Wednesday, 3 September 2025

Small Area Estimates of Physical Activity in England: A Guide for Data Users

Using Small Area Estimates to Support Local Decision-Making

ABSTRACT This guide helps stakeholders understand and use small area estimates of physical activity for effective, data-driven planning. It provides an overview of how estimates are generated, explains practical applications, and offers tips on interpreting results responsibly.

Table of contents

1 Quick Start Guide

1. Identify Target Areas: Use small area estimates to find neighbourhoods with low physical activity rates and identify demographic groups that might benefit from specific interventions.
2. Combine Data Sources: Give context to the estimates by overlaying them with other datasets, such as local deprivation indices or health statistics.
3. Interpret with Caution: Remember, small area estimates are predictions based on models. They're best used for spotting general trends rather than exact figures.

2 Background

2.1 Why Small Area Estimates?

National physical activity surveys, like Sport England's Active Lives Adult Survey and Active Lives Children and Young People Survey, provide robust data for large areas like local authorities. However, for neighbourhood-level insights, these surveys lack the necessary detail. Small area estimates bridge this gap by using statistical techniques to predict activity levels in smaller areas down to the level of a few streets. This helps local stakeholders to:

- Pinpoint neighbourhoods with lower activity levels
- Tailor programs and resources to local needs
- Track and support activity goals at a more granular level

3 Practical Applications

3.1 Identifying and Prioritising Neighbourhoods

Small area estimates help you locate areas with lower activity levels, allowing targeted allocation of resources and interventions. For example, if certain neighbourhoods have lower activity rates among older adults, you might focus on providing age-friendly fitness programs or access to low-impact facilities in those areas.

3.2 Informing Program Development

You can use the data to design specific programs that meet the demographic needs of different communities. For example:

- For younger, high-deprivation areas: Develop sports initiatives or community events that build on existing engagement.
- For older, affluent areas: Provide gentle exercise options, such as walking groups or senior fitness classes, that cater to older populations.

3.3 Overlaying with Other Data for Richer Insights

Combine small area estimates with other data sources, such as the Index of Multiple Deprivation or Census data, to gain a fuller understanding of each neighbourhood's characteristics.

3.4 Supporting Funding Applications and Stakeholder Engagement

Use small area estimates to make a data-driven case for funding, showing exactly where and why resources are needed. These estimates also support conversations with stakeholders by clearly illustrating neighbourhood-level needs.

4 Methodology Overview

For this analysis, small area estimates were generated using Multilevel Regression and Poststratification (MRP), a statistical approach that combines survey data with demographic information. Here's a simplified overview:

1. **Model Building (Multilevel Regression)**: Relationships between physical activity and demographic factors (e.g., age, disability status, socio-economic group) are modelled to predict activity levels in small areas.
2. **Poststratification**: Adjusts predictions to reflect each area's unique population structure. This ensures estimates are tailored to each small area's actual demographics.

4.1 Why MRP?

MRP is particularly effective in producing stable estimates for smaller areas with limited data. By "borrowing strength" from similar population groups and adjusting for each area's demographics, MRP helps generate a more reliable picture of physical activity trends across neighbourhoods.

4.1.1 Missing CYP School-Level Variables

The most significant limitations relate to the children and young people (CYP) small area estimates. A key challenge is that many predictors of physical activity for CYP operate at the school level. However, privacy constraints prevent linking data to individual school locations, limiting the model's ability to capture geographic-specific differences within small areas (LSOAs and MSOAs) for children and young people.

With limited access to school-specific data, the model must rely on broader demographic and socio-economic variables, which do not capture the environmental factors that seem to play stronger roles in children's physical activity. This reduces the model's precision and results in less geographic variation in the estimates, limiting their utility for highly localised decision-making by end users.

5 Key Predictors of Physical Activity: The main drivers of physical activity in small areas

The modelling work identified a series of key drivers—or predictors—of physical activity for individuals across different demographics and geographies. Understanding these predictors helps to explain the variations in physical activity levels observed across local areas.

5.1 Adult Physical Activity Predictors

For adults, the variation in physical activity levels is predominantly driven by individual characteristics rather than geographic location. This means that regardless of where adults live, factors such as disability, age, and socio-economic status play a far greater role in determining physical activity levels. The main predictors are:

- **Disability Status**: Disability status is a strong predictor, though the impact varies by disability type. For instance, certain disabilities, such as visual, learning and behavioural disabilities have little effect on physical activity levels, while mobility disabilities and disabilities linked to chronic pain and chronic illnesses are strongly associated with lower levels of physical activity.
- **Age**: Age is closely linked to physical activity levels, with younger adults generally more likely to engage in physical activity than older adults. Age also interacts strongly with disability status: able-bodied individuals tend to maintain similar activity levels from age 16 to 55, after which activity begins to decline. By contrast, disabled individuals are less likely to be physically active with every year older they are, starting from age 18.
- **National Statistics Socio-economic Classification (NS-SEC)**: Socio-economic status is a significant driver of physical activity, with higher socio-economic groups generally having greater access to resources and environments that promote activity. Notably, students within the NS-SEC framework show particularly high levels of physical activity.

Since these individual characteristics explain most of the variation, geographic area alone is not a key predictor of physical activity in adults.

5.2 Predictors for Children and Young People

For children and young people, the key predictors of physical activity operate differently.

Individual-level factors like disability status and age play a much less significant role in children's activity levels than they do in adults. Family affluence - a rough analogue to NS SEC - does have an impact on children and young people's likelihood of being physically active, but its influence is not as significant as in adults.

Instead, the key drivers seem to be at a level between the individual and geographic level of Local Authority. There is initial evidence that these key drivers are at a school level. However, due to data privacy restrictions, there is a lack of data available at this level so further work is required to better understand how schools influence physical activity.

6 Data Outputs

The modelling developed for this project has produced several sets of estimates that offer a comprehensive view of physical activity patterns across different population groups and geographical levels.

6.1 Key Measures Reported

For adults, three measures have been reported:

- **Physically Active**: Proportion of adults engaging in 150 minutes or more of at least moderate physical activity per week.
- **Physically Inactive**: Proportion of adults engaging in fewer than 30 minutes of at least moderate physical activity per week.

- **Participation in Sport**: Proportion of adults participating in sport and/or physical activity at least twice in the last 28 days.

For children and young people, two measures have been reported:

- **Physically Active**: Proportion of children and young people averaging 60 minutes or more of moderate physical activity per day.
- **Physically Inactive**: Proportion of children and young people averaging fewer than 30 minutes of moderate physical activity per day.

6.2 Geographical Levels

Estimates for both adults and children are provided at three geographical levels:

- **Lower Super Output Area (LSOA)**: Small, neighbourhood-level areas typically consisting of 1,000 to 3,000 people. Estimates at this level help pinpoint localised trends and specific areas of need within a neighbourhood, enabling highly targeted interventions.
- **Middle Layer Super Output Area (MSOA)**: Mid-sized areas with an average population of around 8,200 people. MSOA-level estimates provide a balanced view that is detailed enough for practical planning but broad enough to identify area-wide patterns.
- **Local Authority**: These larger administrative areas, currently numbering 296 across England, support broader, strategic decisions and help align local efforts with wider regional policies.

Providing data at multiple geographic levels allows users to zoom in and out on physical activity trends. For instance, while LSOA-level data can reveal pockets of inactivity within a neighbourhood, Local Authority data can highlight overarching needs and successes across larger districts, supporting both micro- and macro-level decision-making.

6.3 Demographic and Socio-Economic Breakdown

Estimates are also provided for detailed demographic and socio-economic breakdowns across all three geographic levels (LSOA, MSOA, and Local Authority), offering nuanced insights into activity patterns for specific population groups within each area.

For adults, data is available at each geographic level for groups defined by:

- Gender
- Age
- Ethnicity
- Disability status
- Socio-economic group

For children and young people, estimates are provided across the same geographic levels and broken down by:

- Gender
- School stage

- Ethnicity
- Disability status
- Family affluence

These demographic breakdowns allow users to examine physical activity levels for specific groups within neighbourhoods, mid-sized areas, and local authorities, supporting targeted analysis and planning at different scales.

6.4 Accessing the Data and Mapping Tool

All data sets produced through the small area estimates project are publicly accessible on Sport England's website. These data sets are updated regularly, with the latest adult estimates based on the 2023/24 Active Lives survey data.

To make this data more accessible, Sport England has also developed the Active Lives Small Area Estimates Tool. This interactive mapping tool enables users to visually explore and compare estimates across different geographical areas and demographic groups. By selecting specific indicators and applying demographic filters, users can engage directly with the data in a dynamic, map-based format. A more detailed guide on using this tool effectively is provided in the following section.

7 How to Interpret and Use the Estimates

When working with these estimates, it's important to distinguish between interpreting what the data represents and effectively applying it for decision-making. Below are guidelines for both aspects to help users make the most of these estimates.

7.1 Interpreting the Estimates

These estimates are modeled predictions, not direct survey measurements, especially at smaller geographic levels. MRP provides a 'best estimate' based on available data, but small population sizes or unique local characteristics may increase the uncertainty in some areas.

7.1.1 Geographic Comparisons

Comparing different areas can yield valuable insights, but these estimates are best suited for observing general trends rather than precise counts, particularly at the neighbourhood level. Local characteristics, such as community-specific health programmes or facilities, may influence physical activity in ways that are not captured by the model's demographic and socioeconomic predictors. Users should focus on broader patterns rather than exact comparisons of specific values.

7.1.2 Temporal Comparisons

It's important to exercise caution when interpreting changes over time with these estimates. Because they are model-based and incorporate data from multiple sources beyond just local conditions, the estimates may not fully reflect year-to-year changes specific to an individual area. Instead, these estimates are best used for identifying broader trends across areas or demographic groups rather than tracking precise local changes over time.

7.1.3 Consider Uncertainty

Each estimate includes a degree of uncertainty, which is typically higher for smaller or unique population groups. It's recommended to view the estimates as indicative of trends and patterns rather than exact figures. For decision-making purposes, users should consider the estimates as a starting point for further investigation, especially where smaller sample sizes or uncommon demographics are involved.

7.2 Using the Estimates

The estimates offer valuable insights for guiding local policy and programme planning by identifying neighbourhoods or demographic groups with lower physical activity levels.

7.2.1 Informing Decision-Making

The estimates are particularly useful for guiding policy and planning by highlighting neighbourhoods or demographic groups with lower physical activity levels. Comparing areas with similar demographic profiles can help identify regions where additional support, resources, or targeted programmes might increase community activity levels. For example, stakeholders might use the data to allocate funding for physical activity programmes in underserved areas or to develop targeted outreach initiatives.

7.2.2 Combining Estimates with Other Data

Users can enhance the interpretative power of these estimates by combining them with other contextual data, such as census information, deprivation indices, and local health statistics. This combined approach allows for a more comprehensive understanding of physical activity patterns across England's communities.

7.3 Practical Applications of the Mapping Tool

One of the most effective ways to interact with the estimates is through the **Active Lives Small Area Estimates Tool**. This tool allows users to explore the data visually and compare estimates across different areas and demographics, providing a dynamic way to identify trends and potential areas of need. Here's how to leverage the tool effectively:

1. **Select Indicators and Demographics**: Users can filter by specific indicators (e.g., physical activity levels, inactivity) and demographic factors such as age, gender, or socio-economic classification to focus on specific population segments.
2. **Overlay Contextual Data**: The tool enables users to overlay additional datasets, such as census data, deprivation levels, and other public health indicators, to provide richer insights and support more informed planning decisions.
3. **Compare Across Areas**: Users can compare activity levels across multiple areas on the map, identifying neighbourhoods with lower levels of physical activity compared to surrounding areas or similar demographic profiles. This feature is valuable for local authorities aiming to prioritize resources or interventions based on relative need.
4. **Export and Share Visuals**: The mapping tool allows users to export visuals for reports and presentations, making it easy to communicate findings and engage

stakeholders in discussions about community health needs and resource allocation.

These estimates, combined with other contextual data available through the mapping tool, can guide users in developing a thorough understanding of physical activity levels across England's neighbourhoods, supporting targeted, data-driven interventions for improved community health.

8 Case Study: Using Small Area Estimates for GreaterSport in Greater Manchester

GreaterSport, an active partnership in Greater Manchester, aims to improve physical activity levels across diverse communities. By leveraging small area estimates, GreaterSport can better understand where are the areas and demographics with low physical activity and implement targeted interventions. This case study illustrates how GreaterSport can use these data-driven insights for strategic planning.

8.1 Using the Active Lives Mapping Tool to Identify Target Areas

To identify neighbourhoods with low physical activity levels, GreaterSport can use the Active Lives Small Area Estimates Tool. Filtering for active adults at the LSOA level (the most precise level available), the tool shows all LSOAs within GreaterSport's area, colour-coded for their activity levels (see @fig-greatersport-mapping-tool). From this GreaterSport can pinpoint areas requiring additional resources, facilities, or programmes tailored to community needs.

8.2 Exploring Demographic Nuances: A Comparison of Trafford 011C and Stockport 042B

Areas with similar physical activity rates can have very different demographics, impacting approaches to interventions . For example, Trafford 011C and Stockport 042B both report have activity rates of 61.5%, but their demographic profiles contrast significantly. Trafford 011C is in the lowest deprivation decile and has a younger population, while Stockport 042B is in the most affluent decile with an older demographic. @fig-stockport-trafford-age shows the difference in age distribution.

This highlights the importance of placing the physical activity estimates in context. To do so, there are two main approaches. Firstly, you can change filters, looking into the demographic breakdowns within areas. Secondly, you can combine the physical activity estimates with other datasets. For example, the mapping tool supports overlaying data from the census and deprivation datasets. Or you can look at datasets like Nomis' 2021 Census Local Area Report tool, as shown in @fig-nomis-age-table.

Understanding these distinctions helps GreaterSport tailor its efforts. For instance, in Trafford 011C, initiatives might focus on engaging younger residents with accessible facilities, while in Stockport 042B, the emphasis could be on age-friendly, low-impact activities suited to older adults.

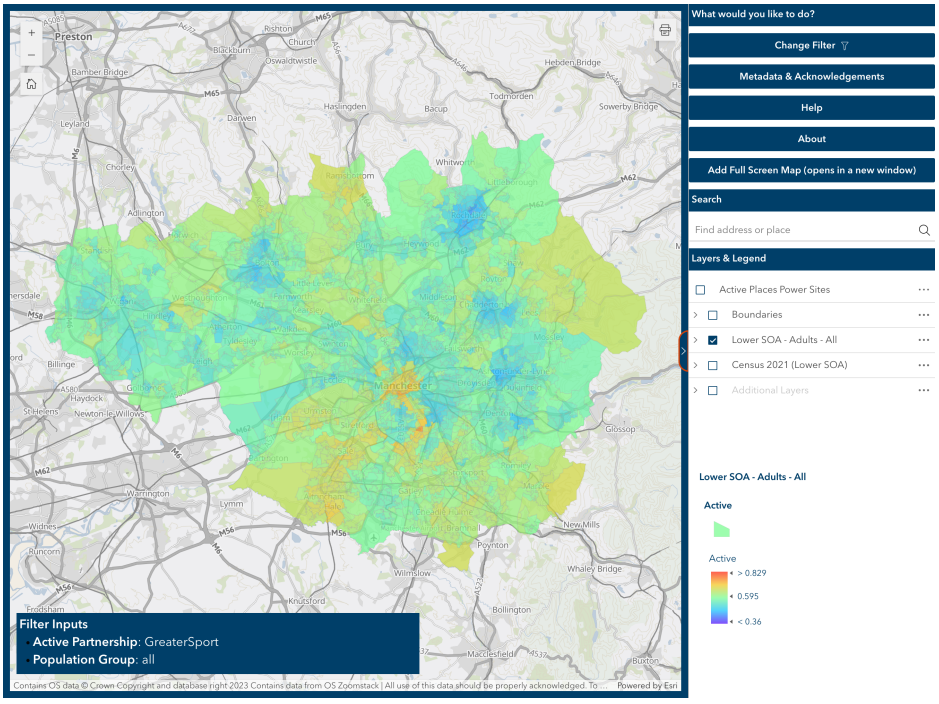

Figure 1: GreaterSport in the Active Lives Mapping Tool.

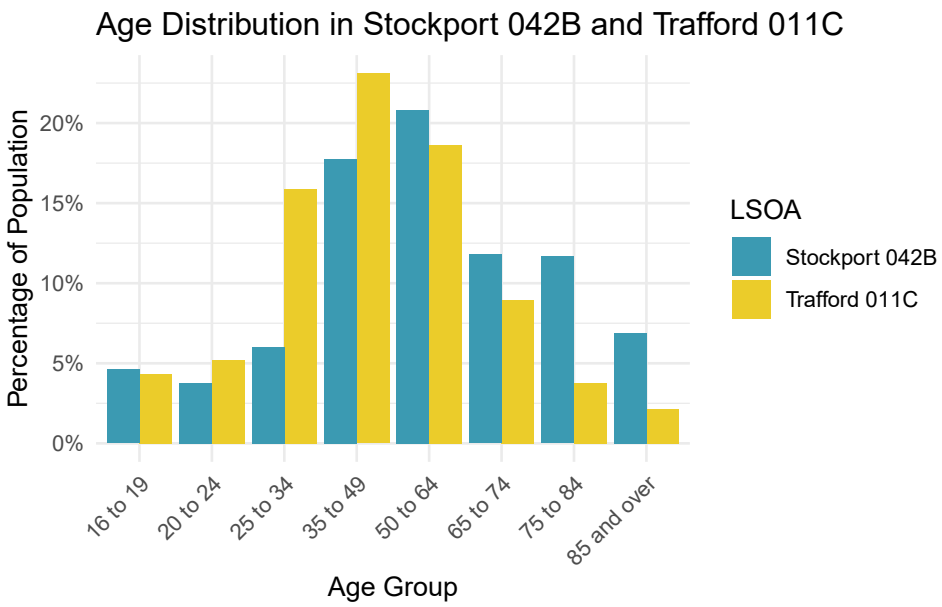

Figure 2: Age Distribution in Stockport 042B and Trafford 011C

Age

Persons

	E01006144 : Trafford 011C 2021 Lower-layer SOA		E01005925 : Stockport 042B 2021 Lower-layer SOA	
	count	%	count	%
All usual residents	1,386	100.0	1,432	100.0
Aged 4 years and under	57	4.1	51	3.6
Aged 5 to 9 years	86	6.2	81	5.7
Aged 10 to 15 years	107	7.7	107	7.5
Aged 16 to 19 years	60	4.3	66	4.6
Aged 20 to 24 years	72	5.2	54	3.8
Aged 25 to 34 years	220	15.9	86	6.0
Aged 35 to 49 years	320	23.1	254	17.7
Aged 50 to 64 years	258	18.6	298	20.8
Aged 65 to 74 years	124	8.9	169	11.8
Aged 75 to 84 years	52	3.8	167	11.7
Aged 85 years and over	30	2.2	99	6.9

Show table notes...

Source: ONS - 2021 Census (TS007B)

Figure 3: A table from Nomis' 2021 Census Local Area Report comparing the age distributions in Trafford 011C and Stockport 042B.

9 Frequently Asked Questions (FAQ)

To conclude this report, below are a series of responses to questions that are commonly raised about the small area estimates and how they are produced.

9.1 How are estimates produced for small populations in areas?

The analysis gives estimates for all possible cross-sections of the variables included in the model. For every possible cross-section, the model gives an estimate of the expected physical activity of someone from this group. This includes large populations - such as white, 35-year-old, university-educated men in a specific area of London - and very small populations - such as Chinese, 99-year-old, female students in a rural, mostly white area. These small populations will have larger uncertainties but are still given in the estimates. Their effect on any of the estimates is proportional to the size of the strata (i.e. the number of people with that description).

9.2 Why are estimates not available for some populations in specific areas?

The poststratification frame - the data on how many people of each cross-section there are - is built using Iterative Proportional Fitting. This technique brings together various datasets from the Office of National Statistics to create a more detailed frame than they release whilst aligning with all their released data. For some cross sections, there are

no people fitting the cross-section's description. While the model can give estimates for these fictional people, we omit them in the released data.

9.3 Why aren't the published Active Lives results completely aligned with the small area estimates?

Sport England publishes Local Authority estimates of physical activity from Active Lives. These estimates are based on survey weighting. The small area estimates are calculated using a different technique - multilevel regression and poststratification - which is especially proficient at providing estimates for areas with limited data, such as small area estimates. The two different calculation methods lead to small variations in the estimates at Local Authority level.

10 Glossary

Small Area Estimates Statistical estimates calculated for small geographic areas, often at neighborhood or sub-regional levels. These estimates help policymakers understand variations in data across small, specific regions.

Multilevel Regression and Poststratification (MRP) A statistical technique that combines survey data with demographic information to produce reliable estimates for smaller geographic areas or specific population groups. MRP helps to account for sparse data in small areas, improving the accuracy of predictions.

Multilevel Regression A type of statistical modeling that accounts for data structured at multiple levels (e.g., individuals within neighborhoods). In this report, multilevel regression is used to model relationships between physical activity and demographic factors, allowing for variation across regions and population groups.

Poststratification A process used in modeling to adjust predictions so they match the demographic structure of the target population. This step ensures that model estimates reflect the distribution of the population accurately across all relevant demographic groups.

Poststratification Frame: A dataset that represents the distribution of a population across key demographic and geographic categories.

Iterative Proportional Fitting (IPF) A method used within poststratification to align model predictions with known population data by adjusting proportions within demographic groups. IPF ensures that estimates reflect the actual population structure more accurately.

Lower Super Output Area (LSOA) A small geographic unit in England, typically made up of between 1,000 and 3,000 residents. LSOAs allow for detailed, neighborhood-level data analysis.

Middle Layer Super Output Area (MSOA) A mid-sized geographic unit in England with an average population of around 8,200 people. MSOAs are used for

community-level analysis, sitting between LSOAs and local authority districts in geographic scale.

Socio-economic Classification (NS-SEC) The National Statistics Socio-economic Classification is a system used to categorize individuals based on occupational and social status, commonly used to analyze socio-economic influences on behaviors such as physical activity.