National Child Measurement Programme
Guidance for data sharing and analysis

About Public Health England

Public Health England's mission is to protect and improve the nation's health and to address inequalities through working with national and local government, the NHS, industry and the voluntary and community sector. PHE is an operationally autonomous executive agency of the Department of Health.

This document is available in other formats on request. Please contact publications@phe.gov.uk.

Contents

1. Introduction

1.1 The National Child Measurement Programme (NCMP) is an annual programme which measures the height and weight of children in Reception (age 4–5 years) and Year 6 (age 10–11 years) within state maintained schools in England. Some independent and special schools also choose to participate, however these records are usually excluded from the analysis for national NCMP reports as the majority of such schools do not take part (see section 4.3).

1.2 The Health and Social Care Information Centre (HSCIC) and Public Health England (PHE) publish analyses of this dataset annually. However, local authorities might want to undertake additional analyses to inform their work on the healthy weight agenda.

1.3 The 2010 Department of Health (DH) obesity strategy *Healthy lives, healthy people: a call to action on obesity in England*[1] and the Public Health Outcomes Framework (PHOF)[2] outline the Government's ambition to achieve a sustained downward trend in child excess weight prevalence. The PHOF uses NCMP data to help local authorities monitor progress towards the ambition, http://www.phoutcomes.info/.

1.4 This document provides guidance to local authorities and other organisations who wish to make use of the NCMP data. It covers:

- an overview of the analyses provided by PHE and the HSCIC
- how to access the NCMP dataset
- which NCMP data can be shared between organisations
- guidance on the appropriate use of the NCMP dataset, to comply with data protection and disclosure rules
- caveats associated with the NCMP data and its interpretation
- suggestions for regional and local analyses
- FAQs for NCMP analysis
- suggestions for local data quality checks

1.5 This guidance document relates predominantly to the 2012/13 NCMP. The HSCIC has launched a new online NCMP data collection system for the 2013/14 and future NCMP years. Local authorities will access their analysis dataset for 2013/14 directly from this online system following the collection period. PHE will continue to receive an analysis dataset in MS Access directly from the HSCIC. This guidance document will be updated in 2015 to reflect these changes.

1.6   If users of the NCMP dataset have additional queries not covered in this guidance, they should contact Public Health England at ncmp@phe.gov.uk.

2. National reports and resources

2.1   A number of NCMP resources to facilitate local use of the NCMP data are produced each year by both the HSCIC and PHE. These consist of reports, data factsheets, and slide sets presenting and describing the NCMP data at national level, and also data tables at national, regional, local authority, and Middle Super Output Area (MSOA) level. Users of the NCMP data are advised to familiarise themselves with these reports and data tables before performing analysis at a more local level. The main resources available are described in the following sections.

NCMP data and analysis provided by the HSCIC

2.2   Headline results from the 2012/13 NCMP, at national and sub-national level, are provided in the HSCIC's report National Child Measurement Programme: England, 2012/13 school year, published in December 2013.

2.3   This report describes the prevalence of child obesity, overweight, healthy weight, and underweight at national and regional level, as well as the variation across demographic and socioeconomic groups. A comparison is also provided with the data from previous years.

2.4   Excel data tables are published alongside the report, they contain data on obesity, overweight, healthy weight and underweight prevalence figures for local authorities (based both on location of school and area of residence), SHAs, regions, and PCTs for 2012/13. Additional national level data are included broken down by ethnic group, rural-urban classification, and Index of Multiple Deprivation (IMD) decile.

2.5   Both the report itself, and the supporting excel files for the 2012/13 and previous years of NCMP data are available on the HSCIC's NCMP pages. http://www.hscic.gov.uk/ncmp

NCMP data and analysis provided by PHE

PHE NCMP analysis reports

2.6   An update to the series of reports describing the trends in obesity prevalence and mean body mass index (BMI) will be published by PHE during 2014 to complement the HSCIC's 2012/13 report.

2.7   This report and the 2008/09 to 2011/12 analysis reports examine the changes that have taken place in the BMI of Reception and Year 6 children over the period for which NCMP data are available. These reports also examine the impact of demographic and socioeconomic variables (principally ethnicity and deprivation) on both the patterns and trends observed.

2.8   Users of the NCMP dataset might also find it useful to read the reports on the 2006/07 and 2007/08 NCMP datasets. These reports examine some of the data quality issues observed within the NCMP database for those years and describe the possible effect of these on reported prevalence of overweight and obesity.

2.9   Previous reports published using NCMP data include short reports describing the use of the ONS Area Classification (ONS-AC) (see section 4.16 and 6.14) in NCMP analysis,[3] and the impact of month of measurement on reported obesity prevalence within the NCMP dataset.[4]

2.10 All NCMP analysis reports published by PHE are available on the PHE Obesity website: http://www.noo.org.uk/NCMP

NCMP local authority profile

2.11 A **new online data tool displays** data from the NCMP 2006/07 to 2012/13. Prevalence of underweight, healthy weight, overweight, and obesity for children in Reception and Year 6 can be examined at local authority level. Data quality indicators are also available in the tool, for example rate of participation in the NCMP. This online tool can be accessed here: http://fingertips.phe.org.uk/profile/national-child-measurement-programme

Other PHE NCMP data publications

2.12 **Spreadsheets** containing NCMP obesity prevalence data for Reception (age 4–5 years) and Year 6 (age 10–11 years) children by 2011 Electoral Ward and 2001 Middle Super Output Area (MSOA) of residence with LA, SHA and England data for comparison is available on the PHE Obesity website

http://www.noo.org.uk/visualisation. These small area prevalence figures use three years of NCMP combined. Data for 2010/11 to 2012/13 will be added to the PHE Obesity website during early 2014. NCMP data at MSOA level are also included in the PHE Local Health tool: http://www.localhealth.org.uk

2.13 **PowerPoint slides** presenting key data and information on child overweight and obesity in clear, easy to understand charts and graphics using NCMP and Health Survey for England data are available on the PHE Obesity website. These can be used freely with acknowledgement to Public Health England. http://www.noo.org.uk/slide_sets

2.14 **Factsheets** which compile up-to-date key information and data about obesity and its determinants in an easily readable format: http://www.noo.org.uk/NOO_pub/Key_data

2.15 **NCMP Operational Guidance** provides more information on the NCMP and advises local commissioners and providers of the NCMP on its implementation. https://www.gov.uk/government/publications/national-child-measurement-programme-operational-guidance-for-the-2013-to-2014-school-year

3. Obtaining and sharing the NCMP dataset

Obtaining the data

3.1 The NCMP regulations[5] allow the HSCIC to make available a cleaned copy of the NCMP dataset to PHE Knowledge and Intelligence Teams (KIT). This is provided via 'Data Depot', a secure, web-based electronic file transfer system. The dataset contains an individual record for each child measured in the NCMP.

3.2 Local authorities can request a cleaned copy of their own NCMP data from PHE KITs. Local authorities are able to access individual child records for the children attending schools within their boundaries.

3.3 The data sharing agreement between the HSCIC and PHE states that all record level data must be treated in accordance with the Data Protection Act 1998,[6] and any data will, therefore, need to be transferred safely and securely between PHE KITs and local authorities by nhs.net or the local authority version gcsx.gov.uk.

3.4 Although local authorities will already have access to their own NCMP data, it is recommended that cleaned data from the HSCIC are used for local analysis, rather than the records held locally. This is important to ensure potentially invalid records are not included in the analysis and for consistency with published figures. Furthermore, the HSCIC dataset is likely to hold a greater number of fields than locally held data (see Appendix 1) which facilitates a greater range of possible analyses.

Sharing data with other organisations

3.5 Other organisations such as the NHS and academic institutions may request NCMP data from local authorities, PHE or from the HSCIC. However there are restrictions over what data can be shared, as outlined below.

3.6 Often the information needed will already be available through the HSCIC and PHE websites, or via the UK Data Archive (UKDA) at http://www.data-archive.ac.uk. Where possible users should be directed towards these existing resources.

3.7   The datasets on the UKDA contain most of the variables contained within the full NCMP dataset, but ethnicity coding and detailed geographic coding (which includes any geographic areas smaller than local authorities) may have been removed to prevent individual children being identified.
http://ukdataservice.ac.uk/

3.8   In some cases the information requested will not be publicly available and therefore will need to be produced using the individual row level dataset. Local authorities are able to provide such information, but this may need to be provided in an aggregated and suppressed form. This ensures individual information is not shared unnecessarily and prevents the identification of individual children.

3.9   PHE is available to co-ordinate requests for NCMP data from academic organisations or individuals on behalf of local authorities if required and will liaise with the HSCIC in responding to the request.

Processing of information by local authorities

3.10 The NCMP regulations[6] state that local authorities (or commissioned services working on behalf of the local authority) may provide NCMP data to any person where it will be used for the purposes of research, monitoring, audit or the planning of services, or of any purpose connected to public health, subject to the condition that the information may be disclosed **only in a form in which no individual child can be identified.**

3.11 Local authorities (or commissioned services working on behalf of the local authority) are able to share personal information from the NCMP necessary to identify a child with a view to the information being:

  (a) communicated to a parent of the child to whom the information relates, together with advisory material relating to the weight of children

  (b) used for the purpose of providing advice and assistance to a parent of the child to whom the information relates with the aim of promoting and assisting improvement of the child's health

  (c) disclosed by the local authority to a health professional who is in a position to provide the advice and assistance referred to in point (b) above and to offer any related treatment to the child

Onward processing of information by the HSCIC

3.12 The NCMP regulations[6] state that the HSCIC and PHE as detailed in the data sharing agreement may provide NCMP data to any person where it will be used for the purposes of research, monitoring, audit or the planning of services, or of any purpose connected to public health, subject to the condition that the information may be disclosed **only in a form in which no individual child can be identified.**

Feeding back NCMP data to schools

3.13 Many schools are keen to receive detailed feedback showing the results of the child measurements that have taken place within their school. Under the regulations, **it is not permissible to share individual child records with schools.**

3.14 Local authorities are advised against routinely feeding back NCMP statistics such as obesity prevalence at a school level on an annual basis. The reasons for this are:

- with small denominator populations, such as those for primary schools, the numbers of overweight and obese children are likely to be small. In many cases publication or sharing of these small number data is therefore not possible as it might allow individual children to be identified.

- most schools will have less than 100% participation. Some groups, such as overweight or obese children, may be more likely to opt-out of being measured than others. This could introduce bias into the results and render them less reliable at school level.

- class sizes in primary schools are small, so school-level prevalence figures will be subject to small number variation and could be affected by unequal sex ratios (see section 7.6 to 7.8). They would, therefore, not provide robust measures of obesity prevalence, even if there was 100% coverage of all children in the relevant age-groups within a school.

3.15 If local areas wish to conduct more detailed analysis of obesity prevalence at school level, this can be done in a number of ways. Data from more than one measurement year can be used to increase the number of child records used in the calculation of prevalence figures. Alternatively, schools could be clustered geographically, or according to shared characteristics such as deprivation or levels of pupil attainment.

3.16 Whilst such techniques are likely to result in more robust estimates of obesity prevalence at school level, it is important to exercise caution if these statistics are to be published or fed back to schools. Care must be taken to ensure schools support the release of these figures and to ensure the data are not misinterpreted as a measure of whether a school provides a 'healthy environment' or not. Most of the variation between schools in statistics such as obesity prevalence is explained by the sociodemographic mix of the pupils attending the school, rather than the school environment.

3.17 Previously the National Obesity Observatory (PHE Obesity Knowledge and Intelligence) provided an Excel Tool to help PCTs produce feedback letters for schools. PHE is currently developing a new resource to support local authorities to share NCMP information with schools. PHE will be consulting with stakeholders to develop a useful product. It is expected that the new school NCMP feedback resource will be available from September 2014.

4. The NCMP national analysis dataset

4.1   The NCMP analysis dataset consists of a single MS Access database with all the information required for local analysis. The database is organised into a series of tables, which are shown in Appendix 1. The tables hold data at pupil, school, local authority (derived from both school location and child residence), and Government Office Region (GOR) (derived from both school location and child residence). In addition, a series of lookup tables are provided to allow the interpretation of coding used within the dataset (for example ethnicity and BMI classification).

4.2   This dataset has been cleaned according to the HSCIC's data-cleaning protocol.[7] All invalid records have been removed but can be viewed in a separate table (Pupil_data_excluded).

4.3   The dataset includes a column labelled 'Exclude_flag'. Records from independent schools and special schools are coded as 1. Local authorities are not required to take measurements at these schools, and many independent and special schools do not participate in the NCMP. As a result these records cannot be seen as a representative sample of pupils attending such schools and are therefore not included in the national analysis. **These records need to be excluded if analysis is to match the HSCIC's figures**.

4.4   It is important that users of the NCMP dataset are aware of the assigned fields provided, and how these are calculated. A full list and description of all fields within the NCMP dataset is provided in Appendix A and the sections below provide more information on key fields.

4.5   From the 2013/14 NCMP collection year local authorities will access their analysis dataset from the new HSCIC NCMP online system.

Geographic coding

4.6   The NCMP dataset contains a number of fields providing geographic information for each child record. These fields can be assigned based on the former PCT that took the measurements, the school attended, or the child's area of residence, as shown in Table 1 below.

4.7   The different methods of assigning geographies may lead to small differences between figures for LAs or GORs in some cases, even where these areas are apparently coterminous.

4.8   Most published NCMP statistics at local authority level use geographic information derived from the school postcode, rather than the child's area of residence. In the first few years of the NCMP, a considerable number of child records were provided without information on area of residence so any summary statistics based on this variable would have to exclude some children from the analysis. Local authority of school location has usually been used to ensure all child records can be included in the published analysis.

Table 1: Geographic coding in the NCMP dataset

Geography:	Derived from:
Primary Care Trust	Former PCT taking measures
Strategic Health Authority	Former PCT taking measures
Local Authority (Tier 1 and 2) of school	School postcode
Government Office Region of school	School postcode
School Lower Super Output Area (LSOA)	School postcode
Child LSOA	Child postcode[a]
Local Authority (Tier 1 and 2) of child	Child postcode
Government Office Region of child	Child postcode

4.9   The proportion of child records with valid coding for area of residence has improved considerably over the years of the NCMP, and in 2012/13 area of residence coding could be assigned to 99.8% of records (see Figure 3).

4.10 As a result of this improvement, NCMP statistics provided by the HSCIC include data based on the area of residence of the child. In addition, Local Authority and Government Office Region coding based on the area of residence of each child has been included in the dataset.

4.11 There are a small number of records in the 2012/13 NCMP dataset where a child record has coding for the urban / rural classification based on child residence, but could not be assigned to an English local authority. This occurs where a child's residence is in Scotland or Wales, but they attend a school in England.

---

[a]Child postcode is converted to a Lower Super Output Area (LSOA) code at the time of upload of data to the HSCIC.

4.12 PHE KITs can share the NCMP 2012/13 analysis dataset with public health teams in upper tier local authorities. Each local authority public health team can receive the records for all children measured in schools within their local authority.

Figure 3: Proportion of records with valid child postcode and ethnicity coding (England totals by year of measurement)

Percentage of records with **valid child postcode**

Percentage of records with **valid ethnicity coding**

2006/07   2007/08   2008/09   2009/10   2010/11   2011/12   2012/13
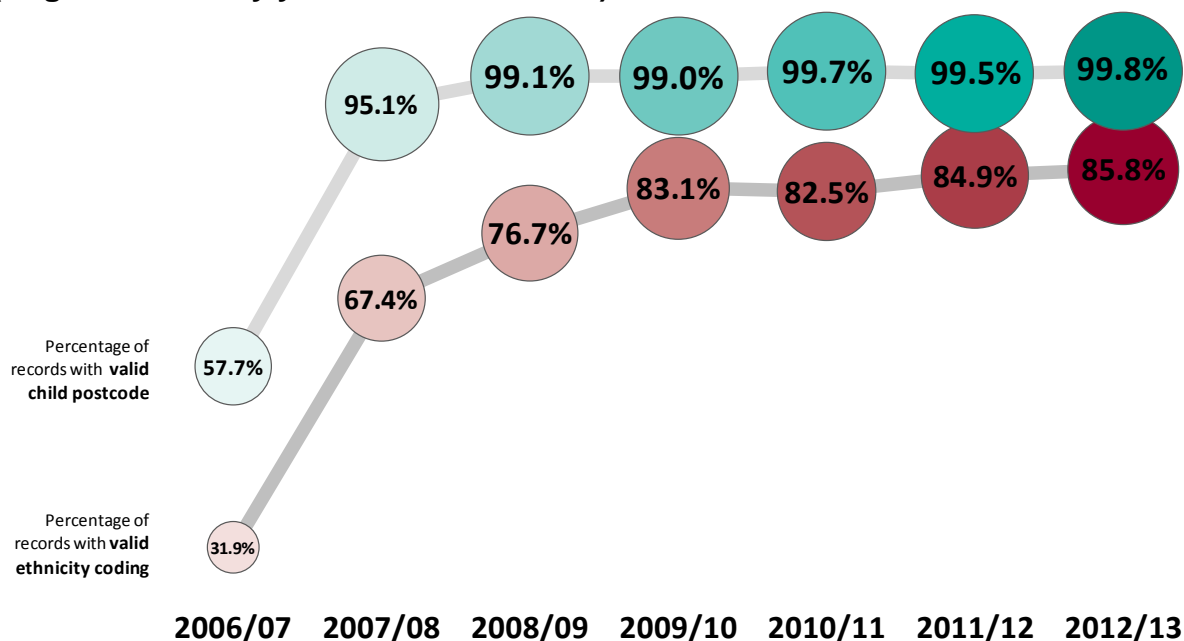
Geographically assigned variables

4.13 The NCMP dataset contains a number of variables assigned based on the geographic coding. These include an indicator of deprivation, an urban/rural classification, and a category based on the Office of National Statistics Area Classification (ONS-AC) (see section 4.16).

4.14 These indicators are provided based both on the LSOA of the school and the LSOA of residence of the child. The former indicators are available for all pupils, whereas the latter is not provided for those children where the LSOA of residence is not available. Indicators based on the LSOA of residence will typically produce more accurate analysis of the resident population in an area. However in areas where a substantial proportion of children measured have no coding for area of residence, indicators based on the school location may need to be used. This is only really an issue with the 2006/07 and 2007/08 NCMP datasets as from 2008/09 more than 99% of child records had a valid postcode of residence.

4.15 The deprivation indicator provided in the dataset consists of an Index of Multiple Deprivation (IMD) 2010 decile. IMD decile one includes those LSOAs within the least deprived 10% of all the LSOAs in England, whereas decile ten contains the most deprived 10%. More information on the IMD is available here:
https://www.gov.uk/government/publications/english-indices-of-deprivation-2010

4.16 The ONS-AC is a system of population stratification that categorises local areas based on a range of sociodemographic characteristics, including deprivation, ethnicity, and urban/rural environment. The categories are named in a way that describes the type of population predominant in those areas, for example 'Disadvantaged Urban Communities' or 'Professional City Life'. More information on the ONS urban/rural and Area Classification indicator is available here:
www.neighbourhood.statistics.gov.uk/dissemination/Info.do?page=nessgeography/areaclassification/area-classification.htm

See sections 2.9 and 6.14.

Child measurements

4.17 In addition to the actual height, weight, and BMI of each child, the NCMP dataset contains the appropriate z scores (standard deviation scores) and centiles for BMI, height and weight for each child. Appendix 3 of this guidance provides more information on how these variables are calculated and how they are used.

4.18 Two fields are provided in the dataset which show whether that child is underweight, healthy weight, overweight, or obese.

4.19 The 'BMI_pop_mon_class' field in the Pupil_data table shows whether individual children are classified as underweight, healthy weight, overweight or obese, for population monitoring according to the 2nd, 85th and 95th centiles on the British 1990 growth reference (UK90).[8,9]

4.20 The 'BMI_clinical_class' field in the Pupil_data table uses the clinical cut-offs of the UK90 BMI growth reference which classify children into these categories according to the 2nd, 91st and 98th centiles of the UK90 BMI growth reference (see Appendix 3).

Table 2a: Codes used within 'BMI_pop_mon_class' field

BMI classification	UK90 population monitoring BMI centile range	School year	
		Reception	Year 6
Obese	>=95th centile	1	2
Overweight	>=85th and <95th centile	3	4
Underweight	<=2nd centile	5	6
Healthy weight	>2nd and <85th centile	7	8

Table 2b: Codes used within 'BMI_clinical_class' field

BMI classification	UK90 clinical cut-offs BMI centile range	School year	
		Reception	Year 6
Obese	>=98th centile	1	2
Overweight	>=91st and <98th centile	3	4
Underweight	<=2nd centile	5	6
Healthy weight	>2nd and <91stcentile	7	8

5. Essentials for NCMP analysis

5.1 When making NCMP data publicly available data must only be released in a form where children cannot be identified. Suppression may need to be applied and corresponding cells providing totals should also be suppressed to avoid disclosure by differencing.

5.2 Users should familiarise themselves with the published data to ensure they are not just duplicating analysis that has already been performed nationally.

5.3 Wherever possible, local analysis should be checked against the figures published by the HSCIC and PHE to ensure consistency.

5.4 Pupils measured attending independent and special schools should be excluded to match published data (see section 4.3).

5.5 Confidence limits are published in the Excel data tables provided by the HSCIC and PHE and these should be used when comparing between areas or monitoring change over time. Users of the NCMP dataset should also apply confidence limits or statistical tests to their own analysis. Methods for doing this are outlined in Appendix 4 of this guidance.

5.6 Any publications using NCMP data should clearly state the thresholds used (usually the 85th and 95th centiles of the UK90 BMI growth reference) to derive obesity, overweight, healthy weight, and underweight prevalence figures.

5.7 Any published analysis should include 'The Health and Social Care Information Centre (HSCIC)' as the data source for NCMP data.

5.8 PHE are keen to see any local NCMP analysis reports in order to be able to share the methods and ideas with other local authorities.

6. Suggested analysis at local level

Data quality and participation

6.1   Previous analysis has shown that low levels of participation in the NCMP and poor data quality may affect the reported prevalence of obesity. At national level both participation and data quality have improved with each year of the NCMP, however there are still some parts of the country where data quality indicators show a need for improvement. It is therefore important that these issues are examined at a local level, both to determine whether further improvements can be made, and also to explore whether any observed changes in obesity prevalence might be related to changes in participation or data quality. A PCT level data quality metric is included in Annex 1 of the HSCIC's 2012/13 NCMP report. Additionally, local authority level data quality indicators are included in the LA NCMP profile:http://fingertips.phe.org.uk/profile/national-child-measurement-programme

6.2   <u>Participation rate and prevalence:</u> Published analyses of the 2006/07 and 2007/08 NCMP datasets suggest that low rates of participation in the NCMP at PCT level may be associated with a lower reported prevalence of obesity, especially for Year 6. This effect may have been due to selection bias in children who were measured, whereby children who do not participate in the NCMP are more likely to be obese than those who do participate.

6.3   As participation rates have increased, the impact of selective opt-out on prevalence rates appears to have declined in more recent years. Information on the effect of participation on prevalence in the current NCMP dataset is included in Annex 5 of the HSCIC's 2012/13 NCMP report. Although the potential impact of selective opt-out on prevalence in the latest data appears to be small at national level, it is possible that the effects could be greater at a local level. This issue is likely to be particularly important when comparing prevalence figures for areas with very different participation rates, or looking at change over time in areas where participation rates have increased or decreased markedly over the same period.

6.4   Local areas may wish to identify areas of low participation within their NCMP data, and also the reasons for these – for example whether this is due to entire schools not taking part (selective opt-out should not be a factor in such cases), or due to individual children within schools not taking part (selective opt-out may be a factor). This information can then be used to increase participation

for future years of the NCMP. The 'schools_data' table in the NCMP dataset contains all eligible state schools, as well as a field to show which schools participated. This can be used to quickly identify which schools within a local area did not take part in the NCMP.

6.5   If investigating the impact of participation on prevalence, it may be preferable to calculate participation rates which relate to just those schools for which measurements were submitted. This is because if an entire school did not take part in the NCMP, selective opt-out of obese children should not be a factor for children with that school. Such a participation rate is calculated by dividing the number of children measured in each school year by the sum of the eligible pupil numbers in schools in which measurements were taken. This figure will be higher than the published participation rates if there were any schools that did not take part in the programme.

6.6   <u>Participation by sex:</u> Previous NCMP analyses have suggested that participation of girls in the NCMP may be lower than that for boys, and differs by area. If this is due to a selective opt-out of overweight and obese girls from the NCMP measurements, any bias could be stronger for girls than for boys. This issue might also benefit from local analysis and should also be considered when looking at differences in prevalence by sex at a local level.

6.7   <u>Data quality:</u> The national NCMP dataset has undergone extensive cleaning but there is a limit on the checks that can be done nationally with such a large dataset. It is therefore advised that local areas assess the quality of their NCMP data before conducting detailed analysis. More information on suggested data quality checks is provided in Appendix 2.

Sociodemographic variation: effect of ethnicity, deprivation and setting

6.8   Published NCMP analyses show that ethnicity, deprivation, and setting (for example urban/rural environment) may influence prevalence of obesity, overweight and underweight. A number of related indicators, such as IMD decile, urban/rural classification, and ONS Area Classification, have been added to the NCMP dataset in order to facilitate the analysis of these factors.

6.9   Local authorities and PHE KITs can also use their local intelligence and data to determine to what extent the variation within their local area can be explained by these variables. It may be useful to make comparisons with the regional or national average to determine whether the relationship between sociodemographic variables and child obesity prevalence in the local population is the same as that for all children in English schools, or whether a

different pattern is observed. Such information should be useful to assist any local targeting of interventions to tackle unhealthy weight among children.

6.10 Ethnicity: The NCMP dataset contains fields showing the ethnicity of individual children. Children have been coded to one of the 16 NHS ethnicity codes as well as one of the more detailed codes, of which there are approximately 230, used by child health systems. Therefore analyses can easily be performed at a local level for different ethnic groups. To avoid problems due to potentially low numbers of children for some ethnic groups, it may be necessary to combine certain ethnic groups, or combine data from a number of years of NCMP measurements when conducting such analysis at local level.

6.11 Socioeconomic status: Analysis of socioeconomic status can be undertaken using the IMD deciles already assigned to child records within the NCMP dataset. Both IMD decile of the school location and of the child's area of residence are included in the NCMP dataset. Coding based on area of residence is preferable for most analysis, although coding based on the location of school can be used where coding for area of residence is missing for a large proportion of children. If other indicators are required for analysis these can be assigned to individual children using the LSOA of residence or school LSOA.

6.12 Analysis is often best performed by grouping child records according to quintiles or deciles of these socioeconomic indicators, depending on the number of child records available for analysis. Analyses can be produced for these groupings to determine the links between factors such as deprivation and prevalence of obesity. An example of such analysis can be found in the *Child Obesity and Socioeconomic Status* data factsheet available at: http://www.noo.org.uk/NOO_pub/Key_data.

6.13 Setting: As shown in the HSCIC's 2012/13 NCMP report, obesity prevalence can be seen to vary between urban and rural areas. Urban and rural coding is available within the NCMP dataset to allow local areas to investigate these patterns. However, it is likely that much of the differences in obesity prevalence between urban and rural areas can be explained by differences in the sociodemographic mix of the urban and rural populations.

6.14 ONS Area Classification: The ONS Area Classification might provide a more useful way of analysing differences in obesity prevalence that takes account of the urban or rural setting, as well as the demographic and socioeconomic mix of the population. This population stratification system is available for free from the ONS website. More information on this can be found in the report *NCMP Analysis using the ONS Area Classification*.[3] ONS Area Classification 'Group',

'Subgroup', and 'Supergroup' have been assigned to child records with a valid LSOA of residence in the NCMP dataset. Also see sections 2.9 and 4.16.

6.15 Users of the NCMP dataset who already have access to commercially available population stratification systems may use these to perform similar analysis.

Monitoring change over time

6.16 Results from previous years of the NCMP can be downloaded from the HSCIC website (see section 2.5), or produced locally through analysis of the datasets provided to PHE by the HSCIC.

6.17 If users want to make comparisons with data from previous years, then the impact of changing participation rates and changes in data quality between the years should always be taken into account. Appropriate statistical testing should be undertaken to ensure any reported differences are indeed significant. The suggested method for establishing the statistical significance of a change in prevalence is described in Appendix 4 of this guidance.

6.18 When examining change in populations over time, a number of papers have suggested looking at change in a measure such as mean BMI z score, rather than change in prevalence figures.[10,11] This can easily be calculated within the NCMP dataset, using the 'BMI_z' field. As this measure takes account of the whole child population, rather than just the proportion above or below a certain threshold, it may allow changes over time to be detected earlier than if prevalence figures alone are used.

6.19 The data collected during the 2012/13 school year is likely to contain a large cohort of children who were previously measured in Reception during 2006/07. However, as identifiable data are not held centrally, at national level there is limited potential to make detailed comparisons between the measurements from these two years for this cohort. However, in some local areas it is possible to compare the 2012/13 NCMP data for Year 6 children with child measurements for the same children taken in Reception year during 2006/07, see section 7.13 to 7.20.

Small area analysis

6.20 Obesity prevalence figures are routinely published at local authority level. However, many practitioners require information for sub-populations and neighbourhoods. Often such small area analysis is required in order to focus resources on the most at risk areas or communities.

6.21 Sub-LA level geographic analysis is possible using NCMP data. Analysis can easily be performed using school or LSOA as the unit of analysis, or by assigning an MSOA, or ward code from the LSOA code. However, such analysis needs to be performed and interpreted with caution.

6.22 Prevalence figures for sub-LA populations are likely to be based on small numbers and so are subject to a high degree of natural variation. Confidence limits should always be used around prevalence estimates and any differences in prevalence between areas should be tested for statistical significance to make sure the differences are not just the result of the small sample size at this level of analysis (see Appendix 4).

6.23 To tackle the issue of small numbers it is advised that, wherever possible, analyses are conducted using more than one year of NCMP data. Combining three years of NCMP data appears to provide relatively robust figures for obesity prevalence at MSOA level. However, for smaller geographies (for example LSOAs or schools), even four years of data may not be enough to provide robust figures. Furthermore, combining data from different years of NCMP measurements reduces the sensitivity of such measures to any change over time.

6.24 An alternative to combining data from different years of NCMP measurements is to combine small areas into clusters of a size which permits robust analysis.

6.25 Certain issues come into play with small area analyses that are less problematic for larger geographic areas. For example, the sex ratio of children measured is likely to show much greater variation across small populations than across local authorities or regions. As obesity prevalence varies significantly by sex, it is possible that statistics based on small populations could be affected by the sex ratio of children measured as well as by the underlying prevalence of obesity.

6.26 PHE has also published further guidance on using NCMP data for small area analysis. This guidance explains the issues associated with presenting NCMP data at small area level as well as exploring what methods are appropriate and available for analysis to ensure more useful and valid results.[12]

Alternatives to small area analyses

6.27 Rather than producing prevalence figures for small areas and then using these to target resources or investigate determinants, obesity prevalence may be better investigated using groups of pupils based on common characteristics.

For example, the population could be grouped using an Index of Deprivation, ethnicity or by community type.

6.28 This approach has the advantage that, rather than resulting in prevalence figures for a large number of small areas based on a small number of children measured in each area, the dataset is divided into a much smaller number of population groups. As a result, any analyses for each group are based on a greater number of children per group, and so can be considerably more robust.

6.29 For example, child obesity prevalence for each LSOA within a local authority will tend to show only a very weak correlation with an index of socioeconomic deprivation at LSOA level. However, if the NCMP data for the same local authority are grouped into deciles of deprivation (based on the LSOA of residence), this will usually show a strong pattern of increasing obesity prevalence with increased levels of socioeconomic deprivation. By using fewer groups, based on similar characteristics, the underlying pattern can be seen much more clearly.

6.30 Analyses of this form can often provide much more useful information regarding the local determinants of obesity and generally provides more robust information on how best to target resources. If the most deprived parts of a local authority are shown to have the highest prevalence of obesity, then resources can be targeted at these areas with a good degree of confidence that they will be reaching those individuals most at risk of future obesity-related ill health.

BMI distribution

6.31 Users of the NCMP dataset may wish to make use of the full range of height, weight and BMI values to examine the distribution of these measured rather than only considering the proportion of underweight, overweight, and obese children. Figure 4 provides an example of the distribution of BMI z scores within the 2012/13 NCMP datasets for children in Year 6, compared to the 1990 baseline.

Figure 4: Distribution of BMI z score for children in Year 6, NCMP 2012/13
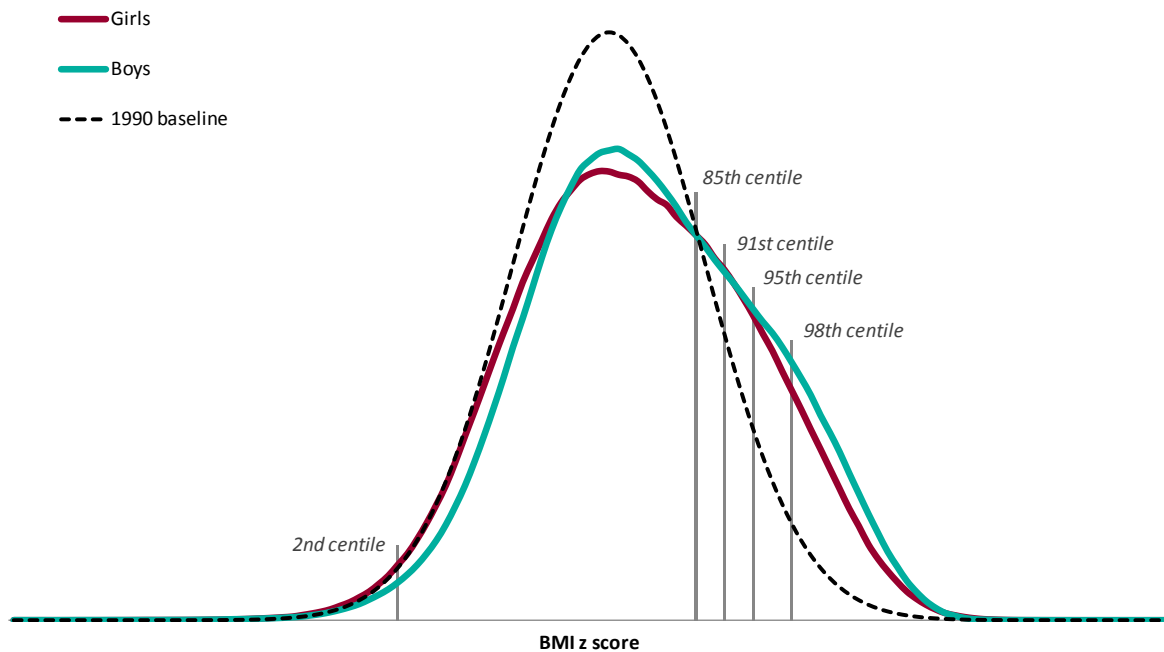
6.32 In this case, the possible confounding effect of age on such analyses has been addressed by using z scores rather than the actual BMI values. The expected height, weight and BMI of children varies substantially with age and sex. Therefore if age and sex are not adjusted for, the shape of the distribution will be affected. These z scores are available in the dataset provided by the HSCIC via PHE.

6.33 It is important to note that the distribution of z scores derived from the UK90 growth reference will be closer in shape to the normal distribution, than the distribution of actual BMI values. These curves should be interpreted in terms of difference from the normal distribution, rather than viewed as representing the current actual population distribution of BMI.

7. Analysis FAQ

7.1   The following questions have been compiled from common queries received by PHE and information included in previous NCMP guidance.

Why should I use the HSCIC's cleaned data?

7.2   Analyses should be based on the cleaned data from the HSCIC rather than using the records held by local authorities, to ensure accuracy and consistency with published figures.

7.3   The HSCIC performs extensive data quality checks before the dataset is distributed to PHE and where data quality issues are identified this information is fed back to local authorities. However local areas may wish to perform further data quality checks of their own. Some suggested checks are outlined in Appendix 2. Furthermore, the HSCIC dataset is likely to hold a greater number of fields than locally held data (see Appendix 1) which facilitates a greater range of possible analyses.

Can I combine data for school years?

7.4   Prevalence figures should usually be produced separately for Reception and Year 6, rather than combining the data. Prevalence of obesity and overweight differ with age, tending to be higher in the older age-groups. As a result, a combined prevalence figure will tend to be lower if a larger proportion of Reception children have been measured, and higher for areas in which a larger proportion of Year 6 children have been measured.

7.5   If combined prevalence figures are produced, they should be age standardised in some way, rather than created by simply combining data for children measured in Reception with that for children in Year 6 to create a 'crude' estimate. A simple way to achieve this involves taking an average of the Year 6 and Reception figures, rather than by calculating a rate in the usual manner by combining data for both school years. This will give a figure which represents the value of the indicator in a hypothetical population where an equal proportion of Reception and Year 6 children were measured.

Do I need to take account of the sex ratio of children measured?

7.6 Obesity and overweight prevalence for children is known to vary by sex. Users of the NCMP dataset may wish to further investigate differences by sex within their local area.

7.7 The 2006/07 NCMP report noted differences between the sex ratios of children measured in different areas. This pattern appears to have continued in more recent years and such differences remain in the 2012/13 datasets. In addition, at least at national level, more boys participate than girls. Although this appears to have a minimal effect on local level prevalence figures, for smaller populations the possible impact of a skewed sex ratio is greater. Users should be aware of this issue and, if prevalence figures are to be compared for boys and girls combined, it should first be ensured that there are no large differences in sex ratio between the populations being examined.

7.8 This issue is likely to be particularly important at school level. Comparing prevalence of obesity at a single sex school with a prevalence figure for the local authority or region that includes girls and boys would not be appropriate. This issue needs consideration, especially if feeding back results to schools.

How can I check to see whether a change or difference in prevalence is meaningful?

7.9 Comparison of prevalence figures with the regional or national rate, between different populations or over time should always take into account the degree of uncertainty around these figures.

7.10 The HSCIC report and published data tables provide approximate confidence limits for local authority prevalence rates. The method used to test for statistical significance is detailed in Annex 2 of the HSCIC report.

7.11 If users of the NCMP dataset want to calculate their own confidence limits for other geographic areas, or if they need to produce confidence limits for prevalence by sex, the 'Wilson Score' method is recommended. See Appendix 4.

7.12 If examining a reported change in rate for statistical significance, the approach recommended by Altman et al. should be used. See Appendix 4, Section A4.5.

What are the best methods for comparing the data for children measured in Reception to Year 6 measurements of the same children?

7.13 In some areas it will be possible to compare the 2012/13 NCMP data for Year 6 children with child measurements for the same children taken in the Reception year during 2006/07. Some local areas may even be able perform such analysis with earlier NCMP datasets – for example where the height and weight Reception year children was routinely measured and recorded prior to the National Childhood Obesity Dataset (NCOD) or NCMP.

7.14 Although it is interesting to compare change in these cohorts of children at a population level (for example comparing obesity prevalence in 2006/07 for children of Reception year with obesity prevalence in 2012/13 for Year 6), such analysis is unlikely to produce a great deal of new information beyond what is already known (that is obesity prevalence increases between the Reception and Year 6).

7.15 Of greater interest would be a comparison of children's BMI at an individual level. Tracking individual children over time opens up the possibility for new avenues of analysis.

7.16 However, such analysis requires individual children to be identified within the dataset so that they can be tracked from Reception to Year 6. As the NCMP dataset supplied by the HSCIC does not contain identifiable data, it is likely that such work can only be performed where child measurements are recorded onto a local system, such as a child health system, and can be extracted in an identifiable form[b].

7.17 Where such data are available, the sort of analysis that is likely to be most productive includes the change in BMI classification over time and the change in BMI z score or centile. If the number of children who can be identified in both school years is large enough it may be possible to perform analysis to investigate whether the change in children's BMI over time differs by sex, area, ethnicity or socioeconomic group.

7.18 Tools such as LMS Growth (see Appendix 3) are likely to prove helpful to anyone performing such analysis – this software can be used to assign BMI z scores, centiles, or BMI classification based on child measurements, which will

---

[b] The Pupil_ID and Local_ID fields within the NCMP MS Access databases are individual row identifiers for each NCMP year and are only unique within each separate dataset. The same pupil identifiers are re-used for different individuals in each year's dataset, and therefore cannot be used to match the same pupil measured in Reception to their Year 6 measurement.

be needed if child measurements are being extracted from a local data source and not the national NCMP dataset. (See A3.11 – A3.12 for further details).

7.19 As PHE does not have access to such data, it is not possible to trial such forms of analysis centrally. This limits the advice that can be provided to local areas. However, some examples of such work have been published, by Hull PCT[13] and Southampton PCT.[14] Local areas who intend to perform such analysis might wish to review what has been done in these areas before deciding on a plan for individual level analysis.

7.20 If any local areas wish to conduct analysis of this sort, PHE would be happy to discuss suitable approaches and methods prior to commencing the work. PHE is also keen to receive links to, or electronic copies of, any local publications which present the results of this sort of analysis, which can then be shared with other areas.

Appendix 1: NCMP 2012/13 data fields

A1.1 The 2012/13 NCMP dataset is supplied to PHE by the HSCIC in a single Access database. The database contains three data tables which contain information at pupil (for both valid and excluded records), school, and PCT.

A1.2 Twelve additional tables are provided which allow users to assign descriptions to the coding used within the dataset. These lookup tables cover BMI classification (clinical), BMI classification (population monitoring), ethnicity codes, local authority codes, ONS Area Classification Group codes, ONS AC Subgroup codes, ONS AC Supergroup codes, removal criteria (description of the criteria under which excluded records were removed), school establishment type and Urban/Rural classification.

A1.3 This table and field structure is now applicable to the most recent versions of NCMP datasets for the period 2006/07 – 2011/12, although some fields are not available for older datasets. The field names and descriptions for the three data tables are outlined below:

Table 1: Pupil_data and Pupil_data_excluded

Field Name	Field Description
Pupil_ID	Unique ID code for each pupil
Local_ID	PCT ID code for each pupil
Exclude_flag	Pupils attending independent and special schools are flagged as 1 (these records need to be excluded to match the published figures)
Sex	Sex of pupil
Age	Age of pupil (in months)
School_yr	School Year of pupil - derived from child age (R: Reception, 6: Year 6)
Ethnicity	Ethnicity recoded to NHS classification (see table Ethnicity_codes)
Ethnicity_full	Ethnicity code as entered by PCT
Height	Height of pupil (in cm)
Height_z	Height z score - derived from British 1990 growth reference, using Age, Sex and Height fields
Height_p	Height centile - derived from British 1990 growth reference, using Age, Sex and Height fields
Weight	Weight of pupil (in kg)
Weight_z	Weight z score - derived from British 1990 growth reference, using Age, Sex and Weight fields
Weight_p	Weight centile - derived from British 1990 growth reference, using Age, Sex and Weight fields
BMI	BMI of pupil in kg/m^2 - derived from height and weight fields
BMI_z	BMI z score - derived from British 1990 growth reference, using Age, Sex and BMI fields
BMI_p	BMI centile - derived from British 1990 growth reference, using Age, Sex and BMI fields

BMI_class_pop_mon	BMI classification to UK90 population monitoring centiles (85th/95th centiles for overweight and obese), plus 2nd centile for underweight (See table BMI_class_pop_mon)
BMI_class_clinical	BMI classification to UK90 clinical centiles (approx 91st/98th centiles for overweight and obese), plus 2nd centile for underweight (See table BMI_class_clinical)
Month_meas	Month of measurement (1: January, 2: February etc)
DayofMeasurement	Day in week measurement took place (1: Sunday, 2: Monday, 3: Tuesday, 4: Wednesday, 5: Thursday, 6: Friday, 7: Saturday)
PCT	PCT that submitted measurement (see table PCT_data) based on PCT codes as at 1st April 2011
SHA	SHA - based on PCT that submitted measurement (see table SHA_codes)
Schl_LA_Tier1	Local Authority (Tier 1) - derived from postcode of school (see table LA_codes_Tier1)
Schl_LA_Tier2	Local Authority (Tier 2) - derived from postcode of school (see table LA_codes_Tier2)
Schl_GOR	Government Office Region - derived from postcode of school (see table GOR_codes)
Schl_Supergroup_ONS_AC	ONS Area Classification Supergroup code - derived from school postcode (see table ONS_AC_Supergroup)
Schl_Group_ONS_AC	ONS Area Classification Group code - derived from school postcode (see table ONS_AC_Group)
Schl_Subgroup_ONS_AC	ONS Area Classification Subgroup code - derived from school postcode (see table ONS_AC_Subgroup)
Schl_LSOA	Lower Super Output Area of school (2001) - derived from postcode of school
Schl_Urban	ONS Urban/ Rural classification - derived from school postcode (see table Urban_rural)
Schl_IMD_Decile	The decile that the Index of Multiple Deprivation (IMD) 2010 score falls into - derived from LSOA of school (0: not available, 1: 1st decile (least deprived),..., 10: 10th decile (most deprived))
Child_LA_Tier1	Local Authority (Tier 1) - derived from postcode of child (see table LA_codes_Tier1)
Child_LA_Tier2	Local Authority (Tier 2) - derived from postcode of child (see table LA_codes_Tier2)
Child_GOR	Government Office Region - derived from child postcode (see table GOR_codes)
Child_MSOA	Middle Super Output Area (2001) - derived from postcode of child
Child_Supergroup_ONS_AC	ONS Area Classification Supergroup code - derived from child postcode (see table ONS_AC_Supergroup)
Child_Group_ONS_AC	ONS Area Classification Group code - derived from child postcode (see table ONS_AC_Group)
Child_Subgroup_ONS_AC	ONS Area Classification Subgroup code - derived from child postcode (see table ONS_AC_Subgroup)
Child_LSOA	Lower Super Output Area of child (2001) - derived from child postcode
Child_Urban	ONS Urban/ Rural classification - derived from child postcode (see table Urban_rural)
Child_IMD_Decile	The decile that the Index of Multiple Deprivation (IMD) 2010 score falls into - derived from LSOA of child (0: not available, 1: 1st decile (least deprived),..., 10: 10th decile (most deprived))
School_child_distance	The straight line distance between the school postcode and child postcode (in km). Please note the child postcode, and all fields derived from it, have been deleted if this distance is greater than 60km, as the child postcode was thought to be erroneous
removal_criteria	Removal criteria (see table removal_criteria)
URN	DfE Unique Reference Number for school (see table School_data)

Table 2: PCT_data

Field Name	Field Description
PCT	PCT code based on PCT codes from 1st April 2011 onwards
PCT_name	PCT name
SHA	SHA code of PCT
PCT_Meas_R	Number of pupils measured - Reception (used for published participation rates)
PCT_Meas_6	Number of pupils measured - Year 6 (used for published participation rates)
PCT_Eligible_R	Eligible Reception year pupils within PCT (derived from sum of school level eligible figures)
PCT_Eligible_6	Eligible Year 6 pupils within PCT (derived from sum of school level eligible figures)
PCT_Particip_R	Participation rate - Reception
PCT_Particip_6	Participation rate - Year 6
Storage_R	Mode of data storage - Reception
Storage_6	Mode of data storage - Year 6
Parentopt_R	Number of pupils not measured due to parental opt-out – Reception
Parentopt_6	Number of pupils not measured due to parental opt-out – Year 6
Childopt_R	Number of pupils not measured due to child opt-out – Reception
Childopt_6	Number of pupils not measured due to child opt-out – Year 6
Unable_R	Number of pupils not measured because child unable to stand on scales - Reception
Unable_6	Number of pupils not measured because child unable to stand on scales - Year 6
Absent_R	Number of pupils not measured because child absent on day of measurement - Reception
Absent_6	Number of pupils not measured because child absent on day of measurement - Year 6
Schl_opt_R	Number of pupils not measured because school opted out of measurement - Reception
Schl_opt_6	Number of pupils not measured because school opted out of measurement - Year 6
nSchl_opt_R	Number of schools which opted out of measurement – Reception
nSchl_opt_6	Number of schools which opted out of measurement – Year 6
Other_R	Number of children not measured for other reasons – Reception
Other_6	Number of children not measured for other reasons – Year 6

Table 3: School_data

Field Name	Field Description
URN	DfE Unique Reference Number for school
Schl_name	School name
Schl_Addr1	School address line 1
Schl_Addr2	School address line 2
Schl_Addr3	School address line 3
Schl_Addr4	School address line 4
Schl_Addr5	School address line 5
Schl_pcode	School postcode
Schl_type	School establishment type
PCT	PCT responsible for taking NCMP measurements
Particip_schl	Schools that participated in NCMP 2012/13 are flagged as 1
Optional_schl	Schools where NCMP measurements are optional, i.e. independent and special schools, are flagged as 1
Schl_LSOA	Lower Super Output Area of school (2001) - derived from school postcode
Schl_LA	Local Authority Tier 1 - derived from postcode of school (see table LA_codes_Tier1)
Schl_GOR	Government Office Region - derived from postcode of school (see table GOR codes)
Schl_Supergroup_ONS_AC	ONS Area Classification Supergroup code - derived from school postcode (see table ONS_AC_Supergroup)
Schl_Group_ONS_AC	ONS Area Classification Group code - derived from school postcode (see table ONS_AC_Group)
Schl_Subgroup_ONS_AC	ONS Area Classification Subgroup code - derived from school postcode (see table ONS_AC_Subgroup)
Schl_urban	ONS Urban/ Rural classification - derived from postcode of school (see table Urban_rural)
Schl_IMD_Decile	The decile that the Index of Multiple Deprivation (IMD) 2010 score falls into - derived from LSOA of school (0: not available, 1: 1st decile (least deprived),..., 10: 10th decile(most deprived))
Schl_Eligible_R	Eligible pupils figure supplied by PCT. Adjusted so not exceeded by number of pupils measured - Reception
Schl_Eligible_6	Eligible pupils figure supplied by PCT. Adjusted so not exceeded by number of pupils measured - Year 6

Appendix 2: Suggested data quality checks at local level

A2.1 Although the NCMP dataset provided to PHE has undergone extensive cleaning at national level, there is a limit on the checks and cleaning that can be done centrally on a dataset with over one million records from around 17,000 schools. As a result there may be some minor remaining data quality issues within the NCMP dataset.

A2.2 In earlier NCMP datasets a number of issues were identified during analysis. These included: duplicate records in the dataset; the same pupil records entered for two adjacent schools; Year 6 pupils entered to infant schools or Reception pupils entered to junior schools and a large proportion of records were found to have height and weight measurements rounded to the nearest whole number.

A2.3 The NCMP upload process is continually being improved to include additional validation checks, and further validation is now done by the HSCIC prior to the data being released. This process is described in the HSCIC's annual NCMP report. However, some data quality issues may remain despite these additional checks.

A2.4 These data quality issues have minor effects on national analyses but may be more important in detailed regional or local analyses by PHE or local authorities. It is therefore important that basic quality checks on the dataset are performed and any anomalies are clarified with the relevant local authorities or with the staff involved in collecting and processing measurements. Please report any errors or issues found from any data quality checks to PHE (ncmp@phe.gov.uk).

A2.5 Users of the NCMP dataset may wish to check for some or all of the following issues before commencing detailed analysis.

   A2.5.1   **Records assigned to the wrong school:** In previous NCMP datasets some child records have been found to be coded to the wrong school. This issue could often only be easily identified in the most obvious cases, such as where infant schools had Year 6 pupils coded to them and where Reception pupils were coded to junior schools. As a result, the true scale of this issue is unknown.

A2.5.2   In many cases this miscoding seems to have occurred where schools share similar names (for example St Mary's Infants and St Mary's Junior) and all records for both schools have been assigned to one of the two institutions.

A2.5.3   The NCMP upload process includes checks to warn local authorities where such miscoding may have occurred. For example, local authorities are warned of the number of schools for which no records are entered and also of the number of schools where the number of pupils measured exceeds the number of pupils reported to be at the school. However, it is still possible that some incorrect school coding may have occurred within the 2012/13 dataset.

A2.5.4   **Duplicate pupils:** Duplicate pupils assigned to a school should have been flagged during the upload process and removed before data were submitted to the NCMP database.

A2.5.5   The HSCIC has also removed any duplicate records identified during data cleaning from the dataset. These suspected duplicate records are included in the NCMP Access database in the Pupil_data_excluded table under 'Removal_criteria' codes 8 and 9. Local areas may wish to examine how many local records have been identified as duplicates. If a substantial number of local child records have been removed for this reason, this information should be fed back to the team responsible for data collection and processing.

A2.5.6   In some situations (for example where duplicate records were submitted to separate schools or where records were actually part of a set of duplicates but contain differing data for a given field) some duplicate records may remain in the dataset. It is therefore worthwhile for local areas to perform some quick checks to see if any such records can be identified.

A2.5.7   Detailed analyses of previous NCMP datasets showed that previously a few PCTs had submitted the same set of records for more than one school. Often this occurred where pupils had been inaccurately coded to schools. For example, in some cases a group of Reception year pupils had been incorrectly added to a similarly named junior school as well as to the correct infant school.

A2.5.8   Although the NCMP dataset is anonymised, it is possible to detect potential duplicate records by matching on fields such as age, sex, height and weight.

A2.5.9   **Rounded records:** Analyses of previous NCMP datasets has shown that incorrectly rounded records (especially those for weight in the Reception year) were associated with a lower reported prevalence of obesity.[15]

A2.5.10   The HSCIC run a validation check for rounded records during the data upload process, but as this warns local authorities of rounded records only after data have been submitted, some local authorities still have a high proportion of rounded records in the 2012/13 dataset if they did not make amendments and resubmit.

A2.5.11   A summary of the proportion of rounded records for every local authority has been provided within the HSCIC's 2012/13 NCMP report (Annex 1), and on the Local Authority NCMP Profile http://fingertips.phe.org.uk/profile/national-child-measurement-programme. Users of the NCMP data at local level are advised to check this list and, if the records for the population being studied have been submitted with rounded heights and/or weights, this issue may need to be taken into consideration when using prevalence figures for those areas. In addition it may prove beneficial to work with the teams responsible for collecting and entering NCMP data to ensure that data are entered to the correct level of precision (1 decimal place) in future years.

A2.5.12   **Day of measurement:** The NCMP dataset contains a field showing day in the week of measurement within the 'Pupil_data' table.

A2.5.13   Analysis of this field has shown that a small proportion of child records are submitted with a date of measurement that corresponds to a Saturday or Sunday. It seems likely that this has resulted where date of measurement has been incorrectly entered for that child.

A2.5.14   Local areas may wish to investigate the proportion of local child measurements which appear to have taken place at a weekend, and feed this information back to the team who collects and processes data if this is found to be occurring locally.

Appendix 3: BMI thresholds, z scores, and p scores

A3.1 The height, weight and BMI of children change as children grow, and also vary between boys and girls. In order to determine whether any individual child's measurements should be considered too low or too high, the child's height, weight or BMI must be compared to a child growth reference. Such references describe the expected pattern of growth for children at different ages and by sex, and are usually based on a relatively healthy historic population (that is one with low obesity prevalence).

A3.2 A child growth reference can be used to convert the height, weight or BMI measurements of individual children into standard deviation scores (z scores) or centiles (p scores). These z scores describe whether the child has a higher or lower value for that measure than would be expected of children of the same age and sex.

A3.3 For example, a child with a BMI z score of 0 (which equates to the 50th centile) has a BMI the same as the average value for children of the same age and sex in the 1990 reference population. A child with a BMI z score of +1.64 (the 95th centile) has a BMI that is higher than 95% of children of the same age and sex in the 1990 reference population.

BMI thresholds used with the NCMP dataset

A3.4 The NCMP published prevalence data use the British 1990 growth reference (UK90) for BMI and the 2nd, 85th and 95th centiles to define children as underweight, overweight or obese according to age and sex. This definition is the most commonly used in England for population monitoring – for example in Health Survey for England (HSE) figures.

A3.5 It is important to note that the 85th and 95th centiles used in the NCMP are intended for population monitoring use only, and do not provide the number or percentage of individual children clinically defined as overweight or obese.

A3.6 In clinical settings or when monitoring the BMI of individual children, the 2nd, 91st and 98th centiles of the UK90 reference are used in the UK to classify the BMI of individual children as underweight, healthy, overweight or obese taking into account the expected variation in BMI by age and sex. The NCMP

37

parental feedback letters issued by PCTs use these clinical cut-offs to assign children to a BMI classification.

A3.7 It is important to note that the clinical cut-offs for child underweight, overweight, and obesity are in fact set at -2 (-6/3), +4/3, and +2 (+6/3) standard deviations. These actually equate to the 2.3rd, 90.9th, and 97.7th centiles, when rounded to one decimal place, although they are usually referred to as the 2nd, 91st, and 98th centiles.

A3.8 Underweight prevalence figures using both population monitoring and clinical thresholds are each presented as having been derived using the 2nd centile. However only those based on population monitoring thresholds use the 2nd centile, whereas those derived using clinical thresholds actually use -2 standard deviations (the 2.3rd centile). For a given population, prevalence of underweight will therefore differ slightly depending on whether the population monitoring or clinical cut offs are used.

A3.9 When presenting prevalence figures based on the 85th and 95th centile cut-offs, or any other BMI thresholds, it is important to state the thresholds and growth reference being used to ensure valid comparisons can be made between the figures being presented and those from other sources.

A3.10   If for any reason users need to calculate their own BMI z scores for NCMP or other data, this can be done quickly and easily using the 'LMS Growth' Microsoft Excel add-in software. This software is available at no charge from Harlow Publishing at http://www.healthforallchildren.com/?product=lmsgrowth

A3.11   LMS Growth can be used to calculate both z scores and centile for child measurements. The two can also be converted within MS Excel, using the 'normsinv' and 'normsdist' commands.

A3.12   There are very slight differences between this Excel add-in and the approach used to assign BMI z scores and centiles in the NCMP dataset. These relate to the precise method used to allocate L, M and S variables to individuals. Hence, the resulting BMI, height and weight z scores assigned may differ by a small amount. However these differences do not have any noticeable impact on prevalence figures.

## Other thresholds for defining children's BMI status

A3.13   Users of the NCMP dataset should also note that other growth references are sometimes used to classify children as overweight or obese. For example,

the World Health Organization (WHO) 2007 or International Obesity Task Force (IOTF) thresholds are sometimes used in the UK. The IOTF thresholds were used in the Foresight obesity modelling[16] and for child obesity prevalence figures from the Millennium Cohort Study.[17]

A3.14   Although these alternative growth references show a broadly similar pattern of changing BMI with age, they produce very different thresholds for underweight, healthy weight, overweight and obesity.

A3.15   Most published NCMP analyses use the recommended UK90 population monitoring thresholds to ensure consistency between published figures. If making comparisons with other published prevalence figures, the same definition of obesity, overweight and underweight must be applied across all figures. Prevalence figures that use different references or thresholds cannot be compared directly.

Figure 5: Obesity and overweight thresholds for boys – UK90 and IOTF
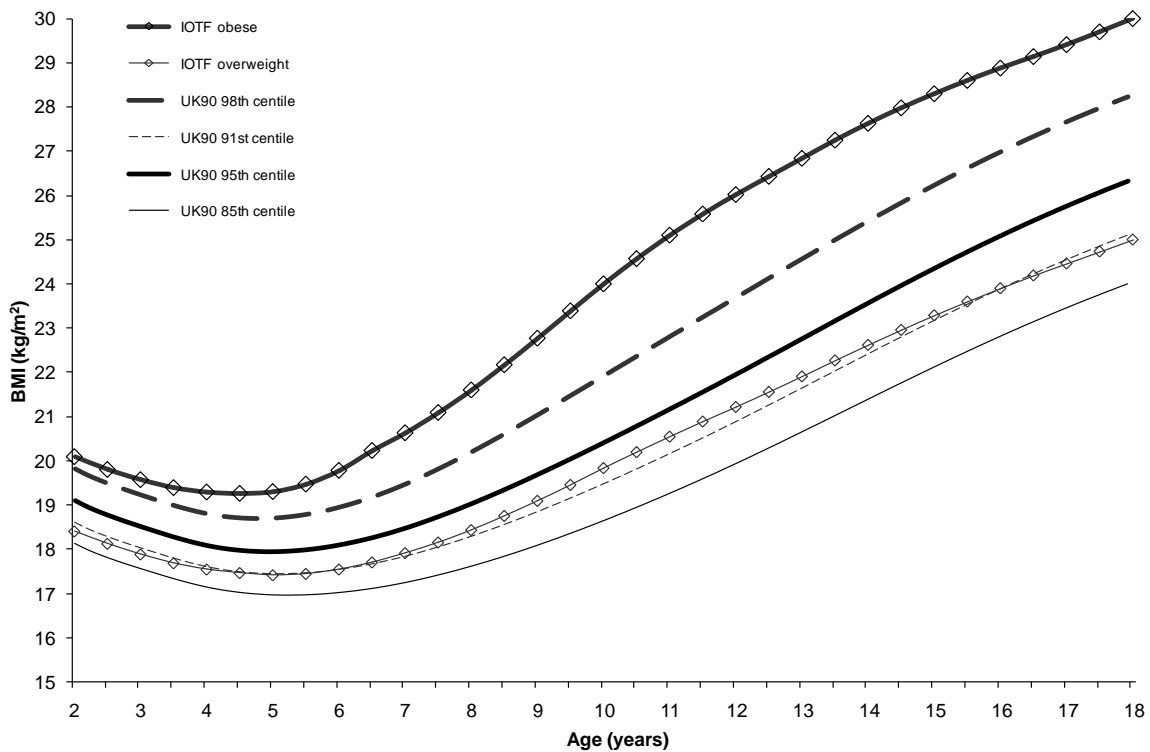
Appendix 4: Methods for confidence limits

A4.1 We recommend that 95% confidence intervals are calculated with the method described by Wilson[18] and Newcombe[19] which is a good approximation of the exact method.

A4.2 The estimated proportions of children with and without the feature of interest were calculated:

observed number of obese children in each area =r
sample size = n
proportion with feature of interest = p = r/n
proportion without feature of interest = q = (1 – p)

A4.3 Three values (A, B and C) were then calculated as follows:

$$A = 2r + z^2; B = z\sqrt{z^2 + 4rq} \text{ ;and } C = 2(n+z^2)$$

where z is the appropriate value, $z_{1-\alpha/2}$, from the standard Normal distribution. Then the confidence interval for the population proportion is given by

(A-B)/C to (A+B)/C

This method is superior to other approaches because it can be used for any data. When there are no observed events, then r and hence p are both zero, and the recommended confidence interval simplifies to *0* to $z^2/(n+z^2)$. When $r = n$ so that
$p = 1$, the interval becomes $n/(n+z^2)$ to *1*.

A4.4 When testing for differences between rates or proportions it is important to use an appropriate statistical test rather than just looking at whether confidence intervals are overlapping. In some circumstances there could be a statistically significant difference even where confidence intervals overlap.

A4.5 The approach outlined by Altman et al. in *Statistics with Confidence (edition 2)*[20] is recommended for such statistical testing.
Where the difference in two rates or proportions, $\hat{D} = \hat{p}_2 - \hat{p}_1$
has confidence limits from:

$$\hat{D} - \sqrt{(\hat{p}_2 - l_2)^2 + (u_1 - \hat{p}_1)^2} \text{ to } \hat{D} + \sqrt{(\hat{p}_1 - l_1)^2 + (u_2 - \hat{p}_2)^2}$$

Where $\hat{p}_i$ is the estimated prevalence for year i, and l_i and u_i are the lower and upper confidence intervals for $\hat{p}_i$ respectively.

A4.6 This method is also provided as 'method 10' in the Newcombe paper 'interval estimation for the difference between independent proportions: comparison of eleven methods'.[21]

References

1. Department of Health. Healthy lives, healthy people: our strategy for public health in England. London: Department of Health, November 2010 http://www.dh.gov.uk/en/Publicationsandstatistics/Publications/PublicationsPolicyAndGuidance/DH_121941.

2. Department of Health. Public Health Outcomes Framework. London: Department of Health, January 2012 http://www.dh.gov.uk/en/Publicationsandstatistics/Publications/PublicationsPolicyAndGuidance/DH_132358.

3. Ridler C, Dinsdale H, Rutter H. National Child Measurement Programme: Analysis using ONS Area Classification. Oxford: National Obesity Observatory, 2011 http://www.noo.org.uk/uploads/doc/vid_12524_NCMP_Analysis%20using%20the%20ONS%20Area%20Classification.pdf

4. Dinsdale H. Variation in children's BMI by month of measurement in the 2008/09 National Child Measurement Programme dataset. Oxford: National Obesity Observatory, 2010 http://www.noo.org.uk/uploads/doc/vid_8785_NCMP_Variation%20by%20month%20of%20measurement_v2.pdf.

5. The Local Authority (Public Health, Health and Wellbeing Boards and Health Scrutiny) Regulations 2013 http://www.legislation.gov.uk/uksi/2013/218/pdfs/uksi_20130218_en.pdf.

6. UK Data Protection Act 1998: http://www.opsi.gov.uk/Acts/acts1998/ukpga_19980029_en_1.

7. National Child Measurement Programme: NHS Information Centre validation process for NCMP data Leeds: HSCIC, July 2009 http://www.hscic.gov.uk/catalogue/PUB00760/nati-chil-meas-prog-eng-2008-2009-data.pdf.

8. Cole TJ, Freeman JV, Preece MA. British 1990 growth reference centiles for weight, height, body mass index and head circumference fitted by maximum penalized likelihood. Stat Med 1998; 17:407-29.

9. Cole TJ, Freeman JV, Preece MA. Bodymass index reference curves for the UK. Arch Dis Child1995; 73: 25-29.

10. Rudolf M, Levine R, Feltbower R, Connor A, Robinson M. The TRENDS Project: development of a methodology to reliably monitor the obesity epidemic in childhood. Arch Dis Child 2006; 91:309-911.

11. Hunt L, Ford A, Sabin A, Crowne E, Shield J. Clinical measures of adiposity and percentage fat loss: which measure most accurately reflects fat loss and what should we aim for? Arch Dis Child 2007; 92:399-403.

12. Dinsdale H, Ridler C. National Child Measurement Programme: Guidance for small area analysis. Oxford: National Obesity Observatory, 2011 http://www.noo.org.uk/uploads/doc/vid_11853_NCMP_Guidanceforsmallarea %20analysisFINAL.pdf.

13. Porter M, Greene T, Taylor A. Childhood obesity in Hull: paired analysis. Hull: Hull PCT, December 2007.

14. King D. Child Growth Briefing Note; National Child Measurement Programme 2001/02 to 2008/09 Summary Report. Southampton: Southampton PCT, 2010 http://www.southamptonhealth.nhs.uk/EasySiteWeb/GatewayLink.aspx?alId=9 7148.

15. Dinsdale H, Rutter H. National Child Measurement Programme: Detailed Analysis of the 2006/07 National Dataset. London: Department of Health, June 2008 http://www.noo.org.uk/uploads/doc168_2_NOO_NCMP_report230608.pdf.

16. McPherson K, Marsh T, Brown M. Tackling Obesities: Future Choices – Modelling Future Trends in Obesity & Their Impact on Health. London: Department of Innovation, Universities and Skills, 2007 http://www.bis.gov.uk/foresight/our-work/projects/current-projects/tackling-obesities.

17. Hanson K, Joshi H (eds). Millennium Cohort Study Third Survey: A User's Guide to Initial Findings. London: Institute of Education, October 2008 http://eprints.ioe.ac.uk/5931/1/MCS_3_Descriptive_Report_Oct_2008.pdf.

18. Wilson EB. Probable inference, the law of succession, and statistical inference.J Am Stat Assoc 1927; 22:209-12.

19. Newcombe RG. Two-sided confidence intervals for the single proportion: comparison of seven methods. Stat Med 1998; 17:857-72.

20. Altman DG, Machin D, Bryant TN, Gardner MJ.Statistics with Confidence, 2nd edn. London; BMJ books, 2000: 49.

21. Newcombe RG, Interval estimation for the difference between independent proportions: comparison of eleven methods. Stat Med 1998; 17:873-90.

All weblinks accessed 21/03/2014

Reader information

Title	NCMP Guidance for data sharing and analysis
Author(s)	Caroline Hancock, Hywell Dinsdale
Reviewer(s)	Paul Eastwood, Health and Social Care Information Centre
	Shireen Mathrani, Public Health England
	Rosie Taylor, Public Health England
Publication date	July 2014
Target audience	Local authorities, Public Health England Knowledge and Intelligence Teams.
Description	Information and guidance for data sharing and analysis by local authorities and Public Health England of the 2012/13 National Child Measurement Programme dataset.
How to cite	Hancock C, Dinsdale H. NCMP Guidance for data sharing and analysis. Oxford: Public Health England Obesity Knowledge and Intelligence, 2014.
Contact	PHE Obesity Knowledge and Intelligence
	www.noo.org.uk
	ncmp@phe.gov.uk
	@PHE_obesity
Electronic location	http://www.noo.org.uk/NCMP/analytical_guidance
Copyright	© Crown copyright 2014