Does More Greenspace Lead to Less Obesity?
Exploring Which Children's Facilities Have the Greatest Impact on Childhood Obesity Rates

1. Introduction

1.1 Background

Childhood obesity has been widely regarded as one of the most critical public health challenges of the 21st century(World Health Organization, 2020). From 1990 to 2022, the obesity rate among adolescents aged 5 to 19 rose from 2% to 8%, reaching 160 million individuals(World Health Organization, 2024).

The prevalence of obesity is influenced by a range of complex factors, including birth weight, preterm birth, cesarean delivery, breastfeeding history, secondhand smoke exposure, outdoor activity time, and parental traits(Mears *et al.*, 2020; Chen *et al.*, 2024).Therefore, green spaces play a vital role in promoting children's physical activity, highlighting the importance of designing environments that attract visits and encourage active participation(Mears *et al.*, 2020; Nicky Hawkins, Kevin Levay, 2021).

1.2 Literature Review

Recent research suggests that exposure to green spaces, like parks and woodlands, may benefit human health, with some studies focusing on the link between green space accessibility and usage.(Cohen *et al.*, 2007; Björk *et al.*, 2008) . Other studies suggest that, residents' satisfaction with the quality of nearby green spaces significantly enhances the likelihood of engaging in exercise within these areas(Pyky *et al.*, 2019) . Researchers have also investigated key mediating factors explaining the link between green spaces and health, covering various types of green spaces and both physical and mental health outcomes.(Lachowycz and Jones, 2013).The use of green spaces may also be influenced by individual factors, such as age and exercise habits (Cohen *et al.*, 2007).

Most studies have focused on general populations rather than specific subgroups. Additionally, the core question of "whether, why, and how green spaces impact health" remains underexplored, with inconsistent and inconclusive empirical findings(Lachowycz and Jones, 2013; Zhou, von Lengerke and Dreier, 2021).

In this paper, we build on existing literature and focus specifically on children in London to further explore the complex relationship between green spaces and health outcomes. Drawing on the mixed empirical evidence described above, we exam the role of accessibility, quality and utilisation of green space as potential mechanisms for driving health benefits. The findings are then used to the local analyse, in order to assist in the planning of research and to reduce childhood obesity rates.

2 Data

2.1 Obesity data

This study uses data from the National Child Measurement Programme (NCMP), tracking childhood weight categories in England for Reception (ages 4–5) and Year 6 (ages 10–11) from 2006/07 to 2019/20.

To calculate the overall childhood obesity rates for each borough in 2019, we used the weighted formula provided by NCMP. This formula aggregates obesity prevalence across Reception and Year 6, considering the total number of children measured in each group. Specifically:

$$Overall\ Obesity\ Rate\ (\%) = \frac{Reception\ Obese\ Children + \ Year\ 6\ Obese\ Children}{Total\ Reception\ Children + Total\ Year\ 6\ Children} \times 100 \qquad (1)$$

2.2 Greenspace data

The greenspace data is sourced from the OS Open Greenspace dataset. But, it excludes certain categories not relevant to the Green Space Index (allotments, community growing spaces, cemeteries, religious grounds, and golf courses). Only relevant typologies such as bowling greens, other sports facilities, play spaces, playing fields, public parks or gardens, and tennis courts are retained for analysis(Fields in Trust, 2018), as shown in Figure 1.

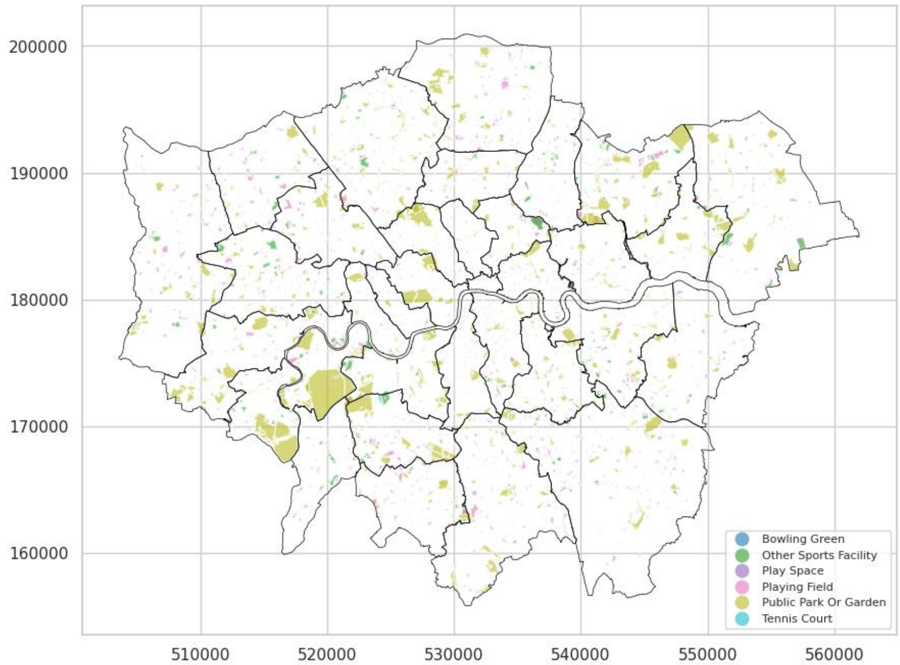

Figure 1 Greenspaces within London Boundary by Type

2.3 Deprivation data

Deprivation in this study was assessed using the 2019 Indices of Deprivation (IMD) dataset from the Ministry of Housing, Communities & Local Government (MHCLG), offering a comprehensive measure of relative deprivation in England.

IMD was included as a variable in this study because social factors, such as poverty levels, are significant contributors to obesity rates. By accounting for deprivation, we aim to improve the explanatory power of the regression equations. Additionally, the IMD dataset provides population counts for each ward, which were used to calculate green space variables.

3 Methodology

To explore the relationship between green space and obesity, we employed a multinomial logistic regression model for data analysis. Based on a review of the literature, we incorporated green space variables such as accessibility, coverage, per capita green area, and the proportion of different green space types (Cohen *et al.*, 2007; Lachowycz and Jones, 2011). Additionally, we included the Index of

Multiple Deprivation (IMD) as a socioeconomic factor to examine its potential interaction with green space variables in influencing childhood obesity rates.

We first conducted variable selection through correlation analysis and Box-Cox transformations to normalize variables. Multicollinearity was addressed using a correlation matrix, VIF, and LASSO regression to remove redundant variables. The remaining variables were analyzed using OLS regression, with spatial autocorrelation tested for model validity.

Then, we focused on the most impactful variables from the regression model. These variables were examined in areas with the highest childhood obesity rates to propose targeted policies for optimizing green space allocation.

Green Space Accessibility Index

First, for each LSOA centroid, Euclidean distances to all green spaces were calculated. Specifically:

$$A_i = \sum_{j \in G(d_{ij} \leq T)} Area(g_j) \tag{2}$$

Where:

- A_i: Total accessible green space area for the i-th LSOA.
- G(d_{ij}≤T): Set of all green spaces g_j within the distance threshold T from the i-th LSOA centroid.
- Area(g_j): Area of the j-th green space.
- d_{ij} Euclidean distance between the i-th LSOA centroid and the j-th green space.
- T: Distance threshold (800 meters (10-minute walking distance) were aggregated(Coombes, Jones and Hillsdon, 2010).).

Then, the total area was summed for each borough as an accessibility indicator.

$$Accessibility\ Index = \sum (Area\ of\ green\ spaces\ within\ 800\ meters\ in\ each\ borough) \tag{3}$$

Green Space Coverage Index

This index measures the proportion of a borough's total land area covered by green spaces, providing an overview of green space availability relative to the overall size of each borough.

$$\text{Coverage Index} = \frac{\text{Total borough greenspace area}}{\text{Total borough land area}} \tag{4}$$

Per Capita Green Space Area Index

This index reflects the amount of green space available per person in each borough, and highlights how equitably green space is distributed among residents within a borough.

$$\text{Per Capita GreenSpace} = \frac{\text{Total greenspace area}}{\text{Total borough Population}} \tag{5}$$

Proportion of Green Space Types Index

This index calculates the share of specific types of green spaces, such as parks or sports facilities, within the total green space area of a borough.

$$\text{Proportion of Type} = \frac{\text{Area of specific greenspace type}}{\text{Total green space area}} \tag{6}$$

4 Result

4.1 Descriptive Statistics

Table 1 Each variable after normalization.

Variable		Mean	Median	Min	Max	Std
Independent	Greenspace coverage box-cox	-3.58	-3.69	-6.23	-1.11	1.36
	Per capita greenspace box-cox	2.91	2.97	1.42	4.29	0.62
	Bowling green	0.00	0.00	0.00	0.01	0.00
	Other sports facility	0.12	0.12	0.01	0.28	0.07
	Play space box-cox	-6.50	-6.52	-9.80	-2.95	1.61
	Playing field box-cox	-1.41	-1.31	-2.01	-0.73	0.30
	Public park or garden	0.70	0.69	0.36	0.94	0.14
	Tennis court box-cox	-10.52	-10.42	-16.05	-4.74	2.49
	IMD average score	20.86	21.79	9.48	32.77	5.66
	Accessible greenspace area box-cox	2.80	2.80	2.80	2.80	0.00
Dependent	Borough obesity rate	16.59	16.78	8.61	22.33	2.93

During the calculation process, Shapiro-Wilk tests are combined with Figure 2, to ensure that each variable follows a normal distribution.

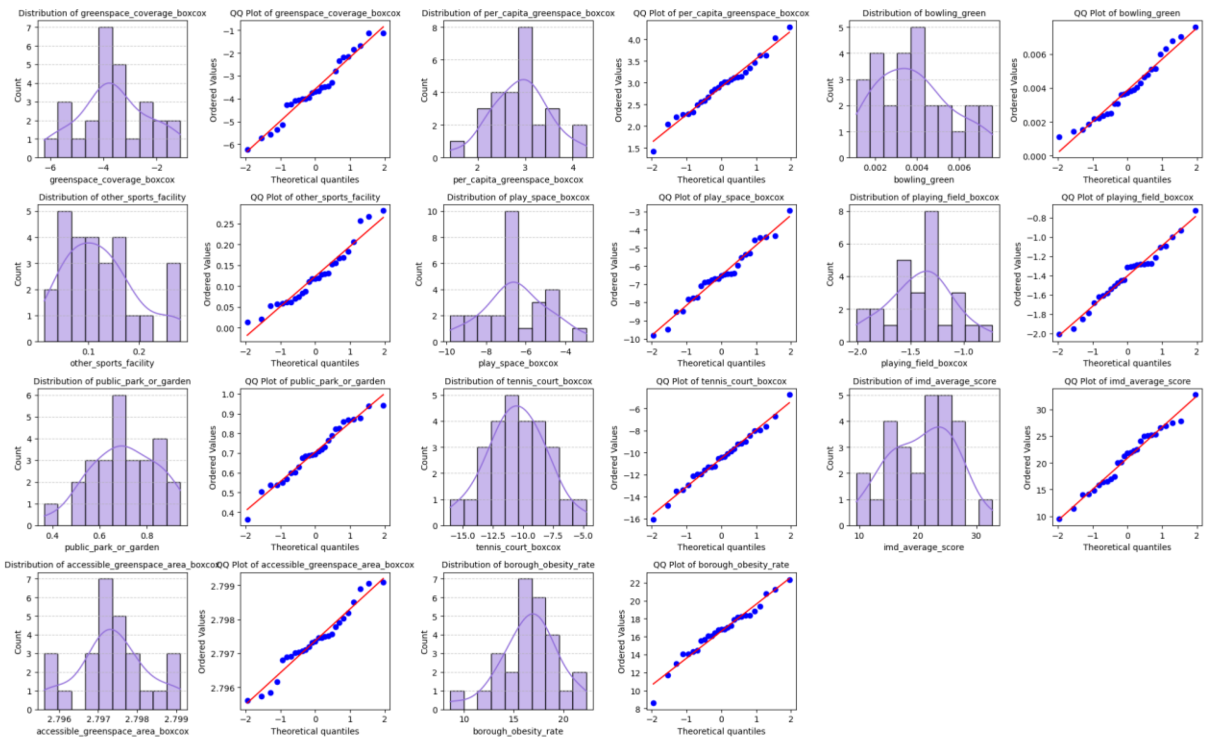

Figure 2 Histograms and QQ plots for each variable after normalization.

4.2 Correlation and Regression

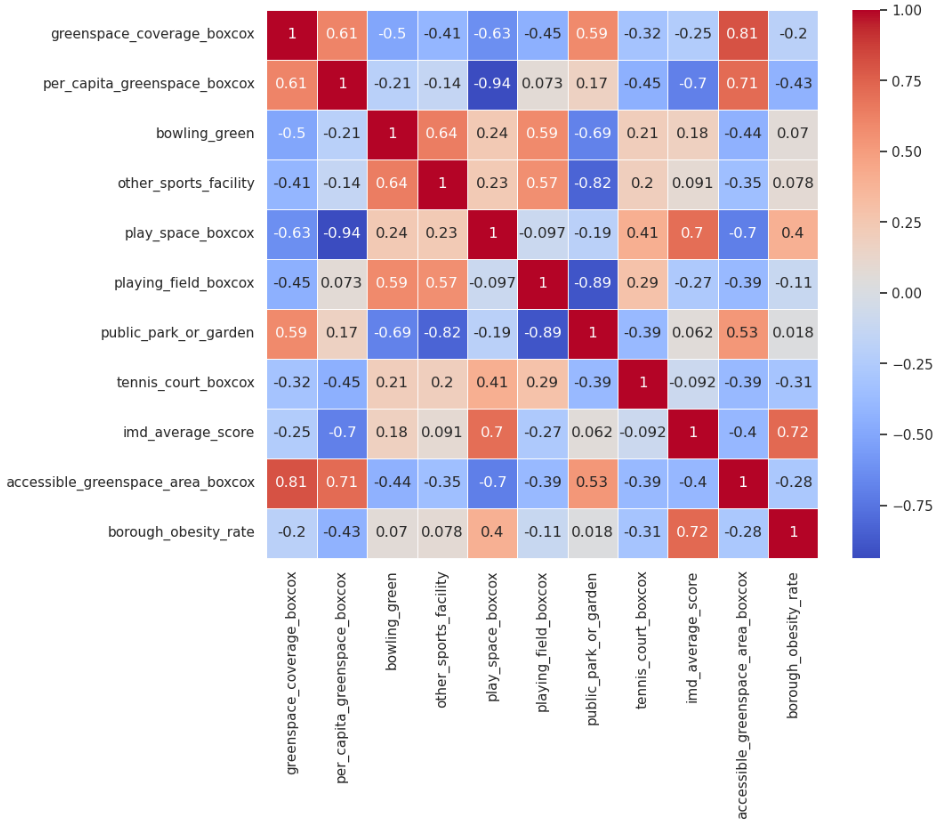

Figure 3 Correlation Matrix Heatmap

Figure 3 shows the presence of multicollinearity among variables. Therefore, using a stepwise elimination method, variables with high VIF values, specifically *public park or garden* and *play space box-cox* were removed. The final VIF values are presented in Table 2.

Subsequently, Lasso regression, in Table 3, was applied to exclude variables with coefficients equal to zero. The remaining variables were then used for the OLS regression analysis.

Table 2 VIF

	Variable	VIF
0	const	4.20 $*10^6$
1	greenspace_coverage_boxcox	4.30
2	per_capita_greenspace_boxcox	8.82
3	bowling_green	2.38
4	other_sports_facility	1.92
5	playing_field_boxcox	2.99
6	tennis_court_boxcox	2.65
7	imd_average_score	4.49
8	accessible_greenspace_area_boxcox	4.35

Table 3 Lasso Regression Coefficient

	Variable	Coefficient
0	greenspace_coverage_boxcox	0.10
1	per_capita_greenspace_boxcox	-1.22
2	bowling_green	-0.00
3	other_sports_facility	-0.00
4	playing_field_boxcox	1.76
5	tennis_court_boxcox	-0.48
6	imd_average_score	0.29
7	accessible_greenspace_area_boxcox	0.00

From Table 4, Tennis Court and IMD Average Score are identified as the most significant factors influencing childhood obesity rates.

Table 4 OLS Regression

(R-squared: 0.642, Adj. R-squared: 0.535)

	coef	std err	t	P>|t|	[0.025	0.975]
const	13.66	8.49	1.61	0.12	-4.00	31.33
greenspace_coverage_boxcox	0.20	0.51	0.40	0.69	-0.86	1.27
per_capita_greenspace_boxcox	-1.58	1.74	-0.91	0.38	-5.20	2.04
playing_field_boxcox	2.14	1.83	1.17	0.26	-1.66	5.94
tennis_court_boxcox	-0.52	0.25	-2.08	0.05	-1.04	0.00
imd_average_score	0.28	0.13	2.07	0.05	0.00	0.55

By calculating Moran's I (-0.087) with a p-value of 0.378, it indicates no significant spatial autocorrelation.

4.3 Local Analysis - Barking and Dagenham

Using the Z-score normalization method, the borough with the highest combined indicator of IMD score and childhood obesity rate (Barking and Dagenham was identified, as shown in Table 5).

$$Combined\ Score = obesity_{norm} + imd_{norm} \tag{7}$$

Then, at the LSOA level, an analysis was conducted incorporating each LSOA's IMD score, population, and tennis court availability, and optimization recommendations were proposed accordingly.

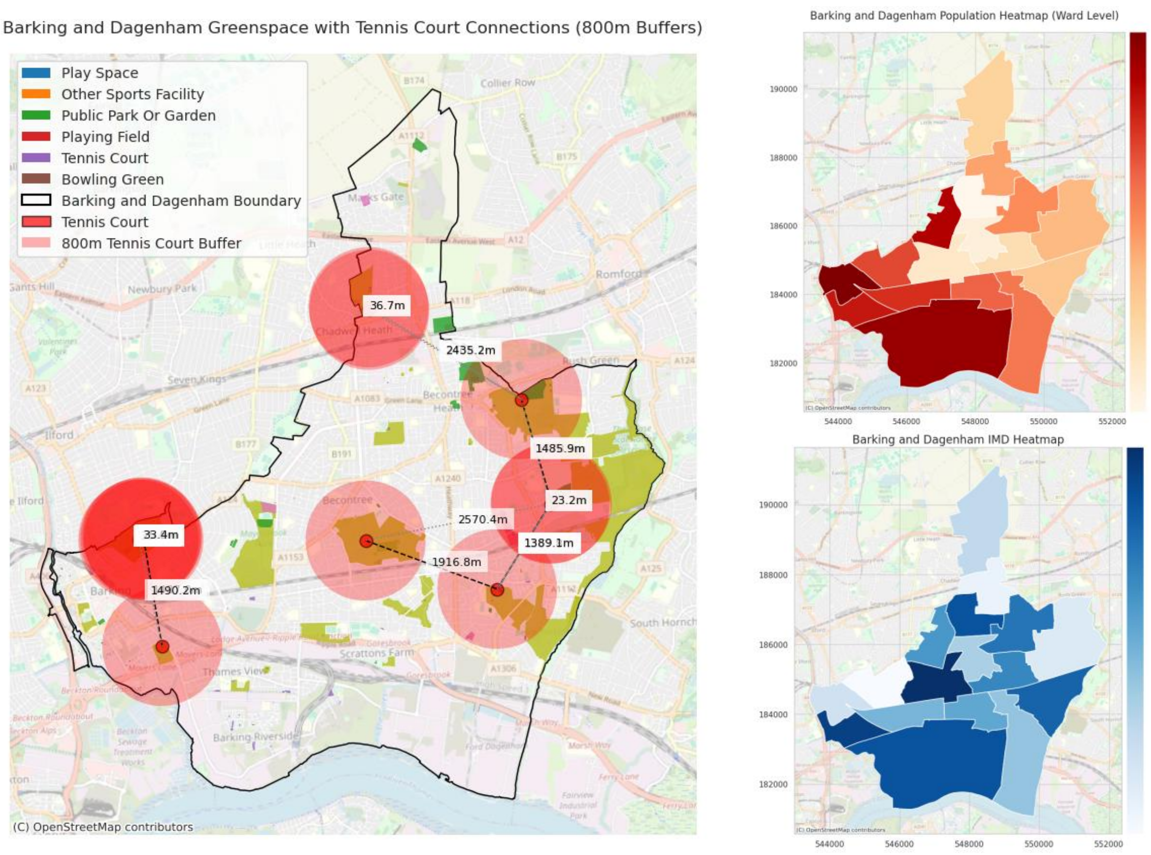

Figure 4 Barking and Dagenham Local Analysis

Table 5 Top 5 Combined score

	Borough	Combined Score
1	Barking and Dagenham	4.06
2	Tower Hamlets	2.81
3	Greenwich	1.98
4	Southwark	1.71
5	Brent	1.35

5 Discussion

5.1 Conclusions

This study examines greenspace accessibility, coverage, type proportions, and per capita usage, alongside childhood obesity and IMD in London.

The OLS regression analysis reveals a strong positive correlation between socioeconomic deprivation and childhood obesity rates, confirming that deprivation remains a key driver of health disparities. Interestingly, the results also suggest that a higher number of tennis courts is associated with reduced childhood obesity rates. This finding highlights the unique role of tennis courts, which, despite being typically located in more affluent areas due to their facility requirements, offer spacious recreational spaces that can accommodate a variety of physical activities. These facilities not only promote tennis but also support broader physical engagement, thereby contributing significantly to reducing obesity rates. The model demonstrates robust explanatory power, with an adjusted R^2 of 53.5%, indicating that the included variables account for a substantial portion of the variation in obesity rates across the study area.

Extending the analysis to Barking and Dagenham, it is observed that green spaces are more concentrated in the central and southern areas, which are typically less deprived and more densely populated. Moreover, areas with high IMD and population indices are often outside the service range of tennis courts. This suggests that resource allocation may not fully address the needs of the most deprived areas.

Suggestion

1. Add tennis courts or sports facilities in highly deprived areas.
2. Expand the service range (buffer zone) of existing greenspace by improving transport and facilities.

5.2 Limitations

1. Establishing a causal relationship with obesity rates is challenging due to the complexity of the factors involved. Therefore, simplistic urban interventions are unlikely to address the fundamental determinants of urban health(Lee and Maheswaran, 2011).

2. The greenspace data is updated in real-time, reflecting the latest information, while other datasets are from 2019. This discrepancy may lead to some deviations from current realities.

3. In the selection of variables, the aesthetic quality and usage frequency of greenspaces could not be evaluated (Cummins and Fagg, 2012).

word count	1593	student number	24143995

References

Björk, J. *et al.* (2008) ‘Recreational values of the natural environment in relation to neighbourhood satisfaction, physical activity, obesity and wellbeing’, *Journal of Epidemiology & Community Health*, 62(4), pp. e2–e2. Available at: https://doi.org/10.1136/jech.2007.062414.

Chen, Y. *et al.* (2024) ‘Examining the non-linear association between the built environment around residence and the risk of obesity among children and adolescents at different school stages in Guangzhou and Shenzhen, China’, *Cities*, 145, p. 104674. Available at: https://doi.org/10.1016/j.cities.2023.104674.

Cohen, D.A. *et al.* (2007) ‘Contribution of Public Parks to Physical Activity’, *American Journal of Public Health*, 97(3), pp. 509–514. Available at: https://doi.org/10.2105/AJPH.2005.072447.

Coombes, E., Jones, A.P. and Hillsdon, M. (2010) ‘The relationship of physical activity and overweight to objectively measured green space accessibility and use’, *Social Science & Medicine*, 70(6), pp. 816–822. Available at: https://doi.org/10.1016/j.socscimed.2009.11.020.

Cummins, S. and Fagg, J. (2012) ‘Does greener mean thinner? Associations between neighbourhood greenspace and weight status among adults in England’, *International Journal of Obesity*, 36(8), pp. 1108–1113. Available at: https://doi.org/10.1038/ijo.2011.195.

Fields in Trust (2018) *Revaluing Parks and Green Spaces*, *Fields in Trust*. Available at: https://fieldsintrust.org/insights/revaluing-parks-and-green-spaces (Accessed: 30 December 2024).

Lachowycz, K. and Jones, A.P. (2011) ‘Greenspace and obesity: a systematic review of the evidence’, *Obesity Reviews*, 12(5), pp. e183–e189. Available at: https://doi.org/10.1111/j.1467-789X.2010.00827.x.

Lachowycz, K. and Jones, A.P. (2013) ‘Towards a better understanding of the relationship between greenspace and health: Development of a theoretical framework’, *Landscape and Urban Planning*, 118, pp. 62–69. Available at: https://doi.org/10.1016/j.landurbplan.2012.10.012.

Lee, A.C.K. and Maheswaran, R. (2011) ‘The health benefits of urban green spaces: a review of the evidence’, *Journal of Public Health*, 33(2), pp. 212–222. Available at: https://doi.org/10.1093/pubmed/fdq068.

Mears, M. *et al.* (2020) ‘Neighbourhood greenspace influences on childhood obesity in Sheffield, UK’, *Pediatric Obesity*, 15(7), p. e12629. Available at: https://doi.org/10.1111/ijpo.12629.

Nicky Hawkins, Kevin Levay (2021) *Changing the Childhood Obesity Conversation to Improve Children’s Health*, *FrameWorks Institute*. Available at: https://www.frameworksinstitute.org/resources/changing-the-childhood-obesity-conversation-to-improve-childrens-health/ (Accessed: 11 January 2025).

Pyky, R. *et al.* (2019) ‘Individual and environmental factors associated with green exercise in urban and suburban areas’, *Health & Place*, 55, pp. 20–28. Available at: https://doi.org/10.1016/j.healthplace.2018.11.001.

World Health Organization (2020) *Noncommunicable diseases: Childhood overweight and obesity*. Available at: https://www.who.int/news-room/questions-and-answers/item/noncommunicable-diseases-childhood-overweight-and-obesity (Accessed: 11 January 2025).

World Health Organization (2024) *Obesity and overweight*. Available at: https://www.who.int/news-room/fact-sheets/detail/obesity-and-overweight (Accessed: 11 January 2025).

Zhou, Y., von Lengerke, T. and Dreier, M. (2021) ‘Comparing different data sources by examining the associations between surrounding greenspace and children’s weight status’, *International Journal of Health Geographics*, 20(1), p. 24. Available at: https://doi.org/10.1186/s12942-021-00278-w.