

Multiple deprivation and geographic distance to community sport events — achieving equitable access to parkrun in England

Schneider PP*,¹, Smith RA¹, Bullas AM², Quirk H², Bayley T¹, Haake SSJ², Brennan A¹, Goyder E¹

¹*School of Health and Related Research, University of Sheffield, Sheffield, UK.*; ²*Advanced Wellbeing Research Centre, Sheffield Hallam University, Sheffield, UK.*

Introduction

Insufficient physical activity is a leading cause of disease and disability worldwide.[???] In the UK, around one in six deaths is attributable to low levels of physical activity.[1] It is also a major contributor to health inequalities, as people from low socio-economic background are more likely to be inactive [2] and are disproportionately affected by physical inactivity-related diseases [3]. Increasing the physical activity levels of the population is therefore high on the public health agenda: it not only has the potential to improve quality of life, reduce mortality rates and alleviate the strain on health and social care services, but also reduce the gap in health inequalities.

However, designing effective public health interventions that increase population physical activity is considerable challenge.[???] Implementing such interventions in a way that does not increase health inequalities might even be more difficult, as studies have shown that programmes to increase physical activity often fail to reach deprived communities and those most in need.[???,??]

parkrun, an international movement which organises free weekly 5km running events, might provide valuable lessons for public health professional with respect to both. Since its founding in 2004, as a small event in London with 13 participants, it has grown to become one of the world's largest mass participation sporting events, with up to 360,000 participants in more than 20 countries.[???] The volunteer-led running events are often characterised as accessible and inclusive.[???] The organisation has been widely praised as being successful in encouraging participation, particularly in individuals who were previously inactive or did not meet public health guidelines.[???,??]

Notwithstanding these subjective accounts, the expansion of parkrun in England, as elsewhere, has been largely grassroots, driven by *demand* rather than *need*. It might therefore be the case that parkrun events are primarily located in areas that are relatively affluent, while people living in more deprived communities may not have the same opportunities to participate. In 2019, Sport England announced funding to support the creation of 200 new parkrun events across England within three years, with the specific aim of increasing participation of individuals from lower socio-economic groups.[???]

The aims of this study are two-fold: firstly, to evaluate whether geographic access to parkrun events in England is equitable across areas with different levels of deprivation. Secondly, to identify 200 optimal locations for future events in order to improve geographic access, in particular for deprived communities.

Methods

Study design

This study is a cross-sectional ecological analysis of the socio-economic disparities in geographic access to parkrun events in England at the end of 2018. All analyses were conducted on the level of Lower Layer Super Output Areas (LSOAs), which divide England into 32,844 geographic units which, on average, have a population of approximately 1,700. We assessed the relationship between access, defined as the distance (as the crow flies) to the nearest parkrun event, and socio-economic deprivation, measured using the Index of Multiple Deprivation. In addition, we used information on public green spaces in England to conducted a simple location-allocation analysis, in order to identify 200 locations for future parkrun events, that maximise access for the population.

Data sources

For this study, we combined data on three types of geospatial entities: 1) LSOAs, 2) parkrun events, and 3) public green spaces.

- 1) For all 32,844 LSOAs, we retrieved geographic locations, defined by the coordinates of its population-weighted centroid, 2017 total population estimates, and the 2015 Index of Multiple Deprivation (IMD) from the Office for National Statistics [???,??,??].
- 2) We included all 465 public parkrun events which were in operation in England by December 12th, 2018 - on this date, Sport England announced their plan to provide funding for setting up 200 additional parkrun events across England.[???] The locations of the events were manually scrapped from the parkrun UK website.[???]
- 3) The locations of public green spaces in England were retrieved from an open dataset of Ordnance Survey.[9] Even though parkrun events do not always require a 5km loop (some courses involve running back and forth) and are held in various settings and terrains, not all green spaces are suitable to host running events. After evaluating existing parkrun event courses, we decided to consider all public parks, gardens, and playing fields with an area of 0.1 km² (e.g. 316m x 316m) or more potentially suitable for hosting events (n=2,842).

Variables

The two variables of interest were access to parkrun events and deprivation of LSOAs.

Access to parkrun was measured in terms of the distance as the crow flies (in km) to the nearest event. For each of the 32,844 LSOAs, we computed the geodesic distances between its population-weighted centroid and all 465 parkrun events that were in operation on December 12th, 2018, and then selected the shortest distance.

The socio-economic deprivation of LSOAs was measured using the 2015 IMD. It is a measure of relative deprivation, which has been used in many similar studies. The IMD combines 37 indicators from seven domains (income, employment, education and skills, health and disability, crime, housing and services, and living environment) into a single score. The score ranges from 0 (least deprived) to 100 (most deprived).[???]

Other covariates, which are likely to affect the availability of parkrun events within an area (e.g. population density or demographics), were not taken into account, because we did not aim to assess to what extent deprivation independently ‘explained’ access. Rather, we sought to evaluate whether or not people living in deprived areas have better or worse access, given the actual circumstances.

Analysis

Mean, standard deviation, median, interquartile range, and range were used as descriptive statistics. We then assessed the association between the IMD and the distance to the nearest parkrun event on the LSOA level. Our hypothesis was that more deprived areas had worse access, i.e. longer distances to the nearest parkrun event than more affluent areas. Pearson and Spearman correlation coefficients were computed using the LSOAs’ total population as weights. We also conducted a stratified analysis, for which we grouped LSOAs into IMD quintiles (most, more, median, less, least deprived) and assessed access to parkrun events in each strata.

Identifying optimal locations for new parkrun events

We conducted a location-allocation analysis to solve the following problem: parkrun UK received funding to start 200 additional parkrun events. There are 2,842 public green spaces in England in which new events could be set up. Which 200 locations should be selected, in order to maximise access for the greatest number of people?

More specifically, the objective was to minimise the population-weighted total sum of distances between all LSOAs and their nearest parkrun event. To identify the optimal 200 green spaces, we applied a simple greedy algorithm that consisted of two steps. Firstly, for each green space, we evaluate how setting up a parkrun event would affect the sum of distances, given the locations of all existing events (i.e. for how many LSOAs would this green space then be the nearest parkrun event, and by how much would it decrease the respective distances). Secondly, the green space with the greatest effect is selected and added to the set of existing parkrun events. This procedure is repeated 200 times.

More formally, the first step of the algorithm is defined by the following equation:

$$\underset{c \in C}{\operatorname{argmin}} \sum_{i=1}^{32,844} d_i(E \cap c) * p_i$$

The function yields the candidate green space c , from the set of all 2,842 green spaces C , which minimises the sum of the products of the minimum distances and the total population of all LSOAs, whereby p_i denotes the population of LSOA i , and $d_i(E \cap c)$ denotes the distance to the nearest parkrun event, which can either be an existing event from the set $E = \{e_1, e_2, \dots, e_{465}\}$, or the candidate green space c , whichever is nearest.

In order to identify the optimal 200 new locations for setting up new parkrun events, the selection procedure is repeated 200 times. At each step, the single best candidate green space location is selected, added to the set of established parkrun events E , and removed from the set of available green spaces C . This means, the optimal green space selected at step k is taken into account when selecting the k th+1 location.

Evaluating the location allocation analysis

We assessed the overall impact of setting up 200 new parkrun events at the identified optimal location on the geographic access to parkrun events in England. In order to conduct a distributional analysis, we investigated the effect of setting up the 200 new events on LSOAs across IMD quintiles.

Data and source code availability

All data and the R source code that were used to generate the results of this study are provided on an open repository.^{[???}

Ethical approval

Ethical approval was obtained from the Sheffield Hallam University Ethics Committee (ER10776545). We did not collect any personal information, but only used aggregate secondary data. The parkrun Research Board approved this research project, and four of its members (AMB, HQ, EG, SSJH) were actively involved in the interpretation of findings and writing of this manuscript.

Results

Descriptive statistics

As of 12th December 2018, approximately 7%, 69%, and 91% of the English population lived within 1, 5, and 10km distance of a parkrun event. Only 578,043 people (1% of the English population) lived more than 20km from an event. The average (SD) and median (IQR) distance to the nearest parkrun event were 4.65 (4.22) and 3.39 (1.99-5.83). The worst access was observed for 2,259 people living on the Isles of Scilly, who live about 76.44km (as the crow flies) away from the next parkrun event on the mainland. Further descriptive statistics are provided in table 1.

Table 1: Descriptive statistics of LSOAs and parkrun events

	Mean (SD)	Median (Q25-Q75)	Range
LSOAs (n = 32,844)			
Population	1,693 (405)	1,612 (1,452; 1,834)	362-13,404
IMD	21.67 (15.59)	17.40 (9.65; 30.07)	0.48-92.60
Distance to the nearest event	4.65 (4.17)	3.39 (1.99; 5.83)	0.04-76.44
Parkrun events (n = 465)			
Catchment area* population	119,612 (74,290)	103,952 (68,837; 151,488)	7,855-628,010
Catchment area* LSOAs	71 (43)	62 (40; 87)	6-350

* Number of LSOAs/total population for which a given parkrun event is the nearest.

Association between Deprivation and Access

There was a negative relationship between IMD and the distance to the nearest parkrun event: the (population weighted) Pearson and Spearman correlation coefficients were -0.15 and -0.18, indicating a small effect. This means, more deprived LSOAs tended to have shorter distances to the nearest parkrun event, i.e. better access.

The distributional analysis of distances by IMD quintile showed that people living in the 20% most deprived LSOAs had the best geographic access, with an average and median distance to the nearest parkrun event of 3.51 and 2.79 km. Depending on the metric, the worst access was observed for LSOAs in the middle (mean distance = 3.36km) or the less deprived group (median distance = 3.93km). Further results of the distributional analysis are shown in Table 2 in the following section.

Optimal locations for new parkrun events

Figure 1 shows the parkrun events (black circles) that existed on 12th December 2018 alongside recommendations for 200 additional event locations (red triangles), which maximise overall access to parkrun for the greatest number of people. The numbers correspond to the rank, where 1 is the location which would improve access the most. The names and exact locations of the selected 200 green spaces are provided in the appendix.

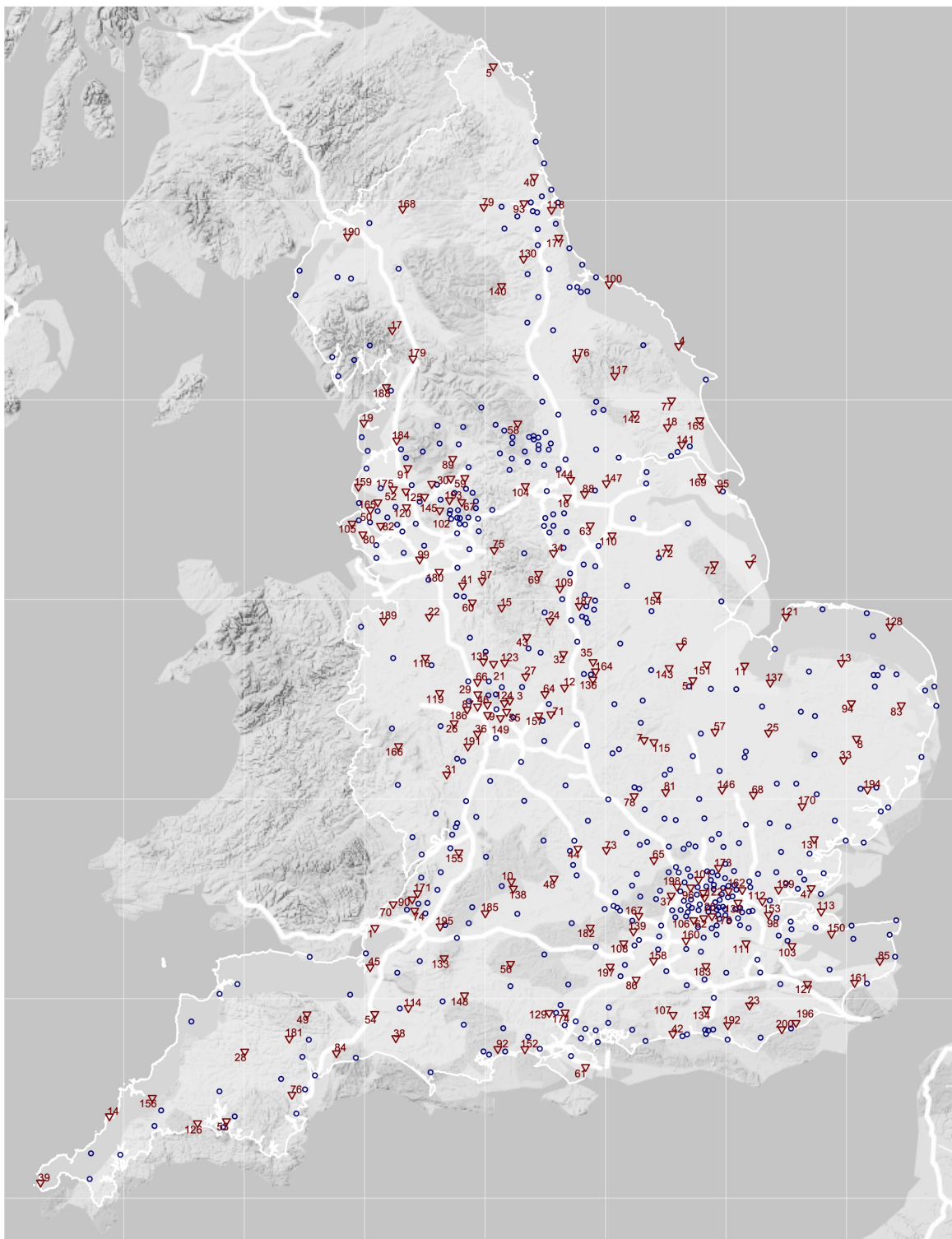


Figure 1: Map of England showing current parkrun events (blue circles) and recommended new event locations (red triangles) ranked in descending order of estimated effect on overall population-weighted access. Information on all 200 identified optimal green space locations are provided in the appendix.

We estimated that setting up new parkrun events in those 200 green spaces would improve access for around 16.5 million people (30% of the population) from 9,854 LSOAs. For those, the distance to the nearest event would, on average, be reduced by 4.09 km ($SD = 3.97$). Overall, it would reduce the average and median distance to the nearest parkrun event from 4.65 and 3.39 km to 3.43 and 2.59 km. The percentage of people who live within 5km of a parkrun would increase from 69% to 81.9%. Figure 2 shows the distribution of distances between LSOAs and their nearest parkrun events before and after the introduction of the 200 new optimal event locations at the identified optimal green spaces.

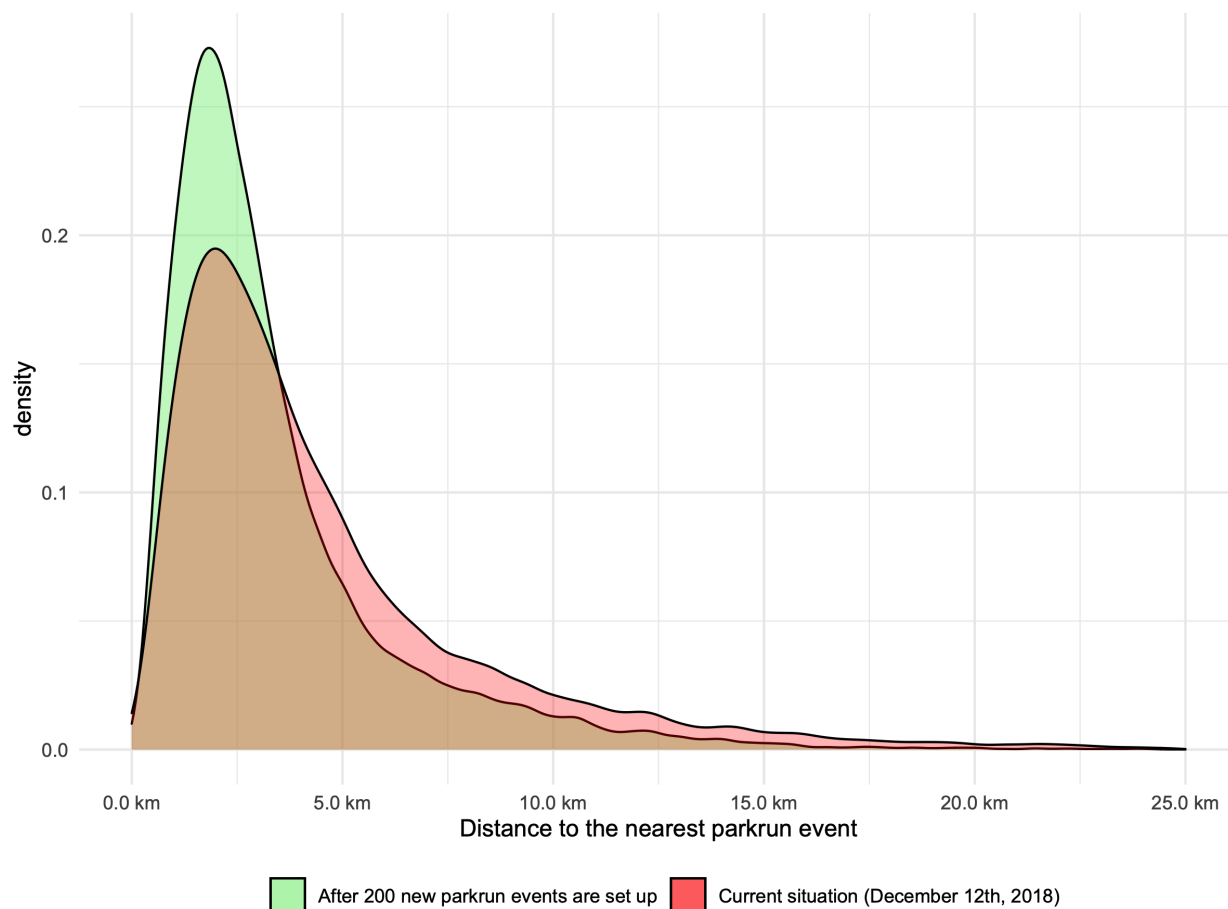


Figure 1: Figure 2

Figure 2: Density plot showing the distance, as the crow flies, to parkrun events before and after the creation of 200 new events in identified green-spaces.

Table 2 shows further results for the effects on LSOAs across IMD quintiles. Overall, setting up 200 new events in the recommended green spaces would amplify the negative association between IMD and geographic access: the population-weighted Pearson and Spearman correlation coefficients changed from -0.15 and -0.18 before, to -0.20 and -0.23 afterwards, indicating that improvements in access to parkrun events were greater for more deprived LSOAs as compared to less deprived LSOAs. Nevertheless, the distributional analysis showed that the improvement in access was smallest for LSOAs in the most deprived quintile.

Table 2: Distributional analysis

	Current situation (December 12th, 2018)			After 200 new parkrun events are set up		
	Mean (SD)	Median (Q25-Q75)	Range	Mean (SD)	Median (Q25-Q75)	Range
Least deprived	4.93 (3.62)	3.91 (2.27; 6.67)	0.12-58.54	3.79 (2.61)	3.09 (1.92; 4.96)	0.12-25.58
Less deprived	5.21 (4.24)	3.93 (2.28; 6.99)	0.14-76.44	3.92 (3.04)	2.99 (1.84; 5.09)	0.14-48.02
Median deprived	5.36 (5.01)	3.68 (2.12; 6.83)	0.11-60.81	3.98 (3.55)	2.79 (1.70; 4.91)	0.11-33.74
More deprived	4.26 (4.38)	2.96 (1.76; 5.00)	0.04-59.44	3.03 (2.78)	2.27 (1.47; 3.49)	0.04-24.07
Most deprived	3.51 (3.01)	2.79 (1.71; 4.39)	0.07-36.17	2.43 (1.68)	2.12 (1.41; 3.02)	0.05-24.30
Overall	4.65 (4.17)	3.39 (1.99; 5.83)	0.04-76.44	3.43 (2.86)	2.59 (1.63; 4.16)	0.04-48.02

Discussion

Overall, geographic access to parkrun in England is good. On 12th December 2018 more than two thirds of the population lived within 5km (the parkrun distance) of a parkrun event. Contrary to our expectation, we did not find that access was better for people living in more affluent areas. In fact, those living in the most deprived areas had the best geographic access to parkrun.

Our analysis has shown that setting up 200 new events in the optimal green spaces in England would increase the percentage of English residents who live within 5k of a parkrun to 82%. This would increase the inequity in access to parkrun events, improving the geographic access of the most deprived more than the access of those living in more affluent areas. It is rare in public health for inequity to exist in this direction.

While our analysis has not been concerned with confounding factors, we are interested only in geographic access based on multiple deprivation, an explanation for the negative relationship between IMD and access is population density. Since deprived areas cluster in cities, geographical access as the crow flies tends to be better for these areas in part because population density is itself a predictor of the existence of parkrun events.

There are several strengths of this study, it is the first study of geographic access to parkrun in England - therefore the approach is novel and the data untapped. Secondly the analysis makes use of large and rich datasets, with over 30,000 LSOA and over 400 existing parkrun events it is unlikely that individual outliers are affecting the results.

However, there are also limitations. Most importantly geographic access is not travel distance, or travel time. In some cases, for example where natural barriers like lakes or rivers block routes, the actual distance travelled may be far in excess of the distance as the crow flies. Also, in rural areas where car use is more prevalent travel time may be shorter, despite travel distances being longer. Further investigation could determine how perceptions of access change with distance in different settings (e.g. rural vs urban).

The choice of 0.1sq-km for greenspaces may also surprise some readers. However, several parkruns exist in very small areas (Mile End park in London, and Cudmore Grove park in Essex are 0.14sq-km and 0.12sq-km respectively). We therefore chose to maximize sensitivity, at the likely expense of specificity, and therefore it is possible that some of the recommendations are in parks that are not suitable to host a parkrun event. Conversely, many blue spaces such as beaches and promenades are not included in the database and therefore there may exist some locations which we have missed, particularly in coastal areas.

The main finding, that geographical access to parkrun events is better in more deprived communities, is surprising given that we know that participation in parkrun has a strong socioeconomic gradient with higher participation rates in more affluent areas (ref). This suggests that providing the opportunity to participate in parkrun events, while a necessary first step to enable participation, has not been enough to engage deprived communities and reduce inequity.

Investigating other barriers to participation in parkrun is likely to improve our understanding of the reasons why physical activity levels are lower in more deprived areas in general, and give parkrun the information it

needs to develop strategies, other than improving access, to engage those who live in deprived communities. Future research should build on this work and develop a model to assess the (cost-)effectiveness of setting up new events, and other strategies, not only in terms of improved potential access, but actual participation. This requires estimating the causal and marginal effects of different interventions on participation, and therefore physical activity levels, using longitudinal data and sophisticated modelling techniques.

parkrun has been hugely successful in encouraging physical activity worldwide, it will achieve its vision of a ‘healthier, happier planet’, but the extent of its impact will be dependent on how well it is able to understand the drivers of participation and develop strategies to over-come the hurdles it identifies.

Conclusion

References

1. Lee I-M, Shiroma EJ, Lobelo F, Puska P, Blair SN, Katzmarzyk PT, et al. Effect of physical inactivity on major non-communicable diseases worldwide: An analysis of burden of disease and life expectancy. *The lancet*. 2012;380(9838):219–29.
2. Farrell L, Hollingsworth B, Propper C, Shields MA. The socioeconomic gradient in physical inactivity: Evidence from one million adults in england. *Social science & medicine*. 2014;123:55–63.
3. Tanaka A, Shipley MJ, Welch CA, Groce NE, Marmot MG, Kivimaki M, et al. Socioeconomic inequality in recovery from poor physical and mental health in mid-life and early old age: Prospective whitehall ii cohort study. *J Epidemiol Community Health*. 2018;72(4):309–13.

Appendix