noo
National Obesity Observatory

NHS

National Child Measurement Programme
Guidance for small area analysis

July 2011

Delivered by NOO on behalf of the Public Health Observatories in England

Contents

1. Key points 3
2. Introduction 4
3. Background 4
4. The geographic areas available for analysis of NCMP data 5
5. Statistical uncertainty and NCMP data 6
6. Methods for increasing the robustness of NCMP analysis for small populations 10
   - A. Combine NCMP data from different years of measurement 10
   - B. Combine small geographic areas into larger population groups and neighbourhoods 12
   - C. Combine populations with similar characteristics 14
   - D. Combine data from Reception and Year 6 16
7. Further information and advice 17
8. Reader information 18

1. Key points

- The National Child Measurement Programme (NCMP) dataset provides height and weight measurements for approximately one million children during each year of measurement. This large sample provides detailed information about patterns of child body mass index (BMI) at both national and sub-national level, and for specific population groups.

- The NCMP dataset contains a very large number of records at national level, but at local level, or in certain subgroups, the numbers may be very much smaller. It is important to ensure the number of children measured used in analysis remains large enough to provide robust data.

- In general, the greater the number of measurements used within any analysis, the more reliable the resulting statistics will be. Analysis based on small samples may be affected by small number variation, and therefore may not provide a reliable estimate of the true value in the underlying population.

- There are techniques available to increase the robustness of NCMP analysis for small populations, including:

    - combining two or more years of NCMP data;
    - combining small geographic areas to create larger neighbourhood areas within a local authority;
    - combining populations with similar characteristics, such as ethnicity or socioeconomic status;
    - combining data from Reception and Year 6.

2. Introduction

This paper provides advice for users of the National Child Measurement Programme (NCMP) dataset who wish to undertake analysis at small area level, such as local neighbourhoods or communities. This guidance supplements the general guidance for NCMP analysis which is circulated with the dataset and available on the National Obesity Observatory (NOO) website.[a]

For more detailed guidance on using small area data in public health intelligence, not specific to the NCMP, the Public Health Observatories in England have produced a Technical Briefing applicable to all sources of public health data.[b]

NCMP analysis is routinely undertaken at national, regional, Primary Care Trust (PCT) and Local Authority (LA) levels. NOO has also recently published NCMP data at Middle Super Output Area (MSOA) level,[c] but prior to this, analysis with national coverage had not been produced below local authority level.

When producing analysis for small population groups, it is important to consider the effect small numbers might have on any resulting analysis. The greater the sample used within an analysis, the more reliable the results. Analysis based on a small sample may be affected by small number variation and may therefore not provide a reliable estimate of the true value in the underlying population.

Although the NCMP dataset contains a very large sample of measurements, when split into smaller geographic or population groups (such as electoral wards or schools, ethnic groups or deprivation deciles), it is important to ensure the number of children measured remains large enough to provide robust data.

This paper explains why caution must be exercised when using NCMP data to provide information about small populations or local areas. It discusses the reliability of estimates of obesity prevalence at different levels of aggregation, and provides advice on approaches to ensure that any analysis is as robust as possible. Alternative ways to identify areas of high obesity prevalence within LA and PCT boundaries are also suggested.

3. Background

The NCMP dataset provides height and weight measurements for approximately one million children during each year of measurement. It includes the majority of the child population across two age groups, Reception (age 4–5 years) and Year 6 (age 10–11 years). This large sample size means the dataset can provide detailed information about patterns of child BMI at both national and sub-national level, and for specific population groups.

There has always been a great deal of interest in using NCMP data to provide an indication of child weight status for areas or populations smaller than local authorities. Many child weight management interventions are commissioned and delivered at local level, and local health organisations, local authorities and others are keen to

[a] www.noo.org.uk/uploads/doc/vid_10612_0910_NCMP_PHO_Analysis_Guidance.pdf
[b] www.apho.org.uk/resource/item.aspx?RID=74894
[c] www.noo.org.uk/visualisation/eatlas

understand as much as possible about the patterns of obesity prevalence within their local area.

The demand for small area analysis of NCMP data was highlighted in a recent survey conducted on behalf of the Department of Health (DH) by Ipsos MORI. In addition, both DH and NOO have received a number of *ad hoc* queries regarding such analysis from PCTs, LAs and other organisations.

There is a demand for small area and population level NCMP data in order to target interventions at the most at risk communities within a local area, or to monitor change over time. In addition, these data are sometimes required to evaluate whether schemes to encourage healthy eating or physical activity that have been implemented in certain areas or schools have led to a change in obesity prevalence.

In many cases NCMP data may be able to fulfil these requirements. However it is important to ensure that any analysis produced is suitable for its intended purpose and that the data are interpreted appropriately.

4. The geographic areas available for analysis of NCMP data

The NCMP dataset contains a very large number of records at national level, but at local level, or in certain subgroups, the numbers may be very much smaller.

On average, PCTs each measure around 3,000 children per age group per year for the NCMP. This equates to around 1,500 children measured per lower tier local authority. Parliamentary constituencies each provide an average sample of nearly 1,000 children per age group. Statistics based on such large numbers of child measurements are likely to provide relatively robust estimates of measures such as obesity prevalence.

However, for smaller geographic areas or population units, the number of children measured is substantially lower and may be inadequate to provide robust estimates. On average at MSOA level there are only 75 children measured per year in each NCMP age group, with only around 15 children of each age group measured per Lower Super Output Area (LSOA) per year. Table 1 shows the average number of children measured for the NCMP at the various levels of aggregation used in England.

Table 1: Levels of aggregation of NCMP data, and number of children measured (per age group)

Common geographies:	PCT	Lower tier LA†	Constituency	MSOA	Ward	School	LSOA
Number of areas	152*	326	533	6,781	7,618[α]	12,800[+]	32,300
Average population	300,000	140,000	100,000	5000+	6,800	N/A[β]	1,500
Average number of children measured for NCMP (2009/10)	3,300	1,500	935	75	65	35	15

*Two Hertfordshire PCTs merged in March 2010, although they submitted NCMP data separately in 2009/10.

[α] Statistical wards, as opposed to Census wards.

[+] Eligible schools only (i.e. not including independent and special schools).

[β] School population varies substantially, depending on whether the school is infant, junior or both.

† Lower tier LAs include district and unitary authorities

As Table 1 shows, there is a substantial difference in size between PCTs, LAs and constituencies and the smaller geographies from MSOAs and below. There have been suggestions that the Office for National Statistics (ONS) create a third tier of output area groupings, an 'Upper Super Output Area', but this does not currently exist.[d]

In addition to the issues around small numbers of children measured, for geographic levels smaller than parliamentary constituency there will be some areas that have no children measured in the NCMP. Therefore it is unlikely that any indicator using NCMP data could be created with 100% geographic coverage.

It is also important to remember that even NCMP-based statistics for geographic areas such as PCTs or LAs, if split by ethnic group or other sociodemographic factors, may also be based on only a small number of child measurements. The guidance in this document also applies to these forms of small sample size analysis.

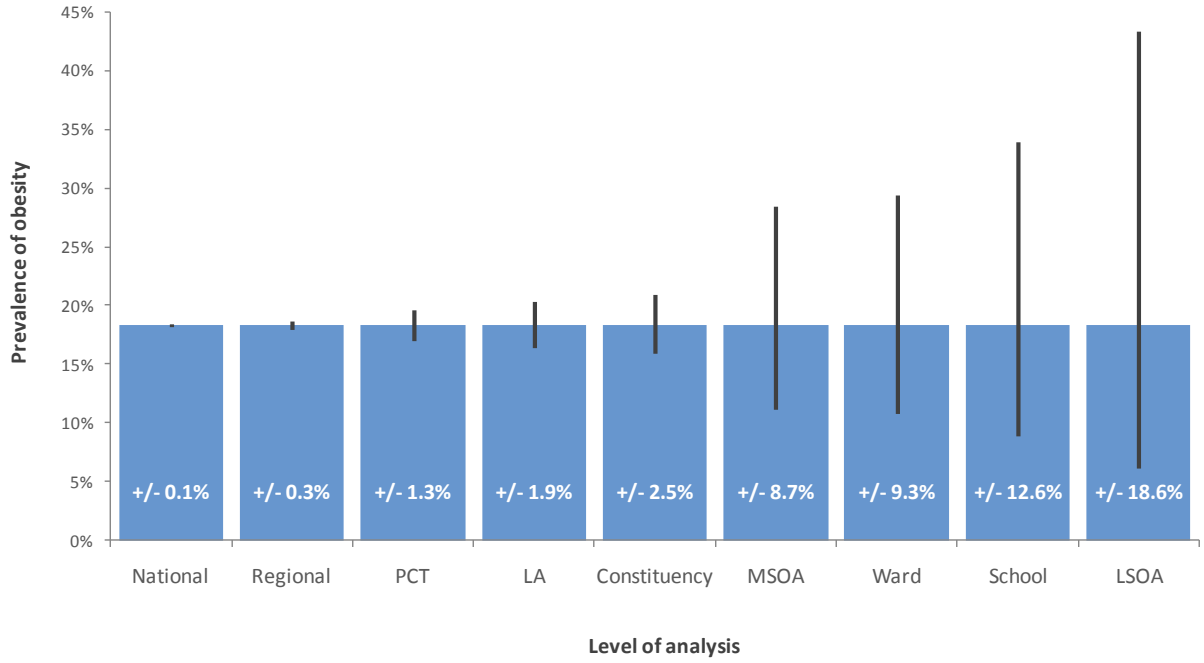
5. Statistical uncertainty and NCMP data

In general, the greater the number of measurements used within any analysis, the more reliable the resulting statistics. Analysis based on small samples may be affected by small number variation, and therefore may not provide a reliable estimate of the true value in the underlying population. This level of uncertainty is illustrated by the confidence limits around such statistics, which increase as the number of children measured decreases.

Figure 1 shows the approximate size of the confidence limits around obesity prevalence figures for children in Year 6 at different levels of geography using a single year of NCMP data. These figures are based on the average number of children measured by the NCMP at these geographic levels.

[d] http://bit.ly/pvPf9t

Figure 1: Average confidence limits around estimates of obesity prevalence at different levels of analysis



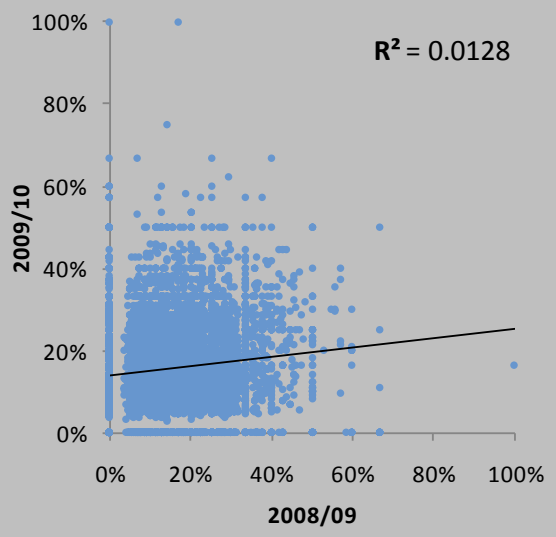
The estimates presented in the chart above are based on the national prevalence of obesity among children in Year 6 and on one year of NCMP measurements

Confidence limits around obesity prevalence estimates at national, regional, PCT, LA and even parliamentary constituency level tend to be of a reasonable size (less than +/- 3%). However for the available English geographic areas below this level (such as electoral wards, MSOAs or LSOAs) confidence limits around obesity prevalence figures increase dramatically to around +/- 10%.

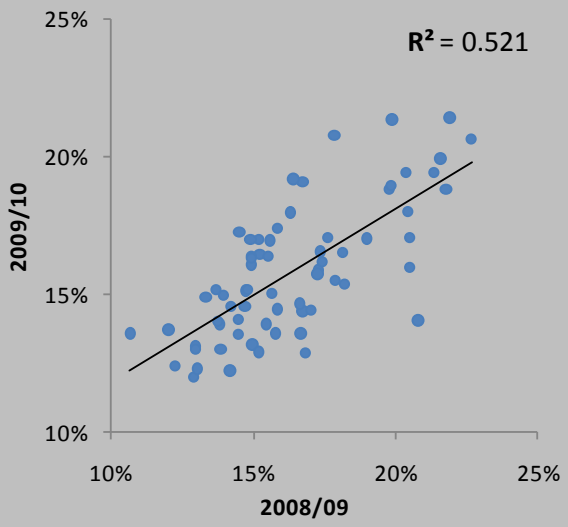
Such a degree of uncertainty is likely to pose a number of difficulties when using such statistics. For example, the data are likely to show substantial year to year variation and unlikely to show a strong association with indicators describing the determinants of obesity (see Case Study 1). Such wide confidence limits will also make it difficult to determine whether any observed differences in prevalence between areas or over time are meaningful, or whether they are likely to have arisen by chance.

Case Study 1: Comparing data at LA and LSOA level within the South East region

Year to year variation in obesity prevalence – LSOAs

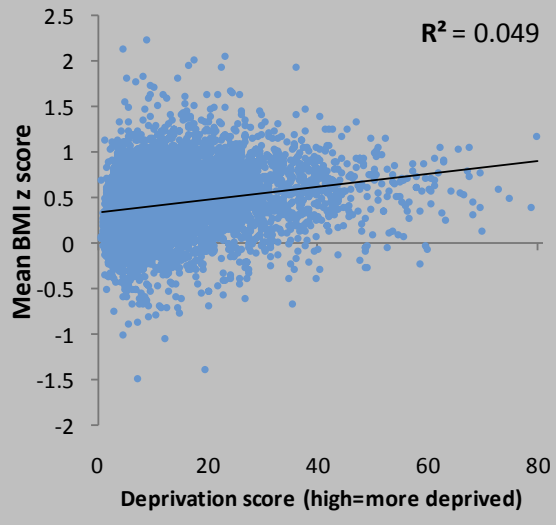


Year to year variation in obesity prevalence – LAs

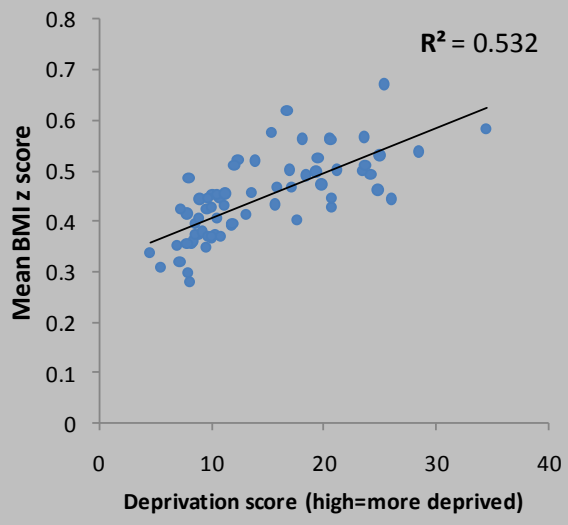


There is substantial year on year variation in obesity prevalence figures produced at LSOA level for children in Year 6, as shown by the low R squared* value. Figures for LAs remain broadly similar from one year to the next. In practice this means that if LSOA data were used to target interventions or resources, the areas targeted are likely to change substantially with each new NCMP dataset.

Relationship between mean BMI z score and deprivation (IMD 2010) – LSOAs



Relationship between mean BMI z score and deprivation (IMD 2010) – LAs

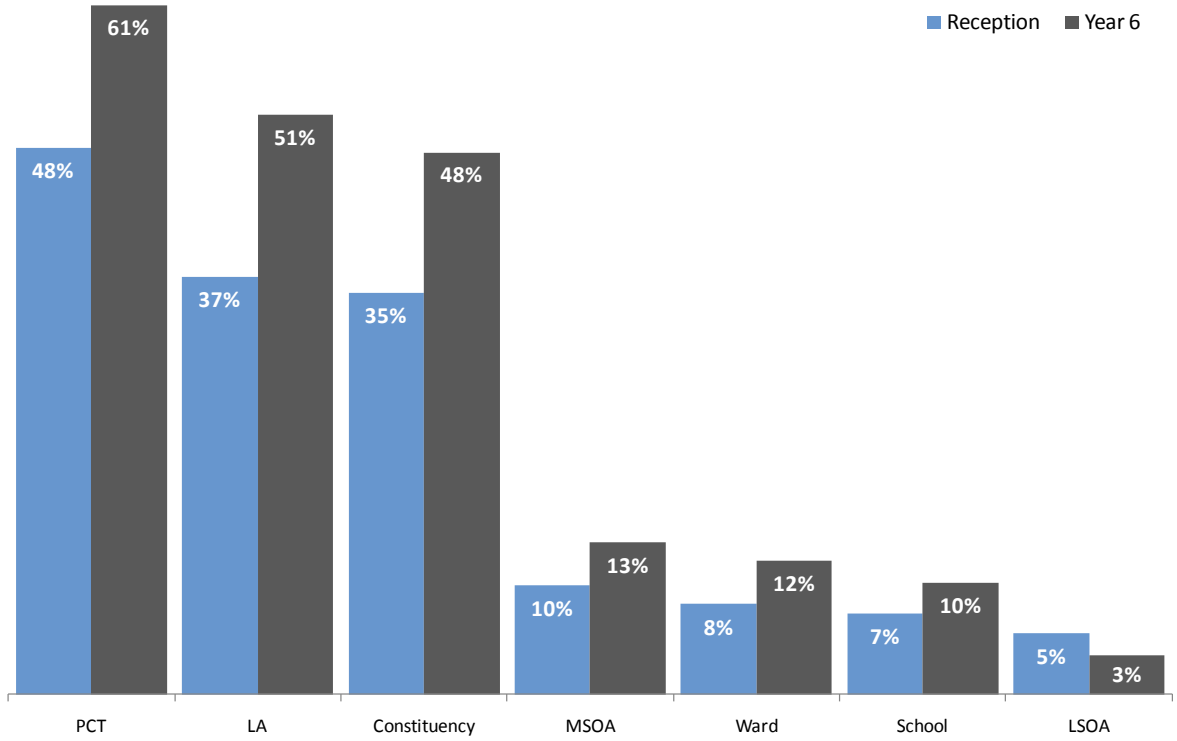


At LSOA level there is only a weak relationship between mean BMI z score (a summary measure of BMI in a child population) and deprivation, as shown by the low R squared value. At LA level this relationship is much stronger. However if analysed in a different way, for example grouping by deciles of deprivation (which are still assigned at LSOA level), it can be shown that children living in more deprived LSOAs do tend to have significantly higher BMI z scores than those living in less deprived LSOAs.

**The R-squared value (the coefficient of determination) shows what proportion of the variation in one variable can be explained by another. The R-squared values range from 0 to 1, with values closer to one implying a stronger relationship between the variables.*

Figure 2 shows the proportion of areas that have a statistically significantly different obesity prevalence from the England average. As the number of children measured decreases, so does the proportion of areas or population groups where we can detect a statistically significant difference in prevalence from the national average.

Figure 2: Proportion of areas with obesity prevalence that is significantly different (higher or lower) to the England average, NCMP 2008/09



For PCTs, LAs, and constituencies, at least 35% of areas show obesity prevalence that is significantly higher or lower than the England average. As some conclusions can be drawn about the obesity prevalence in at least one third of the populations examined, such statistics can therefore provide useful information about the pattern of obesity prevalence across the country.

However, for smaller geographies it is only possible to detect significant differences in obesity prevalence for 13% of areas at most. This means that analysis of obesity prevalence using one year of NCMP data will not be able to detect any difference from the national average for at least seven out of every eight areas at MSOA level or below. As the vast majority of the country is likely to show no significant difference from the national average, such statistics will provide very little useful information for many potential users.

In addition, as a 95% significance level has been used for this analysis, around 5% of areas would be expected show a significant difference by chance. This will become increasingly more of an issue for geographic levels with smaller populations as they tend to show fewer significant differences. For example, the 5% of LSOAs shown (in Figure 2) to have a significantly different prevalence of obesity may well represent nothing more than the level of statistical significance (95%, in this instance) used when checking for differences.

6. Methods for increasing the robustness of NCMP analysis for small populations

In order to create more robust NCMP analysis for small populations, the main aim should be to try and maximise the number of children included in the analysis. This can be achieved in a number of ways:

A. Combine NCMP data from different years of measurement

Most local areas will now have access to multiple years of NCMP data which can be used in analysis. It will often be possible to combine data from these years of measurement in order to increase the number of child measurements available for each area.

This approach is recommended as it is usually relatively easy to perform, especially since the most recent NCMP datasets circulated by the NHS Information Centre cover all years from 2006/07 to 2009/10. Many other health indicators (for example mortality rates) are frequently produced by combining data from a number of years, so this approach should be familiar to most users.

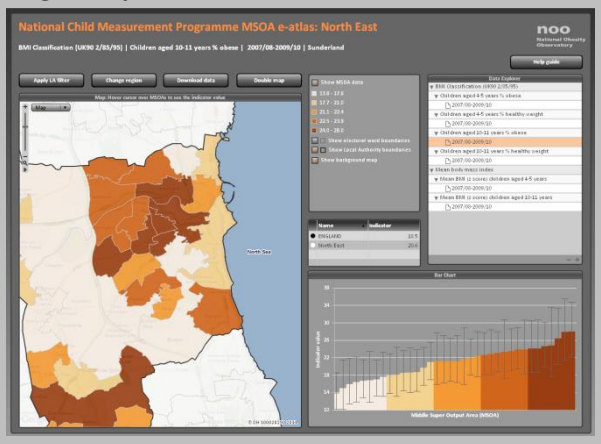
The principal disadvantage of this approach is that it makes it more difficult to monitor change over time. As each statistic includes data from a number of years it is likely to take longer to detect any significant trends. In addition, variables such as area of residence were less complete in earlier years of NCMP measurements which means that it may not always be possible to make full use of data from early years of the NCMP. However as more years of NCMP data become available analysis can be performed using three-year moving averages to examine changes over time.

An example of this approach is the recent MSOA level child e-atlas, published by NOO, which uses three years of NCMP measurements in order to provide more robust data at this level than could be achieved by using one year of NCMP data alone. This approach is described in Case Study 2.

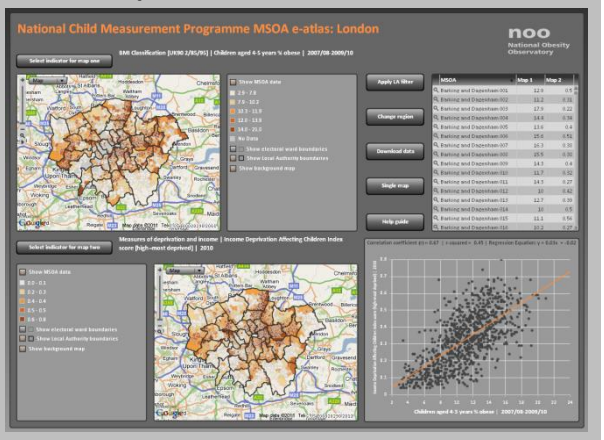
When three years of NCMP data (2007/08, 2008/09, and 2009/10) are combined, the proportion of MSOAs that show a statistically significantly different (either higher or lower) prevalence of obesity to the England average is around double that of analysis undertaken using only one year of data. As shown in Figure 2, using one year of data, for Reception year 10% and for Year 6 13% of MSOAs show a statistically significantly different obesity prevalence. When three years of NCMP data are combined this proportion of MSOAs increases to 19% for Reception year and 28% for Year 6.

Case Study 2: Middle Super Output Area NCMP e-Atlas

Single Map e-Atlas



Double Map e-Atlas



The MSOA e-atlas can be viewed at: http://www.noo.org.uk/visualisation/eatlas

The e-atlases are interactive tools for interrogating data. Data for a series of different indicators can be viewed on maps and charts, and also downloaded in an Excel spreadsheet. NOO has created e-atlases on obesity and its determinants at PCT, LA, and more recently at MSOA level. The MSOA e-atlases enable the examination of NCMP data at neighbourhood level, providing much needed detail on the variation in prevalence of obesity and healthy weight at sub-LA level.

In order to provide robust estimates of obesity and healthy weight prevalence at MSOA level, three years of NCMP data (2007/08, 2008/09, and 2009/10) have been combined. Combining three years of data (rather than just using data from one year) shows a better correlation with the known determinants of obesity (e.g. deprivation). It also allows a greater proportion of areas to be classified as having an obesity prevalence that is higher or lower than the national average with statistical significance.

In order to both prevent potential disclosure of data that might identify individual children, and to ensure a robust indicator of BMI status, data for any MSOA with fewer than 100 children measured across the three years have been suppressed. This is in addition to the suppression of data for MSOAs where only between one and five children have been classified as obese or healthy weight. Where data have been suppressed MSOAs are shaded light grey, indicating 'no data'.

To calculate MSOA prevalence in the NCMP dataset, an MSOA was assigned to each child record based on their LSOA of residence. An LSOA of residence was allocated to each record at the PCT during the process of upload of NCMP data to the Information Centre, where a valid postcode for a child was provided.

In the 2008/09 and 2009/10 NCMP datasets 99% of child records have an LSOA of residence, while in the 2007/08 NCMP 95% of records were assigned an LSOA. Any records without valid LSOA coding are excluded from LSOA and MSOA analysis. The proportion of records with an LSOA of residence for the child varies between PCTs. Additionally, a small number of records have the postcode (and therefore LSOA) of the school or the PCT, instead of the child's postcode/LSOA. These small errors or missing values in the NCMP dataset are unlikely to affect analysis at PCT, LA, or national level, but can be more problematic with small area analysis. It is therefore important to check local data quality before undertaking analysis using the child LSOA field.

B. Combine small geographic areas into larger population groups and neighbourhoods

In some situations it may not be suitable or possible to combine NCMP data from different years of measurement. This may be because the purpose of the analysis is to monitor change over time within small populations, or because NCMP data for previous years have a large proportion of missing data for LSOA of residence within the population being assessed. In other situations even data based on all available NCMP data still does not provide a large enough number of child measurements to create robust analysis. Where this is the case it may be necessary to combine areas or populations to create larger groups or clusters instead of, or in addition to, combining data from different years of measurement (see Case Study 3).

Such clustering is usually based on locality for example, combining neighbouring areas to create larger geographies. Such areas could be created on an *ad hoc* basis, depending on the area of interest, but could also be based on existing areas, such as children's centre areas.

As the smallest geographic indicator available in the NCMP is LSOA of residence, any clustered data will need to be assembled from combinations of LSOAs or larger geographies. In some cases LSOA may not match exactly to the areas required for analysis, but it should usually be possible to create a 'best fit' using LSOAs.

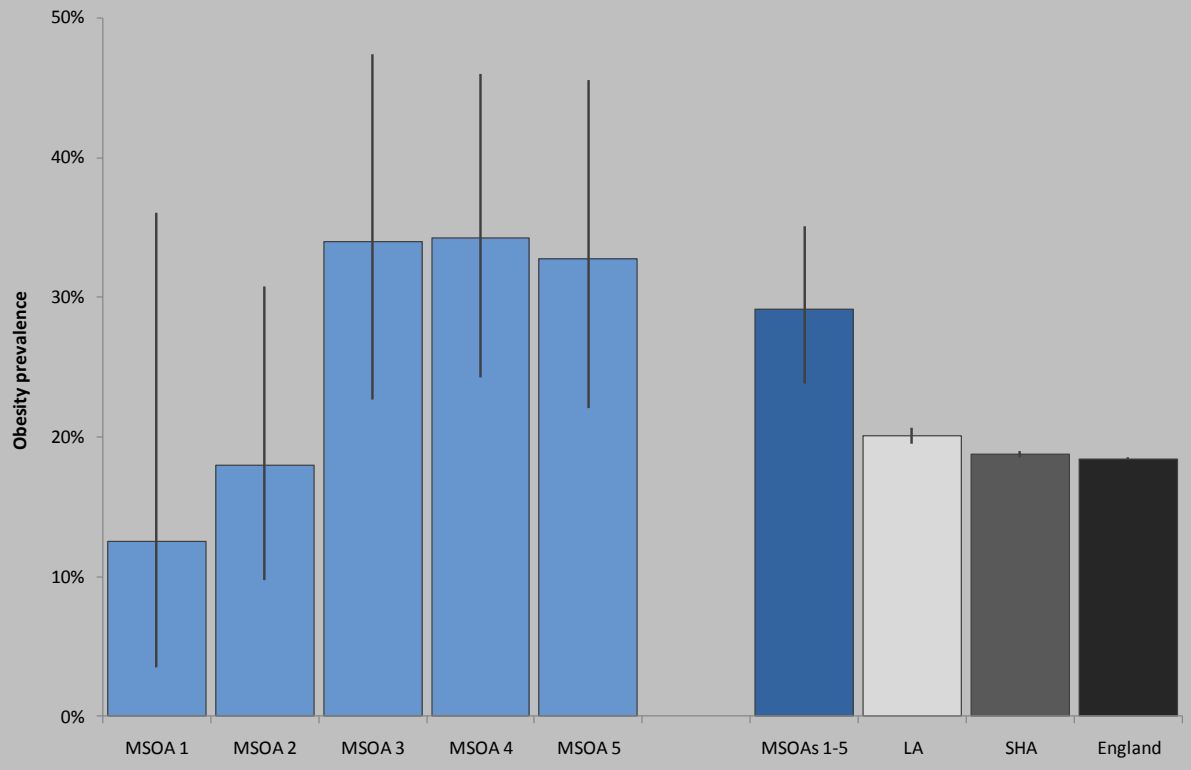
By the very nature of this approach the resulting analysis will have the disadvantage of being less detailed than if areas are not combined. However, in many situations this may be outweighed by the benefits of being more time sensitive or not needing to rely on incomplete data from earlier years.

Case Study 3: Providing data as evidence for targeting resources at neighbourhood level

Recently NOO was contacted by a local level organisation that wished to assess obesity prevalence among children living close to an urban park that was being considered for closure and sale for private development.

The MSOA e-atlas did not show any significant differences from the national average within the area around the park, because of the relatively small number of children who had been measured for the NCMP and lived in that area. However, by combining data from the five MSOAs surrounding the park it was possible to show that obesity prevalence around the park was higher than the obesity prevalence for the LA as a whole, and also higher than the national and regional averages.

Figure 3: Prevalence of obesity among Year 6 children, an example of increased robustness through combining MSOA data, NCMP 2007/08 to 2009/10



C. Combine populations with similar characteristics

For some analyses it may be preferable to create larger population groups based on similar characteristics such as ethnicity or socioeconomic status, rather than confining analysis to neighbouring geographic locations.

Often such groups will be based on the subject under investigation. For example, if the purpose of analysis is to investigate links between socioeconomic status and obesity prevalence within a local area, this can be done by dividing the local population into a limited number of groups based on levels of deprivation. The number of groups can be chosen to maintain an acceptable number of child measurements in each, ensuring that the obesity prevalence estimates for all groups are robust.

This approach is preferable to calculating obesity prevalences for a large number of smaller population groups and then comparing these data with measures of deprivation for each group, as the prevalence estimates for each of the individual small population are unlikely to be robust.

Case Study 4: Office for National Statistics Area Classification

The Office for National Statistics (ONS) has published the 2001 Area Classification, a population stratification system derived from Census data. Further information on this system (including the data files required to perform analysis) can be downloaded from the ONS website

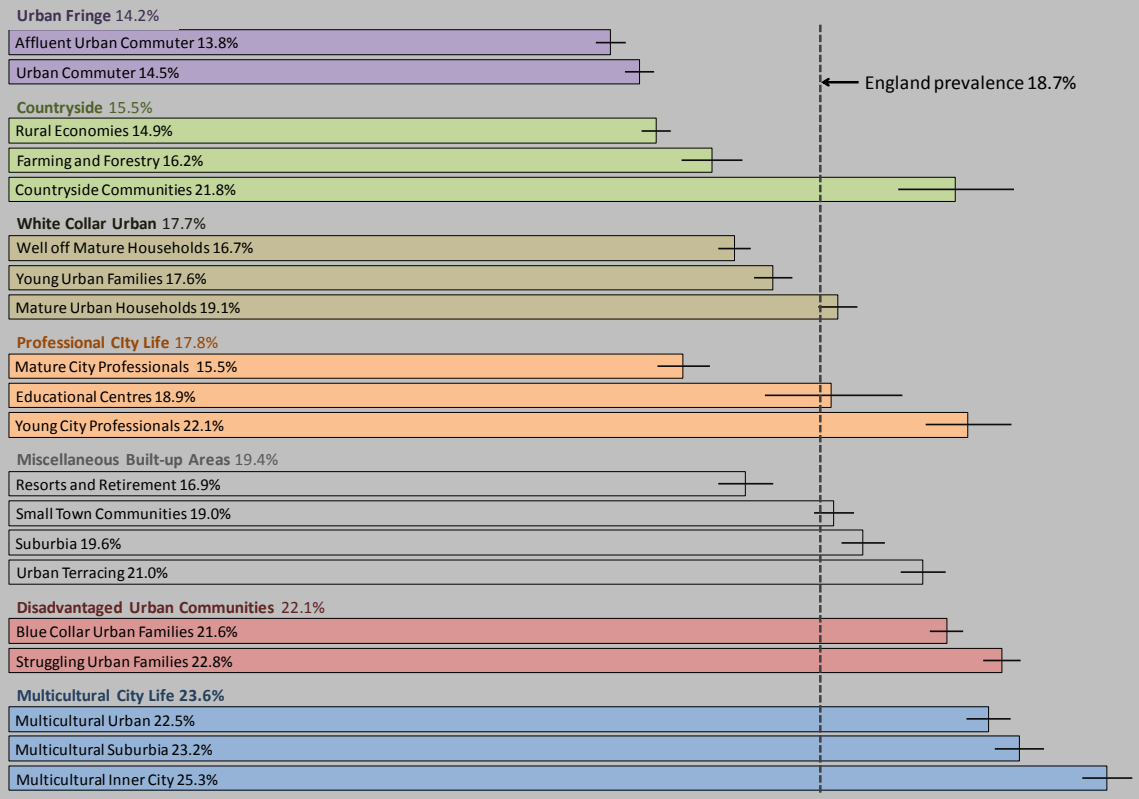
http://www.statistics.gov.uk/about/methodology_by_theme/area_classification/.

The ONS Area Classification groups together geographic areas according to common population characteristics. The LSOA classification divides the UK population into seven 'supergroups' (for example, 'Professional City Life' and 'Disadvantaged Urban Communities') which break down further into 20 'groups' (for example, 'Young City Professionals' and 'Struggling Urban Families').

Figure 4 shows the prevalence of obesity in 2009/10 for Year 6 children in England, from the NCMP, by both LSOA supergroup and group. The prevalence of obesity is generally higher in urban areas than in the 'Countryside' supergroup, with the highest rates in the 'Disadvantaged Urban Communities' and 'Multicultural City Life' supergroups. This confirms earlier findings from NCMP analysis. Moreover, the use of the 2001 Area Classification provides additional detail. For example, the lowest prevalence of obesity is found in areas classed as 'Urban Fringe', showing that the urban environment is not always associated with high obesity prevalence.

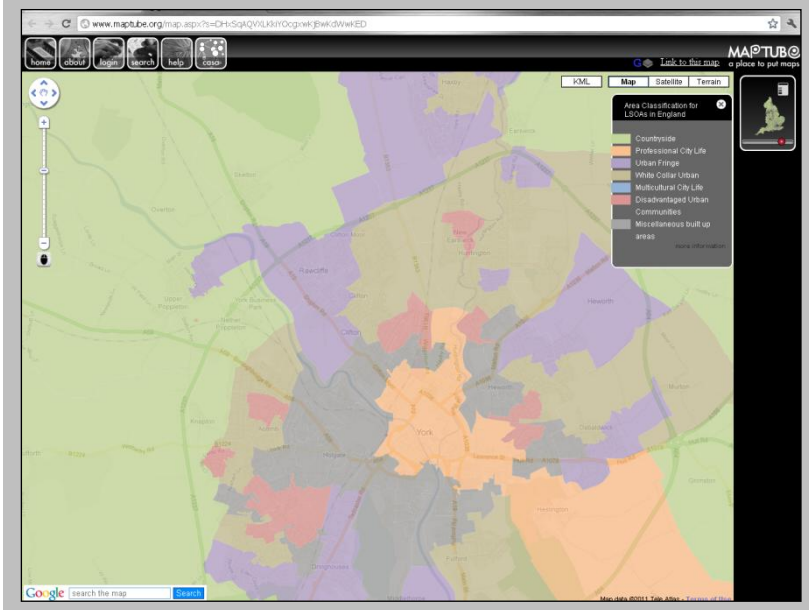
There is also variation within the Area Classification supergroups. Within areas classed as countryside there is a significant difference in obesity prevalence between 'Countryside Communities' and 'Farming and Forestry' or 'Rural Economies'. Similar differences can be seen within the 'Professional City Life' supergroup.

Figure 4: Prevalence of obesity for Year 6 children measured in the 2009/10 NCMP, by ONS LSOA Area Classification supergroup and group, with 95% confidence limits



Local areas may also wish to conduct their own analyses using the 2001 Area Classification with the NCMP data. This will provide more detail on which population groups within these supergroups have the highest prevalence of obesity locally. Similar analyses can be conducted using other population stratification systems such as Acorn from CACI, or Mosaic from Experian.

Area classification map available on MapTube website



NOO has published a map on the MapTube website which shades LSOAs by area classification group.

Using this map it is possible to identify the communities within a local area which are most likely to have high obesity prevalence.

Link to the MapTube website: http://bit.ly/nzoZ6W

D. Combine data from Reception and Year 6

The majority of NCMP data are published separately for Reception and Year 6, and the NCMP guidance for analysis from NOO advises against combining data from school years in the majority of situations. However, combining data from different school years does provide a simple way of maximising the number of child measurements used in analysis, and in some situations may produce an indicator which is more suitable for the purposes required than separate figures for each age group.

As indicators such as obesity prevalence differ substantially between these two age groups, any figure that combines age groups is likely to be affected by the age composition of the population. For example, if age groups are combined, an area where more children were measured in Year 6 than in Reception is likely to report a higher prevalence of obesity than an area with an even split. Some form of weighting or standardisation is therefore likely to be required.

The primary disadvantage of combining data from the two NCMP age groups is that the resulting indicator may not be as easily understood as a straight prevalence figure. It should be stressed in any publications that the resulting figure does not relate to the child population aged 4–11 years, but to an average of the chosen indicator across the 4–5 and 10–11 age groups. A figure for the age group aged 4–11 years would require information about children aged 6–9 years, which the NCMP cannot provide.

This issue is less of a concern for larger populations (where it is likely that the ratio between children measured in Reception and Year 6 is closer to 50:50); however for smaller populations there is likely to be more variation in the age composition of the children measured which would need to be accounted for.

Some form of standardisation is likely to be required in order to account for such variation. A very simplistic 'age standardised' measure could be achieved by taking an average of the indicators for Reception and Year 6, given that the actual ratio between the age groups in the underlying population is likely to be very close to 50:50. However this simple approach does not allow for confidence limits to be calculated easily.

Therefore a slightly more complex approach is required using either direct or indirect age and sex standardisation. The precise method chosen should be influenced by the requirements for the analysis. Such an approach is likely to require some input from an analyst experienced in such procedures and is not covered in detail here. More information on methods of standardisation can be found in the Public Health Observatories in England briefing 'Commonly used public health statistics and their confidence limits'.[e]

[e] http://www.apho.org.uk/resource/item.aspx?RID=48457

7. Further information and advice

Information on the NCMP dataset and general guidance for analysis is available on the NOO website at http://www.noo.org.uk/NCMP. In addition to this guidance and the general guidance for analysis of the NCMP dataset, NOO has recently published a simple guide to the classification of children's BMI which describes the different methods that can be used to interpret BMI in children and young people.[f] ONS has published a useful guidance document on analysing change over time for small areas.[g]

Any queries not covered in these guidance documents can be emailed to ncmp@noo.org.uk. The analysts in the NOO team are available to discuss and provide advice on analysis of NCMP data.

NOO is interested in receiving examples of local analysis of NCMP data. These resources could be made available on the NOO website as illustrations of best practice.

[f] http://www.noo.org.uk/uploads/doc/vid_10612_0910_NCMP_PHO_Analysis_Guidance.pdf
http://www.noo.org.uk/uploads/doc/vid_11762_classifyingBMIinchildren.pdf

[g] http://www.statistics.gov.uk/cci/article.asp?id=2228

8. Reader information

Title	NCMP: Guidance for small area analysis
Authors	Hywell Dinsdale Caroline Ridler
Acknowledgements	Shireen Mathrani, National Obesity Observatory Harry Rutter, National Obesity Observatory Sharif Salah, The NHS Information Centre for health and social care Di Swanston, National Obesity Observatory Rosie Taylor, Department of Health Katherine Thompson, Department of Health
Publication date	July 2011
Target audience	Public Health Observatories Primary Care Trusts Local Authorities
Description	This paper explains why some caution must be exercised when using NCMP data to provide information about small populations. Alternative means of identifying areas of high obesity prevalence at sub local authority level are also examined.
How to cite	Dinsdale H, Ridler C. National Child Measurement Programme: Guidance for small area analysis. Oxford: National Obesity Observatory, 2011.
Contact	National Obesity Observatory www.noo.org.uk ncmp@noo.org.uk
Electronic location	www.noo.org.uk/NCMP
Copyright	

National Obesity Observatory

DELIVERED BY SPH Solutions for Public Health