(formula = run_count ~ imd + mn_dstn + perc_bme, family = poisson(link = "log"),
    data = scaled_df, offset = log(total_pop))

Deviance Residuals:
    Min       1Q   Median       3Q      Max
-42.127   -6.675   -1.900    3.823   46.513

Coefficients:
              Estimate Std. Error z value Pr(>|z|)
(Intercept) -2.7983239  0.0006064 -4614.7   <2e-16 ***
imd         -0.5315137  0.0007063  -752.5   <2e-16 ***
mn_dstn     -0.4242098  0.0007148  -593.5   <2e-16 ***
perc_bme    -0.3203611  0.0007516  -426.3   <2e-16 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

(Dispersion parameter for poisson family taken to be 1)

    Null deviance: 3703296  on 32843  degrees of freedom
Residual deviance: 2297949  on 32840  degrees of freedom
AIC: 2490104

Number of Fisher Scoring iterations: 5

[1] 0.3794865
```

Discussion

Previous analysis by Scheider et al., 2019 has showed that participation in parkrun is lower in more deprived communities, and communities that are further from their nearest parkrun. This paper extends that analysis to include other community level characteristics which were hypothesized to influence parkrun participation. Our findings show that some of the negative effect on participation previously attributed to deprivation can actually be attributed to differences in participation by ethnic minotiries.

We find that communities with higher percentages of ethnic minority residents were less likely to participate in parkrun, even when controlling for deprivation. Since percentage ethnic minorotiy residents and deprivation are positively correlated, it is likely that some of the effect attributed to deprivation in our previous analysis Scheider et al., 2019 is is attributable to ethnicity. In short, the socioeconomic gradient in parkrun participation is shallower than perviously thought, but the ethnic gradient revealed.

That communities with a higher percentage of ethnic minority residents have lower participation rates in parkrun may not be a public health problem in and of itself. Different communities may freely choose to take part in different types of physical activity for many reasons. But given physical actiivty levels in these communities are also lower (REF), it seems there is an opportunity for parkrun to develop methods of reaching groups who have low rates of physical activity engagement, with huge public health potential.

Limitations

This analysis is cross-sectional, using data on participation in parkrun events between 1st January and 10th December. We have not studied how participation rates have changed over time, therefore it is possible that participation is changing at different rates in different LSOAs, and therefore deprived communities and

communities with higher proportions of ethnic minority residents may be decreasing or increasing relative to the population average. It would be interesting to replicate this analysis for each year from 2010-2020 to see how parkrun has grown over time.

Our dependent variable is the number of runs by residents of each LSOA. This is a count variable where each run is treated equally (e.g. 10 runs by one person is equal to 1 run by 10 people). We cannot therefore draw inferences on the number of people who took part within each LSOA at some point in the year, but instead focus on the total run count.

This analysis has been conducted at the LSOA level. Therefore we have been careful to say that participation is lower in areas with higher proportions of ethnic minority residents. It is not possible to make inferences about the effect of deprivation and ethnicity at a individual level from this data. However previous qualitative analysis Goyder et al.,2018 has identified the small number of participants from ethnic minorities as an avenue for further research.

References