



Public Health  
England

Protecting and improving the nation's health

# **National Child Measurement Programme** guidance for data sharing and analysis

## About Public Health England

Public Health England exists to protect and improve the nation's health and wellbeing, and reduce health inequalities. It does this through world-class science, knowledge and intelligence, advocacy, partnerships and the delivery of specialist public health services. PHE is an operationally autonomous executive agency of the Department of Health.

Public Health England

Wellington House

133-155 Waterloo Road

London SE1 8UG

Tel: 020 7654 8000

[www.gov.uk/phe](http://www.gov.uk/phe)

Twitter: [@PHE\\_uk](https://twitter.com/PHE_uk)

Facebook: [www.facebook.com/PublicHealthEngland](https://www.facebook.com/PublicHealthEngland)



© Crown copyright 2016

You may re-use this information (excluding logos) free of charge in any format or medium, under the terms of the Open Government Licence v3.0. To view this licence, visit [OGL](https://www.ogil.io) or email [psi@nationalarchives.gsi.gov.uk](mailto:psi@nationalarchives.gsi.gov.uk). Where we have identified any third party copyright information you will need to obtain permission from the copyright holders concerned.

Any enquiries regarding this publication should be sent to: [ncmp@phe.gov.uk](mailto:ncmp@phe.gov.uk)

Published June 2016

PHE publications gateway number: 2016095



# Contents

|  |    |
|--|----|
| About Public Health England                              | 2  |
| Contents   | 2  |
| 1. Introduction  | 4  |
| 2. National reports and resources                        | 6  |
| 3. Obtaining and sharing the NCMP dataset                | 9  |
| 4. The enhanced NCMP dataset                             | 13 |
| 5. Essentials for NCMP analysis                          | 17 |
| 6. Suggested analysis at local level                     | 18 |
| 7. Analysis FAQ  | 25 |
| Appendix 1: NCMP 2014/15 enhanced extract data fields    | 29 |
| Appendix 2: Suggested data quality checks at local level | 32 |
| Appendix 3: BMI thresholds, z scores, and p scores       | 35 |
| Appendix 4: Methods for confidence limits                | 38 |

# 1. Introduction

- 1.1 The National Child Measurement Programme (NCMP) is an annual programme that measures the height and weight of children in Reception (age four to five years) and Year 6 (age ten to 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 4.5).
- 1.2 The Health and Social Care Information Centre (HSCIC) and Public Health England (PHE) publish analyses of this dataset annually, however, local authorities may want to undertake additional analyses to inform their work to tackle obesity.
- 1.3 The 2010 Department of Health (DH) obesity strategy Healthy lives, healthy people: a call to action on obesity in England<sup>1</sup> and the Public Health Outcomes Framework (PHOF)<sup>2</sup> outline the government's ambition to achieve a sustained downward trend in the prevalence of child excess weight. The PHOF uses NCMP data to help local authorities monitor progress towards this ambition [www.phoutcomes.info](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 2014/15 NCMP. The HSCIC launched a new online NCMP data collection system in 2013/14 and local authorities are now able to access an enhanced analysis dataset directly from this system following the collection period. This extract is available until the HSCIC needs to purge the system to begin another collection year. Local authorities will be made aware of when this will happen so they have sufficient time to download the extract if they have not already done so. PHE continues to receive an analysis dataset directly from the HSCIC which is governed by a

Memorandum of Understanding on data sharing between the two organisations.

- 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](mailto:ncmp@phe.gov.uk).

## 2. National reports and resources

2.1 A number of resources to facilitate local use of NCMP data are produced each year by both the HSCIC and PHE. These consist of reports, data slide sets presenting and describing the NCMP data at national level and also data tables at national, regional, local authority, ward 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 2014/15 NCMP at national and sub-national level are provided in the HSCIC's report; National Child Measurement Programme: England, 2014/15 school year, published in November 2015.

[www.hscic.gov.uk/catalogue/PUB19109](http://www.hscic.gov.uk/catalogue/PUB19109)

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 this report. These contain data on obesity, overweight, healthy weight and underweight prevalence for regions and local authorities (based on location of school, area of residence and submitting authority) for 2014/15. Additional national level data are included broken down by ethnic group, rural-urban classification and Index of Multiple Deprivation (IMD) decile.

2.5 The HSCIC also produces an extract of record level NCMP data with an associated guidance document to accompany the annual report. Some of the data items collected as part of the NCMP are removed from this extract and others are altered to mitigate against the possibility of a child being identified. An equivalent file for the 2006/07 to 2012/13 NCMP years is available from the UK Data Archive (UKDA) at [www.data-archive.ac.uk](http://www.data-archive.ac.uk).

2.6 The 2014/15 HSCIC NCMP report, accompanying data tables and anonymised record level extract (including guidance on how to use it) are available from [www.hscic.gov.uk/catalogue/PUB19109](http://www.hscic.gov.uk/catalogue/PUB19109).

## NCMP data and analysis provided by PHE

### PHE NCMP analysis reports

- 2.7 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 2016 to complement the HSCIC's 2014/15 report.
- 2.8 The 2008/09 to 2013/14 analysis reports, along with the 2014/15 report (when published), 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 observed patterns and trends.
- 2.9 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 with the NCMP database for those years and describe the possible effect of such issues on reported prevalence of overweight and obesity.
- 2.10 Previous reports published using NCMP data include short reports describing the use of the ONS Area Classification (ONS-AC) (see sections 4.14 and 6.14) in NCMP analysis,<sup>3</sup> and the impact of month of measurement on reported obesity prevalence within the NCMP dataset.<sup>4</sup>
- 2.11 All NCMP analysis reports published by PHE are available on the PHE Obesity website: [www.noo.org.uk/NCMP](http://www.noo.org.uk/NCMP).

### NCMP local authority profile

- 2.12 An online data tool displays data from the NCMP 2006/07 to 2014/15. Prevalence of underweight, healthy weight, overweight, and obesity for children in Reception and Year 6 can be examined at local authority level. New data showing prevalence of obesity by deprivation quintile and ethnic group for each local authority were added in February 2016. Data quality indicators are also available in the tool, for example rate of participation in the NCMP. This online tool can be accessed at:  
<http://fingertips.phe.org.uk/profile/national-child-measurement-programme>.

## Other PHE NCMP data publications

2.13 Spreadsheets containing NCMP obesity prevalence data for Reception (age four to five years) and Year 6 (age ten to 11 years) children by 2014 electoral ward and 2011 MSOA of residence is available on the PHE obesity website [www.noo.org.uk/visualisation](http://www.noo.org.uk/visualisation). Local authority, regional and England data are also presented for comparison. These small area prevalence figures use three years of NCMP combined. NCMP data at MSOA level are also included in the PHE Local Health tool: [www.localhealth.org.uk](http://www.localhealth.org.uk).

2.14 PowerPoint slides presenting key data and information on child obesity in clear, easy to understand charts and graphics using NCMP and Health Survey for England (HSE) data are available on the PHE obesity website. These can be used freely with acknowledgement to Public Health England. [www.noo.org.uk/slide\\_sets](http://www.noo.org.uk/slide_sets).

2.15 NCMP Operational Guidance provides more information on the NCMP and advises local commissioners and providers of the NCMP on its implementation. It is available on the PHE website: [www.gov.uk/government/publications/national-child-measurement-programme-operational-guidance](http://www.gov.uk/government/publications/national-child-measurement-programme-operational-guidance).

## 3. Obtaining and sharing the NCMP dataset

### Obtaining the data

- 3.1 The HSCIC launched a new online NCMP data collection system in 2013/14 and local authorities are now able to access an enhanced analysis dataset directly from this system following the collection period. It is important to note that the NCMP data collection system is not a data storage facility so local authorities must make arrangements to download their enhanced analysis dataset and store it locally in line with their own Information Governance policies. **The enhanced dataset is generally available for a few months after the end of the collection period before the system is purged for the following collection year.** The HSCIC emails users of the system to let them know when the dataset is available and also to warn them before the system is purged.
- 3.2 Note that local authorities will not be able to access the enhanced dataset from the NCMP IT system after it has been purged. If enhanced data are required after the system purge then local authorities should contact the HSCIC ([enquiries@hscic.gov.uk](mailto:enquiries@hscic.gov.uk)). The HSCIC will no longer have access to any Personal Identifiable Data (PID) at this point as this is removed as part of the system purge to adhere to the Information Security Policy for the NCMP data collection system. Therefore local authorities who need PID within their extract must download it before the system is purged.<sup>a</sup>
- 3.3 The enhanced dataset contains all the data originally provided by the local authority along with mapped geographic and deprivation-based fields and derived fields such as child age. It also contains information on any data quality issues identified by the HSCIC.
- 3.4 Although local authorities will already have access to their own NCMP data, it is recommended that the enhanced 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 to retain consistency with published figures. The HSCIC dataset also has a greater number of fields than locally held data which facilitates a greater range of

---

<sup>a</sup> Access to the PID within the HSCIC is restricted to a small group of analysts who either deal with local authority queries when submitting NCMP data, or replace the PID with a pseudo identifier to be used to link the measurements taken in reception with those subsequently taken in Year 6. Once this work is complete all PID are removed from the system as the HSCIC no longer has a need for it.

possible analyses and removes the need for local authorities to derive these additional data items themselves.

- 3.5 Further information on the HSCIC NCMP IT system and guidance on how to obtain the enhanced data from the system is given at [www.hscic.gov.uk/ncmpsystem](http://www.hscic.gov.uk/ncmpsystem).
- 3.6 Local authorities should contact the HSCIC in the first instance with requests for historical NCMP data. Local authorities will only be sent historic NCMP data for children they have measured, or who were measured by their predecessor organisations. Note that the HSCIC will no longer have access to PID for these earlier years as explained in section 3.2. The HSCIC may levy a charge for meeting these data requests on a cost recovery basis.

### Sharing data with other organisations

- 3.7 Other organisations such as the NHS and academic institutions may request NCMP data from local authorities directly if they only require data for small areas, or from the HSCIC if they require data for a wider area or at national level, however, there are restrictions over what data can be shared and this is outlined in sections 3.11 to 3.12 below.
- 3.8 Often the information needed will already be available through the HSCIC and PHE websites. Where possible users should be directed to these existing resources.
- 3.9 In some cases the information requested will not be publicly available and therefore will need to be produced using individual row level dataset. Local authorities are able to provide such information, but this may need to be in an aggregated and suppressed form. This ensures individual information is not shared unnecessarily and prevents the identification of individual children.

### Processing of information by local authorities

- 3.10 The NCMP regulations<sup>5</sup> state that information collected for the NCMP ‘may be further processed by or on behalf of the local authority which provided it, with a view to disclosing such information to any person to be used for the purposes of research, monitoring, audit or the planning of services, or for any purpose connected with 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 NCMP regulations<sup>5</sup> state that local authorities (or those working on behalf of the local authority) are only 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 NCMP regulations<sup>5</sup> state that the HSCIC 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.' Typically this may involve providing data in an aggregated and suppressed form in line with the NHS Anonymisation Standard: [www.isb.nhs.uk/documents/isb-1523/amd-20-2010/index\\_html](http://www.isb.nhs.uk/documents/isb-1523/amd-20-2010/index_html). The HSCIC will also apply charges for providing data on a cost-recovery basis.

## Feeding back NCMP data to schools

3.13 Many schools are keen to receive detailed feedback on 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 single year NCMP statistics such as obesity prevalence at a school level. 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 – this will bias the results and make them less reliable at school level
- cohort 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 sections 7.5 to 7.7). Therefore, they would 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 authorities wish to conduct more detailed analysis of obesity prevalence at school level, this may be done in a number of ways. Data from more than one measurement year could 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.
- 3.16 While 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 presented as a measure of whether a school provides a 'healthy environment'. 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 PHE has developed new resources to support local authorities to share NCMP information with schools. These include an individualised school feedback letter which became available in summer 2015.
- 3.18 School-age children health profiles for local authorities are available at <http://atlas.chimat.org.uk/IAS/dataviews/schoolagechildrenprofiles>

## 4. The enhanced NCMP dataset

- 4.1 After the collection year has closed and initial HSCIC validation has been completed, the enhanced NCMP dataset is made available for local authority download through the NCMP IT system. This dataset is provided as a CSV format text file containing all the record-level data that was originally input by users with the role of 'collector' at the local authority. Geographic fields based on child postcode, school postcode and derived fields such as BMI centile are also contained in the dataset. The fields included in this dataset are shown in Appendix 1. For data security reasons only the one user in each local authority holding the role of 'NCMP Lead' will be able to extract their local authority's enhanced NCMP dataset.
- 4.2 The HSCIC also shares a NCMP dataset with PHE, this dataset does not include any PID for the child, for example name, address, date of birth. See Appendix 1 for the list of fields included in the local authority and PHE datasets.
- 4.3 The NCMP system validates data as it is entered. For each record the system checks that all mandatory fields have been populated and that each field contains valid data. Records with missing mandatory fields or invalid data cannot be saved (classified as 'rejections'). Records with possible errors (such as extreme measurements) can be saved, but will generate 'warnings' that the data provider will need to confirm before they can finalise their submission. After the submission deadline date, the HSCIC extracts all non-personally identifiable data for further validation. Local authorities are notified of any data quality problems at this stage and requested to resubmit their data.
- 4.4 The HSCIC then carries out more detailed validation to identify any data quality issues. A document outlining all the validation the HSCIC undertakes is available on the NCMP IT system webpage at [www.hscic.gov.uk/ncmpsystem](http://www.hscic.gov.uk/ncmpsystem). Records with possible errors are flagged in the national dataset so users can decide whether they wish to use them in any analysis they may carry out. As a general principle, such records are used in the national analysis unless their inclusion dramatically impacts on the national level results. Local authorities are contacted if any outliers exist in their data.
- 4.5 The enhanced datasets shared with local authorities and PHE include a column labelled 'NCMPSchoolType' (in the local authority dataset) or 'NCMPSchoolStatus' (in the PHE dataset). Records from independent schools and special schools are coded as 'Independent' and state school records coded as 'State'. Local authorities are not required to take measurements at

independent and special schools and many such schools do not participate in the NCMP. As a result these records are not necessarily a representative sample of pupils attending such schools and are therefore excluded from the national analysis. **These records need to be excluded if analysis is to match the HSCIC's figures.**

- 4.6 The sections below provide more information on key fields in the enhanced data.

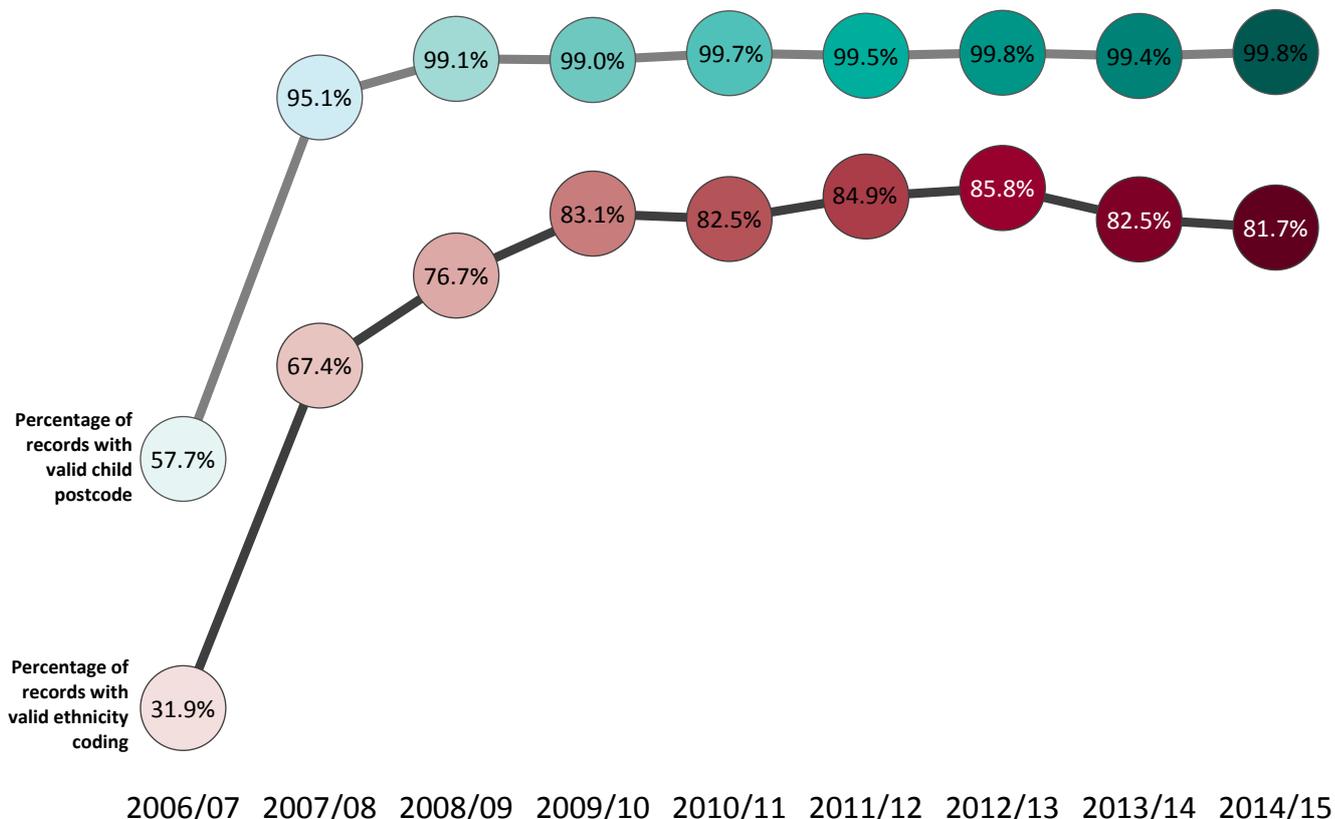
### Geographic coding

- 4.7 The NCMP dataset contains a number of fields providing geographic information for each child record. These fields can be assigned based on the the school attended or the child's postcode.
- 4.8 The different methods of assigning geographies may lead to small differences between figures for local authorities or Government Office Regions (GORs), in some cases, even where these areas have the same boundaries. For example, a pupil may attend a school in one local authority, but be resident in another.
- 4.9 The HSCIC publish NCMP statistics at local authority level using geographic information derived from school postcode, child postcode, and based on the local authority that submitted the data. Statistics included in the PHE NCMP Profile, the Public Health Outcomes Framework, and Health Profiles use local authority derived from child postcode.<sup>b</sup>
- 4.10 The proportion of child records with valid coding for area of residence has improved considerably over the years of the NCMP and in 2014/15 area of residence coding could be assigned to 99.8% of records (see Figure 1).

---

<sup>b</sup> Prior to February 2016 the data in these publications were based on local authority of school.

**Figure 1: proportion of records with valid child postcode and ethnicity coding (England totals by year of measurement)**



## Geographically assigned variables

4.11 The enhanced NCMP dataset contains a number of assigned variables based on the geographic coding. These include an indicator of deprivation; an urban/rural classification; and supergroups, subgroups and groups based on the Office of National Statistics Area Classification (ONS-AC).

4.12 These indicators are based both on the lower super output area (LSOA) of the school and the LSOA of residence of the child (which themselves are derived from the school and child postcodes). The indicators which are based on the LSOA of the school are populated for all pupils, while those which are based on the LSOA of the child's residence are not. Indicators based on the LSOA of residence will typically produce more accurate analysis of the resident population in an area. 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 have a valid postcode of residence.

4.13 The deprivation indicator provided in the dataset consists of an IMD 2010 decile. **IMD decile one includes those lower layer super output areas (LSOAs) within the most deprived 10% of all the LSOAs in England, whereas decile ten contains the least deprived 10%. Note that this labelling is reversed in some of the earlier years' of NCMP data.** The 2015 IMD values for 2011 LSOAs were not available for the 2014/15 NCMP extract but are expected to be included in the 2015/16 dataset. More information on the IMD is available here:

[www.gov.uk/government/collections/english-indices-of-deprivation](http://www.gov.uk/government/collections/english-indices-of-deprivation).

4.14 The supergroup, subgroup and group classification based on 2011 ONS-AC (see sections 2.10 and 6.14) 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](http://www.neighbourhood.statistics.gov.uk/dissemination/Info.do?page=nessgeography/areaclassification/area-classification.htm).

## Child measurements

4.15 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.

4.16 Two fields are provided in the dataset which show whether a child is underweight, healthy weight, overweight, or very overweight (obese).

4.17 The 'BMI\_PopulationCategory' field shows whether individual children are classified as underweight, healthy weight, overweight or very overweight for population monitoring according to the 2<sup>nd</sup>, 85<sup>th</sup> and 95<sup>th</sup> centiles of the British 1990 growth reference (UK90).<sup>6,7</sup> These population monitoring thresholds for BMI are used for reporting population prevalence in publications by the HSCIC and PHE.

4.18 The 'BMI\_ClinicalCategory' field uses the clinical cut-offs of the UK90 growth reference which classify children into these categories according to the 2<sup>nd</sup>, 91<sup>st</sup> and 98<sup>th</sup> centiles of the reference (see Appendix 3). These clinical thresholds are used to provide individual level child measurement feedback to parents.

## 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. The HSCIC ensures that any data it disseminates comply with the NHS anonymisation standard [www.isb.nhs.uk/library/standard/128](http://www.isb.nhs.uk/library/standard/128) and the NCMP regulations [www.legislation.gov.uk/ukxi/2013/218/pdfs/ukxi\\_20130218\\_en.pdf](http://www.legislation.gov.uk/ukxi/2013/218/pdfs/ukxi_20130218_en.pdf).
- 5.2 Users should familiarise themselves with the published data to ensure they are not duplicating analysis that has already been undertaken nationally.
- 5.3 Local analysis of population prevalence should use the population monitoring BMI thresholds (see section 4.17) and where possible be checked against the figures published by the HSCIC and PHE to ensure consistency.
- 5.4 Pupil data from independent and special schools should be excluded to match published data (see section 4.5).
- 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.
- 5.6 Any publications using NCMP data should clearly state the thresholds used (usually the 85<sup>th</sup> and 95<sup>th</sup> centiles of the UK90 growth reference) to derive obesity, overweight, healthy weight, and underweight prevalence figures.
- 5.7 The HSCIC retains copyright of NCMP information and this must be acknowledged in any publication which uses NCMP data. The correct citation is:  
'Copyright © 2015, re-used with the permission of The Health and Social Care Information Centre. All rights reserved.'
- 5.8 PHE and HSCIC are keen to see any local NCMP analysis reports in order to share 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 successive 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, 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 local authority level data quality measures are included in table 8 of the HSCIC's 2014/15 NCMP report. Additionally, local authority level data quality indicators are included in PHE's local authority NCMP profile: <http://fingertips.phe.org.uk/profile/national-child-measurement-programme>.
- 6.2 **Participation rate and prevalence:** published analyses of the 2006/07 and 2007/08 NCMP datasets suggest that low rates of participation in the NCMP at Primary Care Trust (PCT) level may have been 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 – where children who did not participate in the NCMP are more likely to be obese than those who did participate.
- 6.3 In more recent years, as participation rates have increased, the impact of selective opt-out on prevalence rates appears to have declined. Information on the effect of participation on prevalence in the current NCMP dataset is included in Annexes B and D of the HSCIC's 2014/15 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 local level. This issue is likely to be particularly important when comparing prevalence figures for areas with very different participation rates, or when considering change over time in areas where participation rates have increased or decreased markedly over the same period.
- 6.4 Local authorities may wish to identify areas of low participation within their NCMP data and also the reasons for this – for example whether it is due to entire schools not taking part (selective opt-out may 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.

- 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 at 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 from which measurements were taken. This figure will be higher than the published participation rates if there were schools that did not take part in the programme.
- 6.6 **Participation by sex:** previous NCMP analyses has 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 be considered when looking at differences in prevalence by sex at a local level. Participation rates split by sex cannot be calculated centrally by the HSCIC as the numbers of children eligible for measurement in each local authority is not available by sex.
- 6.7 **Data quality:** the national NCMP dataset has undergone extensive cleaning, but there is a limit to the checks that can be undertaken nationally with such a large dataset. It is therefore advised that local authorities 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-AC have been added to the NCMP dataset in order to facilitate the analysis of these factors.
- 6.9 Local authorities can also use their local intelligence 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 used in child health systems, of which there are approximately 230. Therefore analyses can easily be performed at a local level for different ethnic groups. To avoid potential problems due to low numbers of children in some ethnic groups, it may be necessary to combine ethnic groups where appropriate, or combine data from a number of years of NCMP measurements. This technique has been used to produce obesity prevalence by ethnic group for all local authorities and published in the PHE NCMP Profile, <http://fingertips.phe.org.uk/profile/national-child-measurement-programme>.

- 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 enhanced NCMP dataset made available to local authorities. 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. In 2014/15 the postcode of residence was available for 99.5 per cent of children with nearly all missing postcodes in one local authority. 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. The PHE NCMP profile includes obesity prevalence by IMD quintile for all local authorities available at: <http://fingertips.phe.org.uk/profile/national-child-measurement-programme>. An example of such analysis at national level can be found in the PHE child obesity and socioeconomic status data factsheet available at: [www.noo.org.uk/NOO\\_pub/Key\\_data](http://www.noo.org.uk/NOO_pub/Key_data).
- 6.13 **Setting:** as shown in the HSCIC's 2014/15 NCMP report, obesity prevalence can vary between urban and rural areas. Urban and rural coding is available within the enhanced NCMP dataset to allow local areas to investigate these patterns, however it is likely that many of the differences in obesity prevalence between urban and rural areas can be explained by differences in the sociodemographic mix of urban and rural populations.
- 6.14 **ONS Area Classification:** the ONS-AC 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. More information on this population stratification system can be

found in the report NCMP Analysis using the ONS Area Classification.<sup>3</sup> 2011 ONS Area Classification 'Group', 'Subgroup' and 'Supergroup' variables have been assigned to records using both LSOA of school and LSOA of child (also see sections 2.10 and 4.14).

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 or produced locally through analysis of historical local NCMP datasets.

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 significant. The suggested method for establishing the statistical significance of a change in prevalence is described in Appendix 4.

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.<sup>8,9</sup> Mean BMI z score is available in the 'BMIZScore' field in the enhanced NCMP dataset. 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 Data collected during the 2014/15 school year is likely to contain a large cohort of children who were previously measured in Reception during 2008/09. 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. In some local areas it is possible to compare the 2014/15 NCMP data for Year 6 children with child measurements for the same children taken in Reception year during 2008/09 using data held locally which contains a child identifier.

6.20 Now that the NHS number is part of the NCMP collection it will be possible to carry out this type of analysis on a national basis but this can only start when the Reception year children measured in 2013/14 are remeasured when they reach Year 6 in 2019/20.

## Small area analysis

- 6.21 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.
- 6.22 Sub-local authority 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 a MSOA or ward code from the LSOA code. Such analysis needs to be performed and interpreted with caution.
- 6.23 Prevalence figures for sub-local authority 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).
- 6.24 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. PHE publishes NCMP statistics at Ward and MSOA level using three years of data at: [www.noo.org.uk/visualisation](http://www.noo.org.uk/visualisation). 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.
- 6.25 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.
- 6.26 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.
- 6.27 PHE has 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 and explores what methods are appropriate and available for analysis to ensure useful and valid results.<sup>10</sup>

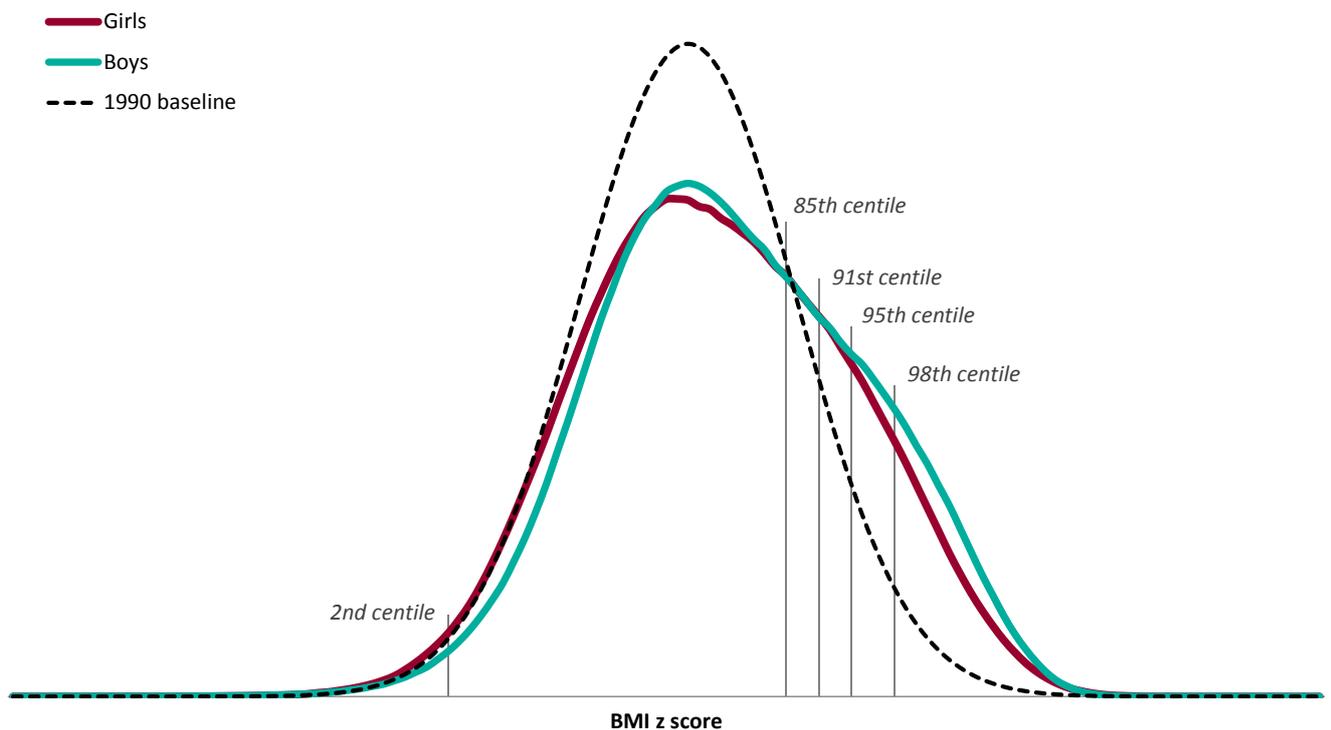
## Alternatives to small area analyses

- 6.28 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.
- 6.29 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.
- 6.30 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 greater levels of socioeconomic deprivation. By using fewer groups, based on similar characteristics, the underlying pattern can be seen much more clearly.
- 6.31 Analyses of this kind can often provide much more useful information regarding the local determinants of obesity and generally provide 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 towards these areas with a good degree of confidence that they will reach those individuals most at risk of future obesity-related ill health.

## BMI distribution

6.32 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 those measured rather than only considering the proportion of underweight, overweight, and obese children. Figure 2 provides an example of the distribution of BMI z scores within the 2014/15 NCMP dataset for children in Year 6, compared to the 1990 baseline.

**Figure 2: distribution of BMI z score for children in Year 6, NCMP 2014/15**



6.33 In this example, 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 vary 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 enhanced NCMP dataset.

6.34 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 and related answers have been compiled from common queries received by PHE.

### Why should I use the HSCIC enhanced NCMP extract?

7.2 The HSCIC performs extensive data quality checks and where data quality issues are identified this information is fed back to local authorities. Analysis of the enhanced NCMP extract is recommended to ensure accuracy and 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.

### Can I combine data for school years?

7.3 Prevalence figures should usually be produced separately for Reception and Year 6, rather than combining the data. Prevalence of child obesity varies with age and is generally 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.

7.4 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?

7.5 Prevalence of child obesity is known to vary by sex. Users of the NCMP dataset may wish to further investigate differences by sex within their local area.

7.6 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 2014/15 dataset. 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.

- 7.7 This issue is likely to be particularly important at school level. Comparing the 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?

- 7.8 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.
- 7.9 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 E of the HSCIC report. See [www.hscic.gov.uk/catalogue/PUB19109](http://www.hscic.gov.uk/catalogue/PUB19109).
- 7.10 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.
- 7.11 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 Reception to Year 6 measurements of the same children?

- 7.12 In some areas it will be possible to compare the 2014/15, 2013/14 and 2012/13 NCMP data for Year 6 children with measurements for the same children taken in the Reception year during 2008/09, 2007/08 and 2006/07 respectively using data held locally which includes a child identifier. Some local authorities may even be able perform such analysis with earlier NCMP datasets – for example where the height and weight of Reception year children were routinely measured and recorded prior to the NCMP, in the 2005/06 National Childhood Obesity Dataset (NCOD).

- 7.13 Although it is interesting to compare change in these cohorts of children at a population level (for example comparing obesity prevalence in 2008/09 for children of Reception year with obesity prevalence in 2014/15 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).
- 7.14 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.
- 7.15 Such analysis requires individual children to be identified within the dataset so that they can be tracked from Reception to Year 6. The enhanced NCMP dataset now contains child NHS number (where available) in order to facilitate such analysis but the historic NCMP data do not have this field. It is thus likely that, at least until 2019, such tracking work can only be performed where child measurements were recorded onto a local system (such as a child health system) and can be extracted in an identifiable form.<sup>c</sup>
- 7.16 Where identifiable data are available, the sort of analysis that is likely to be most productive includes calculating 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.
- 7.17 Tools such as LMS Growth (see Appendix 3) are likely to be helpful to anyone performing such analysis. This software can be used to assign BMI z scores, centiles, or BMI classification based on child measurements. Such categorisation will be needed if child measurements are being extracted from a local data source and not the national NCMP dataset. (See Appendix 3, sections A3.11 to A3.12 for further details.)
- 7.18 As PHE does not have access to data that enables tracking of individual children, 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,<sup>11</sup> Southampton PCT<sup>12</sup> and South Gloucestershire.<sup>13</sup> Local authorities that intend to perform such analysis might wish to review that undertaken in these areas before deciding on a plan for individual level analysis.

---

<sup>c</sup> The NCMPSystemID field within the NCMP datasets is an individual row identifier for each NCMP year and is only unique within each separate dataset. The same pupil identifier is 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.

7.19 If any local authorities wish to conduct analysis of this sort, PHE would be happy to discuss suitable approaches and methods prior to commencement of 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. These can then be shared with other areas.

## Appendix 1: NCMP 2014/15 enhanced extract data fields

A1.1 The 2014/15 NCMP enhanced dataset is supplied to local authorities as a single data table which contains information about both pupil (for both valid and excluded records) and school. Fields marked with a † are not included in the PHE dataset.

A1.2 The field names included in the data table are given below:

**Table A1: local authority enhanced extract data fields**

| Field                        | Description  |
|------------------------------|--|
| LACode                       | Code of the local authority that submitted the data  |
| LAName                       | Name of the local authority that submitted the data  |
| NCMPSystemId                 | Unique code for each record generated by the NCMP System                                     |
| PupilReference               | Local authority pupil reference code (if provided)   |
| NHSNumber†                   | NHS number   |
| FirstName†                   | Pupil first name   |
| LastName†                    | Pupil last name  |
| Sex                          | Pupil sex  |
| DateOfBirth†                 | Pupil date of birth  |
| AgeInMonths                  | Age in months on date of measurement   |
| SchoolYear                   | Reception or Year 6  |
| Ethnicity                    | Ethnicity code provided by local authority   |
| NHSEthnicCode                | Your ethnicity code mapped to NHS ethnic code  |
| NHSEthnicDescription         | Your ethnicity code mapped to NHS ethnic description   |
| NCMPEthnicity                | Your ethnicity code mapped to NCMP ethnic category   |
| Height                       | Pupil height   |
| HeightZScore                 | The number of standard deviations the height is above or below the mean                      |
| Weight                       | Pupil weight   |
| WeightZScore                 | The number of standard deviations the weight is above or below the mean                      |
| Bmi                          | Pupil body mass index  |
| BmiZScore                    | The number of standard deviations the BMI is above or below the mean                         |
| ClinicalBMlcategory          | Weight status based on clinical thresholds   |
| GroupedClinicalBMlcategory   | Weight status based on clinical thresholds grouped to combine overweight and very overweight |
| BmiCentile(PScore)           | BMI centile  |
| PopulationBMlcategory        | Weight status based on population thresholds   |
| GroupedPopulationBMlcategory | Weight status based on population thresholds grouped   |

| Field   | Description   |
|---|---|
|   | to combine overweight and very overweight   |
| DateOfMeasurement   | Date pupil was measured   |
| DayOfMeasurement  | Day of measurement  |
| MonthOfMeasurement  | Month of measurement  |
| NonMeasurementReasonCode  | Non-measurement reason code   |
| NonMeasurementReasonDescription   | Non-measurement reason description  |
| FeedbackDateSent  | Date that feedback extract for parental letter was generated                            |
| Address1†   | Pupil address line 1  |
| Address2†   | Pupil address line 2  |
| Address3†   | Pupil address line 3  |
| Address4†   | Pupil address line 4  |
| Address5†   | Pupil address line 5  |
| Postcode†   | Pupil postcode  |
| Notes†  | Any notes entered about pupil   |
| ParentTelephone†  | Parent telephone number   |
| ParentEmail†  | Parent email address  |
| SchoolUrn   | 6 digit DfE school reference number   |
| SchoolName  | School name   |
| SchoolPostcode  | School postcode   |
| NcmpSchoolType ( <i>This field is named NCMPSchoolStatus in the PHE dataset</i> ) | Description of whether the school is a 'state' or 'independent' school                  |
| PupilSchoolPostcodeDistance   | Distance in km between child's home and school  |
| PupilGovernmentOfficeRegion   | The region code based on pupil postcode   |
| PupilGovernmentOfficeRegionCode   | ONS GOR Code  |
| PupilGovernmentOfficeRegionName   | ONS GOR Name  |
| Pupil2001CensusOutputArea   | Census output area 2001   |
| Pupil2001LowerSuperOutputArea   | Lower Super Output Area 2001  |
| Pupil2001MiddleSuperOutputArea  | Middle Super Output Area 2001   |
| Pupil2011CensusOutputArea   | Census Output Area 2011   |
| Pupil2011LowerSuperOutputArea   | Lower Super Output Area 2011  |
| Pupil2011MiddleSuperOutputArea  | Middle Super Output Area 2011   |
| PupilNationalGridReferenceNorthing†   | National Grid Reference (100m) – Northing – the Ordnance Survey postcode grid reference |
| PupilNationalGridReferenceEasting†  | National Grid Reference (100m) – Easting - the Ordnance Survey postcode grid reference  |
| PupilElectoralWard  | The administrative/electoral area based on pupil postcode                               |
| PupilUrbanRuralIndicator  | Urban and rural classification of pupil output area                                     |
| PupilUrbanRuralDescription  | Urban and rural classification of pupil output area                                     |
| PupilUrbanRuralGroup  | Urban and rural classification grouping 'sparse' and 'less sparse' together             |
| PupilTier2LocalAuthority  | Lower tier local authority based on pupil postcode                                      |
| PupilIndexOfMultipleDeprivationScore  | Relative measure of deprivation based on pupil postcode (IMD2010)                       |

| <b>Field</b>                           | <b>Description</b>  |
|--|---|
| PupilIndexOfMultipleDeprivationDecile  | Relative measure of deprivation based on pupil postcode (IMD2010)                       |
| PupilOnsSupergroupCode                 | ONS classification based on population characteristics (2011 ONS Area Classification)   |
| PupilOnsSupergroupDescription          | ONS classification based on population characteristics (2011 ONS Area Classification)   |
| PupilOnsGroupCode                      | ONS classification based on population characteristics(2011 ONS Area Classification)    |
| PupilOnsGroupDescription               | ONS classification based on population characteristics(2011 ONS Area Classification)    |
| PupilOnsSubgroupCode                   | ONS classification based on population characteristics(2011 ONS Area Classification)    |
| PupilOnsSubgroupDescription            | ONS classification based on population characteristics(2011 ONS Area Classification)    |
| SchoolGovernmentOfficeRegion           | The region code based on school postcode  |
| SchoolGovernmentOfficeRegionCode       | ONS GOR Code  |
| SchoolGovernmentOfficeRegionName       | ONS GOR Name  |
| School2001CensusOutputArea             | Census output area 2001   |
| School2001LowerSuperOutputArea         | Lower Super Output Area 2001  |
| School2001MiddleSuperOutputArea        | Middle Super Output Area 2001   |
| School2011CensusOutputArea             | Census output area 2011   |
| School2011LowerSuperOutputArea         | Lower Super Output Area 2011  |
| School2011MiddleSuperOutputArea        | Middle Super Output Area 2011   |
| SchoolNationalGridReferenceNorthing    | National Grid Reference (100m) – Northing – the Ordnance Survey postcode grid reference |
| SchoolNationalGridReferenceEasting     | National Grid Reference (100m) – Easting – the Ordnance Survey postcode grid reference  |
| SchoolElectoralWard                    | The administrative/electoral area based on school postcode                              |
| SchoolUrbanRuralIndicator              | Urban and rural classification of school output area                                    |
| SchoolUrbanRuralDescription            | Urban and rural classification of school output area                                    |
| SchoolUrbanRuralGroup                  | Urban and rural classification grouping 'sparse' and 'less sparse' together             |
| SchoolTier2LocalAuthority              | Lower tier local authority based on school postcode                                     |
| SchoolIndexOfMultipleDeprivationScore  | Relative measure of deprivation based on school postcode (IMD2010)                      |
| SchoolIndexOfMultipleDeprivationDecile | Relative measure of deprivation based on school postcode (IMD2010)                      |
| SchoolOnsSupergroupCode                | ONS classification based on population characteristics (2011 ONS Area Classification)   |
| SchoolOnsSupergroupDescription         | ONS classification based on population characteristics (2011 ONS Area Classification)   |
| SchoolOnsGroupCode                     | ONS classification based on population characteristics (2011 ONS Area Classification)   |
| SchoolOnsGroupDescription              | ONS classification based on population characteristics (2011 ONS Area Classification)   |
| SchoolOnsSubgroupCode                  | ONS classification based on population characteristics (2011 ONS Area Classification)   |
| SchoolOnsSubgroupDescription           | ONS classification based on population characteristics (2011 ONS Area Classification)   |

## Appendix 2: Suggested data quality checks at local level

- A2.1 Although the enhanced NCMP dataset provided to local authorities has undergone extensive cleaning at national level, there is a limit to the checks and cleaning that can be undertaken centrally on a dataset with over one million records annually from around 17,000 schools. As a result there may be some minor remaining data quality issues.
- A2.2 In earlier NCMP datasets a number of issues were identified during analysis. These included: duplicate records in the dataset; Year 6 pupils entered to infant schools or Reception pupils entered to junior schools; and a large proportion of records with 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 undertaken by the HSCIC prior to the data being released. A document outlining all the HSCIC validation checks is available on the NCMP IT system webpage at [www.hscic.gov.uk/ncmpsystem](http://www.hscic.gov.uk/ncmpsystem), however, some data quality issues may remain despite these additional checks.
- A2.4 These data quality issues have a minimal effect on national analyses but may be more important in detailed regional or local analyses by PHE or local authorities. It is therefore key 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 HSCIC ([ncmp@hscic.gov.uk](mailto:ncmp@hscic.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 **Validation warnings flagged by NCMP IT System:** it is important that local authorities look carefully at each validation warning and check to make sure the data being flagged are correct. There have been recent examples, particularly around extreme weights, where local authorities have confirmed data as correct in response to a validation warning only to find it was an error after the data have been published.

- A2.5.2 **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.
- A2.5.3 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.
- A2.5.4 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.
- A2.5.5 **Duplicate pupils:** duplicate pupils assigned to the same school will be blocked by the NCMP IT system.
- A2.5.6 In some situations (for example where duplicate records were submitted to separate schools) some duplicate records may remain in the dataset. It is therefore worthwhile for local authorities to perform some quick checks to see if any such records can be identified.
- A2.5.7 Detailed analyses of previous NCMP datasets showed instances where the same set of child records were submitted 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 This can also occur when a school closes and reopens as an academy. If the local authority wishes to assign the records to the new academy school then it is important that the records for the predecessor school are removed otherwise they will be counted twice in the dataset.
- A2.5.9 **Rounded records:** the NCMP operational guidance states that height and weight measurements should to be recorded to the

nearest one decimal place. Analyses of previous NCMP datasets has shown that incorrectly rounded records (especially those for weight in Reception year) were associated with a lower reported prevalence of obesity.<sup>14</sup>

A2.5.10 The HSCIC runs 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 2014/15 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 2014/15 NCMP report (table 8), 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 will 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 enhanced NCMP dataset contains a field showing day in the week of measurement.

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 which affects the calculated age of the child at measurement and ultimately affects the BMI score.

A2.5.14 Local authorities 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.

A2.5.15 A similar data quality measure exists where the month of measurement is given as August when a school would be closed.

## 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 50<sup>th</sup> 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 +1.645 (the 95<sup>th</sup> 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 2<sup>nd</sup>, 85<sup>th</sup> and 95<sup>th</sup> 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 85<sup>th</sup> and 95<sup>th</sup> 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 2<sup>nd</sup>, 91<sup>st</sup> and 98<sup>th</sup> centiles of the UK90 reference are used in the UK to classify individual children as underweight, healthy, overweight or very overweight (obese) taking into account the expected variation in BMI by age and sex.

The NCMP parental feedback letters issued by local authorities 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 2.3<sup>rd</sup>, 90.9<sup>th</sup> and 97.7<sup>th</sup> centiles when rounded to one decimal place, although they are usually referred to as the 2<sup>nd</sup>, 91<sup>st</sup> and 98<sup>th</sup> centiles.
- A3.8 Underweight prevalence figures using both population monitoring and clinical thresholds are each presented as having been derived using the 2<sup>nd</sup> centile, however, only those based on population monitoring thresholds use the 2<sup>nd</sup> centile, whereas those derived using clinical thresholds actually use  $-2$  standard deviations (the 2.3<sup>rd</sup> centile). For a given population, prevalence of underweight will therefore differ slightly depending on whether the population monitoring or clinical thresholds are used.
- A3.9 When presenting prevalence figures based on the 85<sup>th</sup> and 95<sup>th</sup> centile thresholds, 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 free of charge from Harlow Publishing at [www.healthforallchildren.com/?product=lmsgrowth](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 'NORM.S.DIST' and 'NORM.S.INV' functions.
- 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.

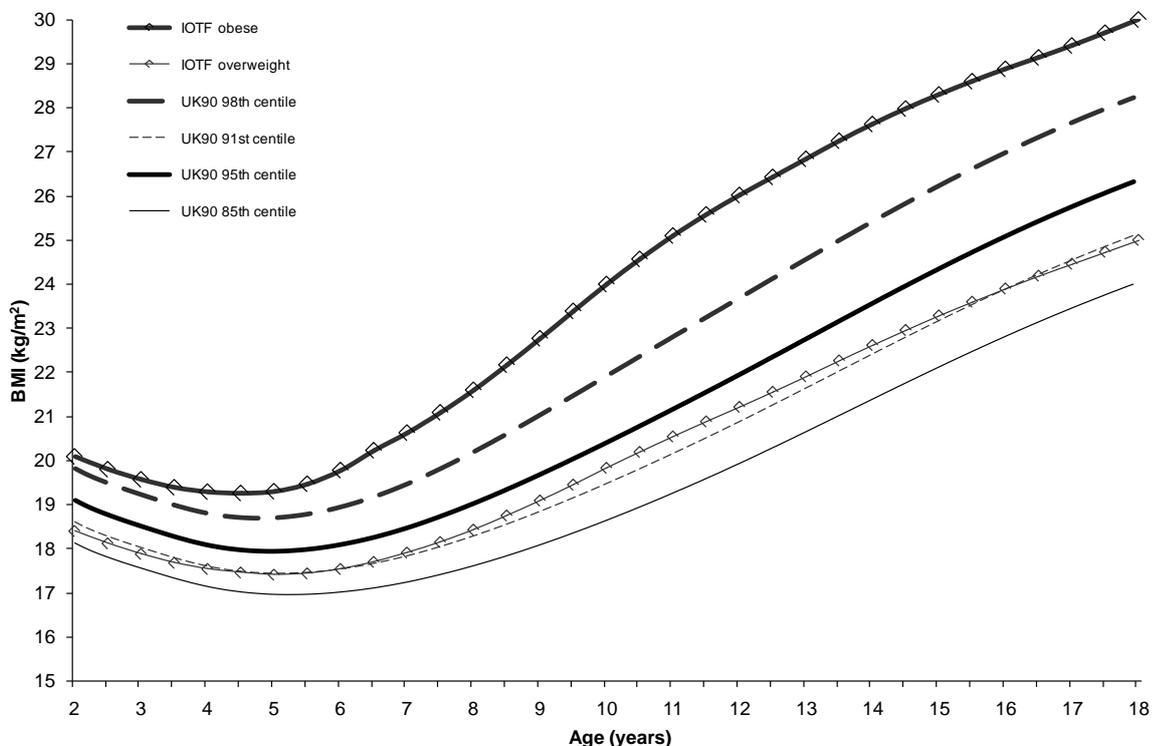
## Other thresholds for defining children’s BMI status

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<sup>15</sup> and for child obesity prevalence figures from the Millennium Cohort Study.<sup>16</sup>

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. (see Figure 3).

A3.15 Most published NCMP analyses use the recommended UK90 population monitoring thresholds to ensure consistency between published figures. When 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 3: 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<sup>17</sup> and Newcombe<sup>18</sup> 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

$$(A-B)/C \text{ to } (A+B)/C$$

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)*<sup>19</sup> 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’.<sup>20</sup>

## References

1. Department of Health. Healthy lives, healthy people: our strategy for public health in England. London: Department of Health, November 2010  
[www.dh.gov.uk/en/Publicationsandstatistics/Publications/PublicationsPolicyAndGuidance/DH\\_121941](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  
[www.dh.gov.uk/en/Publicationsandstatistics/Publications/PublicationsPolicyAndGuidance/DH\\_132358](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  
[www.noo.org.uk/uploads/doc/vid\\_12524\\_NCMP\\_Analysis%20using%20the%20ONS%20Area%20Classification.pdf](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  
[www.noo.org.uk/uploads/doc/vid\\_8785\\_NCMP\\_Variation%20by%20month%20of%20measurement\\_v2.pdf](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.
6. 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.
7. Cole TJ, Freeman JV, Preece MA. Bodymass index reference curves for the UK. *Arch Dis Child* 1995; 73: 25–29.
8. 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.
9. 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.
10. Dinsdale H, Ridler C. National Child Measurement Programme: Guidance for small area analysis. Oxford: National Obesity Observatory, 2011  
[www.noo.org.uk/uploads/doc/vid\\_11853\\_NCMP\\_Guidanceforsmallarea%20analysisFINAL.pdf](http://www.noo.org.uk/uploads/doc/vid_11853_NCMP_Guidanceforsmallarea%20analysisFINAL.pdf).
11. Porter M, Greene T, Taylor A. Childhood obesity in Hull: paired analysis. Hull: Hull PCT, December 2007.

12. King D. Child Growth Briefing Note; National Child Measurement Programme 2001/02 to 2008/09 Summary Report. Southampton: Southampton PCT, 2010 [www.publichealth.southampton.gov.uk/Images/Child%20Growth%20Report%20SC%20PCT%20\(2010\).pdf](http://www.publichealth.southampton.gov.uk/Images/Child%20Growth%20Report%20SC%20PCT%20(2010).pdf).
13. Pearce M, Webb-Phillips S, Bray I. Changes in objectively measured BMI in children aged 4–11 years: data from the National Child Measurement Programme. *J Public Health* 2015 (epub ahead of print).
14. Dinsdale H, Rutter H. National Child Measurement Programme: Detailed Analysis of the 2006/07 National Dataset. London: Department of Health, June 2008. [www.noo.org.uk/uploads/doc168\\_2\\_NOO\\_NCMP\\_report230608.pdf](http://www.noo.org.uk/uploads/doc168_2_NOO_NCMP_report230608.pdf).
15. 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 [www.bis.gov.uk/foresight/our-work/projects/current-projects/tackling-obesities](http://www.bis.gov.uk/foresight/our-work/projects/current-projects/tackling-obesities).
16. 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](http://eprints.ioe.ac.uk/5931/1/MCS_3_Descriptive_Report_Oct_2008.pdf).
17. Wilson EB. Probable inference, the law of succession, and statistical inference. *J Am Stat Assoc* 1927; 22:209–12.
18. Newcombe RG. Two-sided confidence intervals for the single proportion: comparison of seven methods. *Stat Med* 1998; 17:857–72.
19. Altman DG, Machin D, Bryant TN, Gardner MJ. *Statistics with Confidence*, 2nd edn. London; BMJ books, 2000: 49.
20. Newcombe RG, Interval estimation for the difference between independent proportions: comparison of eleven methods. *Stat Med* 1998; 17:873–90.

All weblinks accessed: 25/04/2016

## Reader information

|                            |  |
|----------------------------|--|
| <b>Title</b>               | NCMP Guidance for data sharing and analysis  |
| <b>Author(s)</b>           | Caroline Hancock, Vicky Copley   |
| <b>Reviewer(s)</b>         | Paul Niblett, Health and Social Care Information Centre<br>Sam Widdowfield, Health and Social Care Information Centre<br>Di Swanston, Public Health England        |
| <b>Publication date</b>    | June 2016  |
| <b>Target audience</b>     | Local authorities and Public Health England analysts   |
| <b>Description</b>         | Information and guidance for data sharing and analysis by local authorities and Public Health England of the 2014/15 National Child Measurement Programme dataset. |
| <b>How to cite</b>         | Hancock C, Copley V. NCMP Guidance for data sharing and analysis. London: Public Health England, Obesity Risk Factors Intelligence, 2016.                          |
| <b>Contact</b>             | <a href="http://www.noo.org.uk">www.noo.org.uk</a><br><a href="mailto:ncmp@phe.gov.uk">ncmp@phe.gov.uk</a><br><a href="#">@PHE_obesity</a>                         |
| <b>Electronic location</b> | <a href="http://www.noo.org.uk/NCMP/analytical_guidance">www.noo.org.uk/NCMP/analytical_guidance</a>   |
| <b>Copyright</b>           | © Crown copyright 2016   |