



Public Health
England

Protecting and improving the nation's health

National Child Measurement Programme

Guidance for analysis and data sharing

About Public Health England

Public Health England exists to protect and improve the nation's health and wellbeing, and reduce health inequalities. We do this through world-leading science, knowledge and intelligence, advocacy, partnerships and the delivery of specialist public health services. We are an executive agency of the Department of Health and Social Care, and a distinct delivery organisation with operational autonomy. We provide government, local government, the NHS, Parliament, industry and the public with evidence-based professional, scientific and delivery expertise and support.

Public Health England
Wellington House
133-155 Waterloo Road
London SE1 8UG
Tel: 020 7654 8000
www.gov.uk/phe
Twitter: [@PHE_uk](https://twitter.com/PHE_uk)
Facebook: www.facebook.com/PublicHealthEngland

For queries relating to this document, please contact: ncmp@phe.gov.uk



© Crown copyright 2018

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). Where we have identified any third party copyright information you will need to obtain permission from the copyright holders concerned.

Published September 2018
PHE publications
gateway number: 2018458

PHE supports the UN
Sustainable Development Goals



Contents

About Public Health England	2
Contents	3
1. Introduction	4
2. National reports and resources	5
3. Obtaining and sharing the NCMP dataset	8
4. The enhanced NCMP dataset	13
5. Essentials for NCMP analysis	17
6. Local level analysis	18
Appendix 1: Example school feedback letter	29
Appendix 2: NCMP enhanced extract data fields	30
Appendix 3: Suggested data quality checks at local level	33
Appendix 4: BMI thresholds, z scores, and p scores	36
Appendix 5: Calculating confidence limits	40
References	42

1. Introduction

The National Child Measurement Programme (NCMP) is an annual programme that measures the height and weight of children in Reception (age 4 to 5 years) and Year 6 (age 10 to 11 years) within state maintained schools in England. Some independent and special schools also choose to participate, but these records are usually excluded from the analysis for national NCMP reports as the majority of such schools do not take part (see p.13).

NHS Digital (NHSD) and Public Health England (PHE) publish analyses of this dataset annually and many local authorities undertake additional analyses to inform their work to tackle obesity.

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 NCMP data and analyses provided by PHE and NHSD (p.5)
- how to access the NCMP dataset (p.8)
- which NCMP data can be shared between organisations (p.9)
- guidance on the appropriate use of the NCMP dataset to comply with data protection and disclosure rules (p.10-11, 17)
- suggestions for regional and local analyses and caveats associated with the NCMP data and its interpretation (p.18)
- suggestions for local data quality checks (p.33)

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

2. National reports and resources

A number of resources to facilitate local use of NCMP data are produced every year by both NHSD and PHE. These consist of reports and data slide sets presenting and describing the NCMP data at national level; and data tables at national, regional, local authority, Clinical Commissioning Group (CCG), ward and Middle Super Output Area (MSOA) level. Users of the NCMP data are advised to familiarise themselves with these reports, slide sets 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 NHSD

Headline results from the NCMP at national, regional and local authority level are provided annually in NHSD's reports and data tables <https://digital.nhs.uk/data-and-information/publications/statistical/national-child-measurement-programme>.

These reports describe the prevalence of child severe obesity, obesity, overweight, healthy weight and underweight at national, regional and local authority level, as well as the variation across demographic and socioeconomic groups. A comparison is also provided for local authority data from previous years.

Excel data tables are published alongside the reports. These contain data on severe obesity, obesity, overweight, healthy weight and underweight prevalence for regions and local authorities (based on location of school, area of residence and submitting authority). Additional national level data is included, broken down by ethnic group, rural-urban classification and Index of Multiple Deprivation (IMD) decile.

NHSD also publishes a non-disclosive extract of record level NCMP data, for datasets 2013 to 2014 onwards, with an associated guidance document to accompany the annual report. Some of the data items collected as part of the NCMP are removed and others are altered to ensure there is no risk of any child being identified. Files for the 2006/07 to 2012/13 NCMP datasets are available from the UK Data Archive (UKDA) at <https://discover.ukdataservice.ac.uk/>.

NCMP data and analysis provided by PHE

PHE NCMP analysis reports

The PHE NCMP trends analysis reports and slide sets examine the changes that have taken place in the body mass index (BMI) of Reception and Year 6 children

over the period for which NCMP data is available. These reports also examine the impact of demographic and socioeconomic variables (such as ethnicity and deprivation) on observed patterns and trends.

www.gov.uk/government/publications/national-child-measurement-programme-ncmp-trends-in-child-bmi

Users of the NCMP dataset might also find it useful to read the [older 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.

[Previous reports published using NCMP data](#) include short reports describing the use of the ONS Area Classification (ONS-AC) (see pp.15 and 20) in NCMP analysis, and the impact of month of measurement on reported obesity prevalence within the NCMP dataset.

In March 2017 PHE published a report examining how weight status tracks in individual children during primary school over time, between Reception and Year 6. This analysis used NCMP data from local authorities. The report also provides an analytical protocol and step by step guidance on undertaking local level tracking analyses (see p.24). www.gov.uk/government/publications/weight-change-in-primary-school-age-children

NCMP and Child Obesity Profile

The NCMP and Child Obesity Profile online tool displays data for all years of the NCMP. Prevalence of severe obesity, obesity, overweight, healthy weight, and underweight for children in Reception and Year 6 can be examined at local authority level. The tool also presents trend data and enables easy comparison of local authority data, allowing users to compare regional neighbours and local authorities with similar characteristics. It includes inequalities data (sex, deprivation and ethnicity) for child obesity by local authority. Data quality indicators are also available in the tool, for example rate of participation in the NCMP. Contextual indicators provide local authority estimates for several topic areas that are determinants of, or related to, child obesity. For each topic area, where possible, the information is shown by gender, ethnicity, sexuality, region and local authority. This online tool can be accessed at: <http://fingertips.phe.org.uk/profile/national-child-measurement-programme>.

Other PHE NCMP publications

The Public Health Outcomes Framework (PHOF) indicators on child excess weight use data from the NCMP and are available in the PHE online PHOF data tool <https://fingertips.phe.org.uk/profile/public-health-outcomes-framework>

Spreadsheets containing NCMP obesity prevalence data for Reception (age 4 to 5 years) and Year 6 (age 10 to 11 years) children by CCG, electoral ward and MSOA of residence are available at www.gov.uk/government/statistics/child-obesity-and-excess-weight-small-area-level-data. Local authority and England data are also presented for comparison. These small area prevalence figures use three years of NCMP data combined. NCMP data at MSOA, ward, and CCG level are also included in the PHE Local Health tool: www.localhealth.org.uk.

National and regional slide sets on child obesity are available from PHE. These PowerPoint slides present key data and information on child obesity in clear, easy to understand charts and graphics. The national child obesity slide set summarises the latest national level data from the NCMP and the Health Survey for England (HSE) while the regional child obesity slide sets show customised NCMP data for each of the nine regions of England.

The slides can be **downloaded** and used freely with acknowledgement to Public Health England. Notes accompany each of the national slides and are available in the downloaded versions. The slides are a useful tool for practitioners and policy makers working on obesity and physical activity at local, regional and national level. They can be used in presentations to health and wellbeing boards, other committees and to elected members as well as in regional or national conference and workshop presentations. All the slide sets are also available to download at: <https://khub.net/web/phe-obesity-intelligence/public-library>

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.

3. Obtaining and sharing the NCMP dataset

Obtaining the data

NHSD launched an online NCMP data collection system in 2013 to 2014, and local authorities are 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.** NHSD sends emails to users of the system to let them know when the dataset is available and also to warn them before the system is purged (see section 4 and Appendix 2).

The enhanced dataset contains all the data originally provided by the local authority along with mapped geographic and deprivation-based fields and calculated fields such as child age and weight, height and BMI z-scores and equivalent centiles. It also contains information on any data quality issues identified by NHSD.

Although local authorities will already have access to their own NCMP data, it is recommended that the enhanced data from NHSD is 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 NHSD dataset also has a greater number of fields than locally held data which facilitates a greater range of possible analyses and removes the need for local authorities to calculate these additional data items themselves.

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 is required after the system purge then local authorities should contact NHSD (enquiries@nhsdigital.nhs.uk). NHSD 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.^a

^a Access to the PID within NHSD 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 is removed from the system as NHSD no longer has a need for it.

Further information on the NHSD NCMP IT system and guidance on how to obtain the enhanced data from the system is given at

<https://digital.nhs.uk/services/national-child-measurement-programme/it-system>.

PHE receives an analysis dataset directly from NHSD which is governed by a Memorandum of Understanding or a Data Sharing Agreement between the two organisations. The PHE dataset does not include any PID for the child, for example name, address, date of birth. See Appendix 2 for the list of fields included in the local authority and PHE datasets.

Local authorities should contact NHSD 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 NHSD will no longer have access to PID for these earlier years as explained on p.8. NHSD may levy a charge for meeting these data requests on a cost recovery basis.

Sharing data with other organisations

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 NHSD 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 on pp.10 and 11 below.

Often the information needed will already be available through the NHSD and PHE websites. Where possible users should be directed to these existing resources (see section 2).

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 in an aggregated and suppressed form. This ensures individual information is not shared unnecessarily and prevents the identification of individual children (see p.17 on suppression).

If more information is required than is publicly available, an application can be made directly to NHS Digital via their Data Access Request Service

<https://digital.nhs.uk/services/data-access-request-service-dars>.

NCMP, General Data Protection Regulation and data processing

The General Data Protection Regulations (GDPR) became UK law on 25 May 2018. All processing of personal data – meaning all aspects of the collection, analysis and dissemination of data about identifiable individuals – must have a lawful basis under the GDPR. <https://ico.org.uk/for-organisations/guide-to-the-general-data-protection-regulation-gdpr/key-definitions/>

All local authorities in England are required to collect information on the height and weight of Reception and Year 6 schoolchildren. The statutory authority^b for the NCMP means that the lawful basis for processing this data is considered to be provided by the GDPR Articles covering ‘compliance with a legal obligation’ and ‘the provision of health care or treatment’. See the [NCMP Operational Guidance](#) chapter 3 and Appendix 1 for further information.

Roles and responsibilities for NCMP data processing

Local authorities (as local level data controllers) are responsible for making decisions on how and who collects the data guided by the NCMP regulations, ensuring it is stored securely and any local sharing of the data conforms with the regulations. A local NCMP provider may be contracted to carry this out; this could be a school nursing team working in schools or a local health care provider (local level data processor). The team collecting the data enters it into the NCMP IT system (hosted by NHSD) and/or the local child health information database. Local authorities are responsible for sending the data to NHSD where it is stored securely. NHSD is responsible for the information processed in the national NCMP data collection tool (as data controllers for the national dataset).

<https://digital.nhs.uk/about-nhs-digital/our-work/keeping-patient-data-safe/gdpr/gdpr-register/national-child-measurement-programme-ncmp-gdpr-information>

Processing of information by local authorities

The NCMP regulations¹ 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.”

^b The statutory authority for processing NCMP data is provided by The Local Authorities (Public Health Functions and Entry to Premises by Local Healthwatch Representatives) Regulations 2013 and The Local Authority (Public Health, Health and Wellbeing Boards and Health Scrutiny) Regulations 2013.

NCMP regulations¹ 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 NHSD

NCMP regulations¹ state that NHSD 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**. NHSD will also apply charges for providing data on a cost-recovery basis. More information is available from <https://digital.nhs.uk/services/data-access-request-service-dars>.

Feeding back NCMP data to schools

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**.

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 this 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 p.21). 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.

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 combined 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.

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 is 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.

PHE has developed new resources to support local authorities to share NCMP information with schools. These include an annual individualised school feedback letter, combining the most recent three years of data to create a large enough sample size to provide a reliable picture of the weight status of the school population. An example of this is shown in Appendix 1.

School-age child health profiles for local authorities are available at <https://fingertips.phe.org.uk/profile-group/child-health/profile/child-health-school-age>.

4. The enhanced NCMP dataset

After the collection year has closed and initial NHSD validation has been completed, the enhanced NCMP dataset is made available for local authorities to 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 height, weight and BMI z-scores and centiles are also contained in the dataset. The fields included in this dataset are shown in Appendix 2. 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. NHSD also shares an anonymised NCMP dataset with PHE. <https://digital.nhs.uk/services/national-child-measurement-programme/it-system>

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 should look into and either change the record or confirm it is correct before they can finalise their submission. After the submission deadline date, the NHSD 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.

The NCMP system also provides data quality indicators at local authority level which local authorities should examine each time they upload data, so that early corrective action can be taken if any of these indicators are falling outside their allowable values.

NHSD then carries out more detailed validation to identify any data quality issues. A document outlining all the validation NHSD undertakes is available on the [NCMP IT system webpage](#). Records with potential 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.

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 mainstream state-maintained schools with

children in Reception and Year 6 are coded as 'NCMP' schools and all other schools are coded as 'Non-NCMP' schools.^c Local authorities are not required to take measurements in non-NCMP 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 NHSD's figures.

Geographic coding

The NCMP dataset contains a number of fields providing geographic information for each child record. These fields can be assigned, based on the school attended or the child's postcode.

In some cases, the different methods of assigning geographies may lead to small differences between figures for local authorities or Government Office Regions (GORs/Regions), 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.

NHSD publishes 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 and Child Obesity Profile, PHOF, and Health Profiles use local authority derived from child postcode.^d

The proportion of child records with valid coding for area of residence has improved considerably over the years of the NCMP and has been greater than 99% each year since 2008 to 2009. The proportion of child records with valid coding for area of residence at local authority level is available in the **NCMP and Child Obesity Profile**.

Geographically assigned variables

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 for National Statistics Area Classification (ONS-AC).

^c Prior to the 2017 to 2018 collection these categories were referred to as 'state' and 'independent' schools.

^d Prior to February 2016 the data in these publications was based on local authority of school.

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 records, while those which are based on the LSOA of the child's residence are only populated if the child's postcode was supplied. Indicators based on the LSOA of residence will typically produce more accurate interpretation 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 to 2007 and 2007 to 2008 NCMP datasets as from 2008 to 2009 more than 99% of child records have a valid postcode of residence.

The deprivation indicator provided in the dataset consists of an IMD decile. IMD decile 1 includes those LSOAs within the most deprived 10% of all the LSOAs in England, whereas decile 10 contains the least deprived 10%. Note that this labelling is reversed in some of the earlier years of NCMP data. More information on the IMD is available here: www.gov.uk/government/collections/english-indices-of-deprivation.

The supergroup, subgroup and group classification based on 2011 ONS-AC (see pp.6 and 20) 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-AC indicator is available here: www.ons.gov.uk/methodology/geography/geographicalproducts/areaclassifications/2011areaclassifications

Child measurements

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 4 of this guidance provides more information on how these variables are calculated and how they are used.

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

The 'BMI_PopulationCategory' field shows whether individual children are classified as underweight, healthy weight, overweight or very overweight for population monitoring according to the 2nd, 85th and 95th centiles of the British 1990 growth reference (UK90).^{2,3} These population monitoring thresholds for BMI are used for reporting population prevalence in publications by NHSD and PHE.

The 'BMI_ClinicalCategory' field uses the clinical cut-offs of the UK90 growth reference which classify children into categories based on their BMI z-scores. These do approximate very roughly to the 2nd, 91st and 98th centiles of the growth reference (see Appendix 4), but it is important to use the z-scores to assign the BMI clinical category rather than these approximate centiles. Clinical thresholds are used to monitor individual level child height, weight, and BMI and are therefore used to feedback NCMP measurements to parents.

5. Essentials for NCMP analysis

Before planning any local analysis, users should familiarise themselves with the published data to ensure they are not duplicating analysis that has already been undertaken by NHSD or PHE (see section 2).

When making NCMP data publicly available, data must only be released in a form such that children cannot be identified. Secondary suppression may need to be applied and corresponding cells providing totals should also be suppressed to avoid disclosure by differencing. NHSD ensures that any data it disseminates complies with the **NHS anonymisation standard** and the **NCMP regulations**.

Suppression should be applied to prevalence figures and corresponding counts where the number of children in the category ranges from 0 up to and including 5.

Local analysis of population prevalence should use the population monitoring BMI thresholds (see p.15 and Appendix 4) and where possible be checked against the figures published by NHSD and PHE to ensure consistency.

Pupil data from independent and special schools should be excluded to match published data (see p.14).

Confidence limits are published in the Excel data tables provided by NHSD 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 5.

Any publications using NCMP data should clearly state the thresholds used (usually the 2nd, 85th, 95th and 99.6th centiles of the UK90 growth reference) to derive underweight, healthy weight, overweight, obesity, and severe obesity prevalence figures.

NHSD retains copyright of NCMP information, and this must be acknowledged in any publication which uses NCMP data. The correct citation is: 'Copyright © [YEAR], re-used with the permission of The Health and Social Care Information Centre. All rights reserved. NHS Digital is the trading name of the Health and Social Care Information Centre'.

6. Local level analysis

Data quality and participation

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. Local authority level data quality measures are included in NHSD's NCMP reports. Additionally, local authority level data quality indicators are included in PHE's local authority NCMP and Child Obesity Profile: <http://fingertips.phe.org.uk/profile/national-child-measurement-programme>.

Participation rate and prevalence: published analyses of the earliest NCMP datasets (2006 to 2007 and 2007 to 2008) suggested 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 were more likely to be obese than those who did participate.

In more recent years, as participation rates have increased, the impact of selective opt-out on prevalence rates has declined. However, despite the potential impact appearing 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.

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.

Participation by sex: previous NCMP analysis has suggested that participation of girls in the NCMP may be lower than for boys, and this can differ by local authority. If this is the case, it may be due to a selective opt-out of overweight or obese girls (more than overweight or obese boys) from the NCMP measurements. 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 NHSD as the numbers of children eligible for measurement in each local authority are not available by sex.

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 3.

Sociodemographic variation: effect of ethnicity, deprivation and setting

Published NCMP analyses show that ethnicity, deprivation, and setting (for example urban/rural environment) may influence the 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 analyses that explore these factors.

Local authorities can also use their local knowledge to determine the extent to which 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 England, 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.

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. The coding of ethnicity in the NCMP dataset is not complete, but since 2008 to 2009 more than 80% of records have ethnicity coding. Users of the data are advised to check completeness of ethnicity coding before performing any analysis. 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 and Child

Obesity Profile: <https://fingertips.phe.org.uk/profile/national-child-measurement-programme>.

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. If other indicators are required for analysis these can be assigned to individual children using the LSOA of residence or school LSOA.

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 and Child Obesity 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 patterns and trends slide set available at: www.gov.uk/guidance/phe-data-and-analysis-tools#obesity-diet-and-physical-activity.

Setting: 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, but 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.

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](#). 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 p.15).

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

Combining data for school years

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.

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 are measured.

Taking account of the sex ratio of children measured

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.

The 2006 to 2007 NCMP report noted differences between the sex ratios of children measured in different areas. This pattern appears to have continued in more recent years. 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.

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. In the case of single sex schools, analysis could utilise published England figures for boys and girls for comparison which are available in the NHSD annual report.

Checking to see whether a change or difference in prevalence is meaningful

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.

The published data from NHSD and PHE provide 95% confidence limits for local authority prevalence rates. The method used for confidence intervals and to test for statistical significance is detailed in the Appendices report in each NCMP annual publication by NHSD, <https://digital.nhs.uk/data-and-information/publications/statistical/national-child-measurement-programme>.

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 5.

If examining a reported change in rate for statistical significance, the approach recommended by Altman et al. should be used. See Appendix 5, p.40.

Monitoring change over time

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 5.

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.^{4,5} Mean BMI z score can be calculated using 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.

Since the 2012 to 2013 NCMP, the datasets are likely to contain a large cohort of Year 6 children who were previously measured in Reception from 2006 to 2007 onwards. As identifiable data is 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 NCMP data for Year 6 children with child measurements for the same children taken in Reception year. This could be done using data held locally which contains a child

identifier. Some local authorities may even be able to 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 to 2006 National Childhood Obesity Dataset (NCOD).

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 for records submitted with NHS number when the Reception year children measured in 2013 to 2014 are remeasured when they reach Year 6 in 2019 to 2020.

The best methods for comparing Reception to Year 6 measurements of the same children

Although it is interesting to compare changes in these cohorts of children at a population level (for example comparing obesity prevalence in 2009 to 2010 for children of Reception year with obesity prevalence in 2016 to 2017 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 Reception and Year 6).

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.

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 datasets 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.^e

Where identifiable data is 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.

^e 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.

In March 2017 PHE published a report examining how weight status tracks over time in individual children during primary school using the NCMP data from 4 local authorities. The report also provides an analytical protocol and step by step guidance on undertaking local level tracking analyses. By following the guidance, local data findings can be directly compared with those from the 4 local authorities involved in the PHE study. www.gov.uk/government/publications/weight-change-in-primary-school-age-children

Some other local examples of such work produced by local authorities have been published, by Hull PCT,⁶ Southampton PCT⁷ and South Gloucestershire.⁸

Small area analysis

The NCMP dataset contains a very large number of records at national level, but, at local level or in certain subgroups, the number may be much smaller. Obesity prevalence figures are published annually at national and local authority level using a single year of NCMP data, but many practitioners also 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. Table 1 shows the average number of children measured for the NCMP at the various levels of aggregation used in England.

Table 1: Levels of aggregation of NCMP data and average number of children measured per school year group

Common geographies:	Upper tier LA	Lower tier LA	CCG	MSOA	Ward*	School**	LSOA
Number of areas	152	326	207	6,781	7,445	17,038	32,844
Average population	360,000	170,000	270,000	8,000	7,400	n/a [†]	1,600
Average number of children measured for NCMP (2016/17)	3,901	1,819	3,041	93	88	39	18

* Electoral wards

** Includes eligible state schools only (not including participating independent or special schools)

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

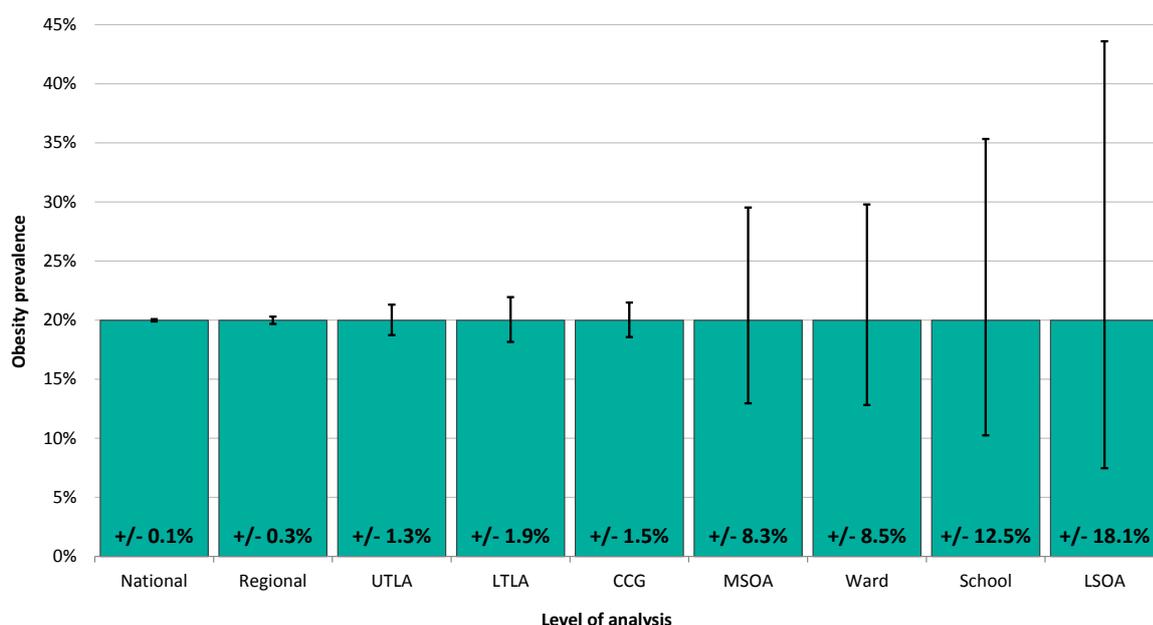
On average, upper tier local authorities measure around 3,900 children per age group (Reception or Year 6) per year for the NCMP. This equates to around 1,800 children measured per lower tier local authority. CCGs each provide on average around 3,000 children measured per age group. Statistics based on such large numbers of child measurements are likely to provide relatively robust estimates of measures such as obesity prevalence.

Sub-local authority level analysis is possible using NCMP data. Analysis can be performed using school or LSOA as the unit of analysis, or by assigning an MSOA or ward code from the LSOA code. Such analysis needs to be performed and interpreted with caution. Prevalence figures for sub-local authority populations are based on small numbers and so are subject to a higher degree of natural variation. On average at MSOA level there are only around 90 children measured per year in each NCMP age group, with only around 20 children of each age group measured at LSOA level. 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 5).

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

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

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



The estimates presented in the chart above are based on the national prevalence of obesity among Year 6 children in one year of NCMP measurements (2016 to 2017).

Confidence limits around obesity prevalence estimates at national, regional, local authority and even CCG level tend to be of a reasonable size (under +/-2%). However for the available English geographic areas below this level (such as electoral wards, MSOAs or LSOAs) confidence limits around obesity prevalence figures increase dramatically from +/-8% up to +/-18%.

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

Other 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 (see p.21).

Methods for increasing the robustness of NCMP analysis for small populations

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, MSOA, and CCG level using three years of data at: www.gov.uk/government/statistics/child-obesity-and-excess-weight-small-area-level-data and ward level data is also available in the [PHE Local Health tool](#). For smaller geographies (for example LSOAs or small schools) even 4 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.

In some situations it may not be suitable or possible to combine NCMP data from different years of measurement. This may be because the purpose of the analysis is to monitor change over time within small populations, or because NCMP data for previous years may have a large proportion of missing data for LSOA of residence. Where this is the case it may be necessary to combine areas or populations to create larger groups or clusters instead of, or in addition to, combining data from different years of measurement. Such clustering is usually based on locality, for example combining neighbouring areas to create larger geographies. Such areas

could be created on an ad hoc basis, depending on the area of interest, but could also be based on existing areas, such as children's centre areas.

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

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

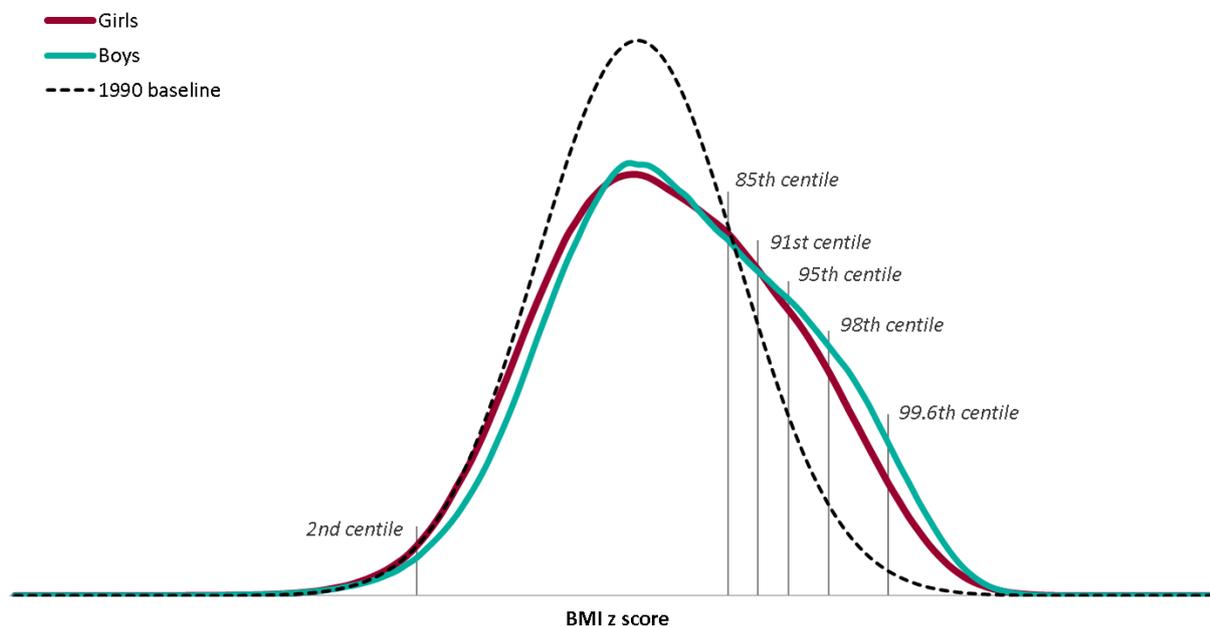
An alternative to combining geographic areas is to combine data for populations with similar characteristics such as ethnicity or socioeconomic status (using the IMD) or by using data from population stratification systems such as ONS-AC (see p.20).

When producing estimates of prevalence by child BMI category from combining data for small geographies or population characteristics, it is advised that ideally a sample of at least 50 children is used as the denominator. This is to ensure a large enough sample size to provide a reliable picture of prevalence for the population being examined.

BMI distribution

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, obese or severely obese children. Figure 2 provides an example of the distribution of BMI z scores within the NCMP dataset for children in Year 6, compared to the 1990 baseline.

Figure 2: Distribution of BMI z score for children in Year 6



This chart was produced using 2016 to 2017 NCMP data

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.

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.

Appendix 1: Example school feedback letter



Public Health
England

Protecting and improving the nation's health

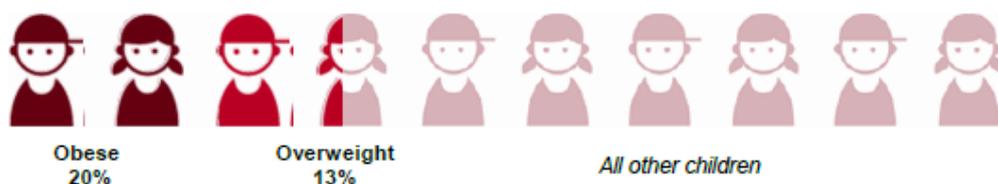
National Child Measurement Programme

[YEAR] summary of results for 123456 - Example Primary School

These results are based, where data is available, on the 20XX/XX, 20XX/XX and 20XX/XX years combined to give a sample large enough to produce reliable estimates of overweight and obesity.

Nationally, schools achieve an average participation rate of 96% in Reception and 94% in Year 6. In your school **XX.X% of children in Reception** and **XX.X% of children in Year 6** were measured. High participation rates (over 90%) ensure that the estimates of overweight and obese children in your school are reliable. Participation rates below 90% may result in unreliable estimates.

Proportion of **Reception** children in your school who were overweight or obese NCMP 20XX/XX, 20XX/XX, and 20XX/XX combined



The proportion of Reception children who were overweight or obese (33.5%) is higher in your school than in most other schools across England (22%). 27% of Reception children were overweight or obese in Example local authority.

Proportion of **Year 6** children in your school who were overweight or obese NCMP 20XX/XX, 20XX/XX, and 20XX/XX combined



The proportion of Year 6 children who were overweight or obese (35.9%) is statistically similar compared to other schools across England (34%). 42% of Year 6 children were overweight or obese in Example local authority.

In any population of children of all ages, it would be expected that 10% would be overweight and a further 5% would be obese. However current levels across England are far higher.

Appendix 2: NCMP enhanced extract data fields

The NCMP enhanced datasets are 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. The field names included in the data table are given below:

Table A1: Local authority enhanced extract data fields

Field	Description
SubmitterLocalAuthorityCode	Code of the local authority that submitted the data
SubmitterLocalAuthorityName	Name of the local authority that submitted the data
SubmitterRegionCode	Code of the region of the local authority that submitted the data
SubmitterRegionName	Name of the region 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
HeightPScore	Height centile
Weight	Pupil weight
WeightZScore	The number of standard deviations the weight is above or below the mean
WeightPScore	Weight centile
Bmi	Pupil body mass index
BmiZScore	The number of standard deviations the BMI is above or below the mean
ClinicalBMIcategory	Weight status based on clinical thresholds
GroupedClinicalBMIcategory	Weight status based on clinical thresholds grouped to combine overweight and very overweight
BmiCentile(PScore)	BMI centile
PopulationBMIcategory	Weight status based on population thresholds
GroupedPopulationBMIcategory	Weight status based on population thresholds grouped to combine overweight and very overweight

Field	Description
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	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
PupilTier1LocalAuthorityCode	Upper tier local authority code based on pupil postcode
PupilTier1LocalAuthorityName	Upper tier local authority name based on pupil postcode
PupilTier2LocalAuthorityCode	Lower tier local authority code based on pupil postcode
PupilTier2LocalAuthorityName	Lower tier local authority name based on pupil postcode
PupilIndexOfMultipleDeprivationScore	Relative measure of deprivation based on pupil postcode (IMD2015)
PupilIndexOfMultipleDeprivationDecile	Relative measure of deprivation based on pupil postcode (IMD2015)
PupilOnsSupergroupCode	ONS classification based on population characteristics (2011 ONS Area Classification)
PupilOnsSupergroupDescription	ONS classification based on population characteristics (2011 ONS Area Classification)

Field	Description
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
SchoolTier1LocalAuthorityCode	Upper tier local authority code based on school postcode
SchoolTier1LocalAuthorityName	Upper tier local authority name based on school postcode
SchoolTier2LocalAuthorityCode	Lower tier local authority code based on school postcode
SchoolTier2LocalAuthorityName	Lower tier local authority name based on school postcode
SchoolIndexOfMultipleDeprivationScore	Relative measure of deprivation based on school postcode (IMD2015)
SchoolIndexOfMultipleDeprivationDecile	Relative measure of deprivation based on school postcode (IMD2015)
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)
DQFlagCode	Data quality flag code
DQFlagDescription	Data quality flag description

Appendix 3: Suggested data quality checks at local level

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.

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 or half number.

The NCMP upload process is continually being improved to include additional validation checks and further validation is now undertaken by NHSD prior to the data being released. A document outlining all the NHSD validation checks is available on the NCMP IT system webpage at <https://digital.nhs.uk/services/national-child-measurement-programme/it-system#related-information> however, some data quality issues may remain despite these additional checks.

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 NHSD quoting 'NCMP' in the subject heading (enquiries@nhsdigital.nhs.uk).

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

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 is 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 was published.

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.

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 incorrectly been assigned to one of the two institutions.

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

Duplicate pupils

Duplicate pupils assigned to the same school will be blocked by the NCMP IT system.

In some situations duplicate children are allowed, for example if they have moved schools during the year and were measured in each school (ie the duplicate records were submitted from 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.

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 when pupils had been inaccurately coded to schools; for example, in some cases a group of Reception pupils had been incorrectly added to a similarly named junior school as well as to the correct infant school.

The most common reason is 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.

Rounded records

The NCMP operational guidance states that height and weight measurements should 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.⁹

NHSD 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 may still have a high proportion of rounded records in their dataset if they did not make amendments and resubmit.

A summary of the proportion of rounded records for every local authority has been provided within NHSD's NCMP reports and on the local authority NCMP and Child Obesity Profile <http://fingertips.phe.org.uk/profile/national-child-measurement-programme>.

Day of measurement

The enhanced NCMP dataset contains a field showing day in the week of measurement.

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 the BMI score.

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

Appendix 4: BMI thresholds, z scores, and p scores

The height, weight and BMI of children change as they 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).

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). 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. The p score describes the BMI centile value; for example a BMI p score of 0.85 means the child's BMI is on the 85th centile.

As an 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 +1.645 (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

The NCMP published prevalence data uses 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. Table A3-1 shows the p scores that should be used to classify child BMI into categories to calculate prevalence for population monitoring purposes.

Table A3-1: P scores used for population monitoring BMI classification

Population monitoring BMI centile category	BMI centile score (p-score)	BMI centile
Severe obesity	≥0.996	≥99.6 th
Obese	≥0.95	≥95 th
Overweight	≥0.85	≥85 th
Healthy weight	>0.02 to <0.85	>2 nd to <85 th
Underweight	≤0.02	≤2 nd
Very thin	≤0.004	≤0.4 th

It is important to note that the 2nd, 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.

In clinical settings or when monitoring the BMI of individual children, z-scores are used to classify individual children as underweight, healthy, overweight or very overweight (obese) taking into account the expected variation in BMI by age and sex, see table A3-2 below. The NCMP parental feedback letters issued by local authorities use these clinical cut-offs to assign children to a BMI classification.

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 (z-score). These actually equate to the 2.275th, 90.879th and 97.725th centiles when rounded to three decimal places, although they are usually referred to as the 2nd, 91st and 98th centiles, but the z-scores must be used to assign the pupils to the correct category.

Table A3-2: Z scores used for clinical BMI classification

Clinical BMI centile category	BMI Standard Deviation Score (z-score)	Approximated BMI centile line on growth chart
Severe obesity	≥8/3	≥99.6 th
Very overweight (clinical obesity)	≥6/3	≥98 th
Overweight	≥4/3	≥91 st
Healthy weight	>-6/3 to <4/3	>2 nd to <91 st
Underweight (Low BMI)	≤-6/3	≤2 nd
Very thin	≤-8/3	≤0.4 th

Note the right hand column in this table shows the approximated centile line on the growth reference. Centiles (p-scores) should not be used to calculate prevalence for individual children using the clinical cut-offs as they are approximations and will misclassify some children. The z-scores shown in the middle column should be used to assign a BMI clinical category.

Underweight prevalence figures using both population monitoring and clinical thresholds are each presented as having been derived using the 2nd centile,

however, those based on population monitoring thresholds use the exact 2nd centile. Those derived using clinical thresholds use -2 standard deviations (the 2.275th centile shown to three decimal places). For a given population, prevalence of underweight will therefore differ slightly depending on whether the population monitoring or clinical thresholds are used.

Severe obesity prevalence figures using both population monitoring and clinical thresholds are each presented as having been derived using the 99.6th centile. The exact 99.6th centile (p score of 0.996) is used for population monitoring whereas any prevalence figures derived using clinical thresholds use 8/3 standard deviations. For a given population, prevalence of severe obesity will therefore differ slightly depending on whether the population monitoring or clinical thresholds are used.

When presenting prevalence figures based on the 85th and 95th 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.

If users need to calculate 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.

LMS Growth can be used to calculate both the z scores and the centile for child measurements. The two can also be converted within MS Excel, using the 'NORM.S.DIST' and 'NORM.S.INV' functions.

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 minimal amount. However, these differences do not have any noticeable impact on prevalence figures.

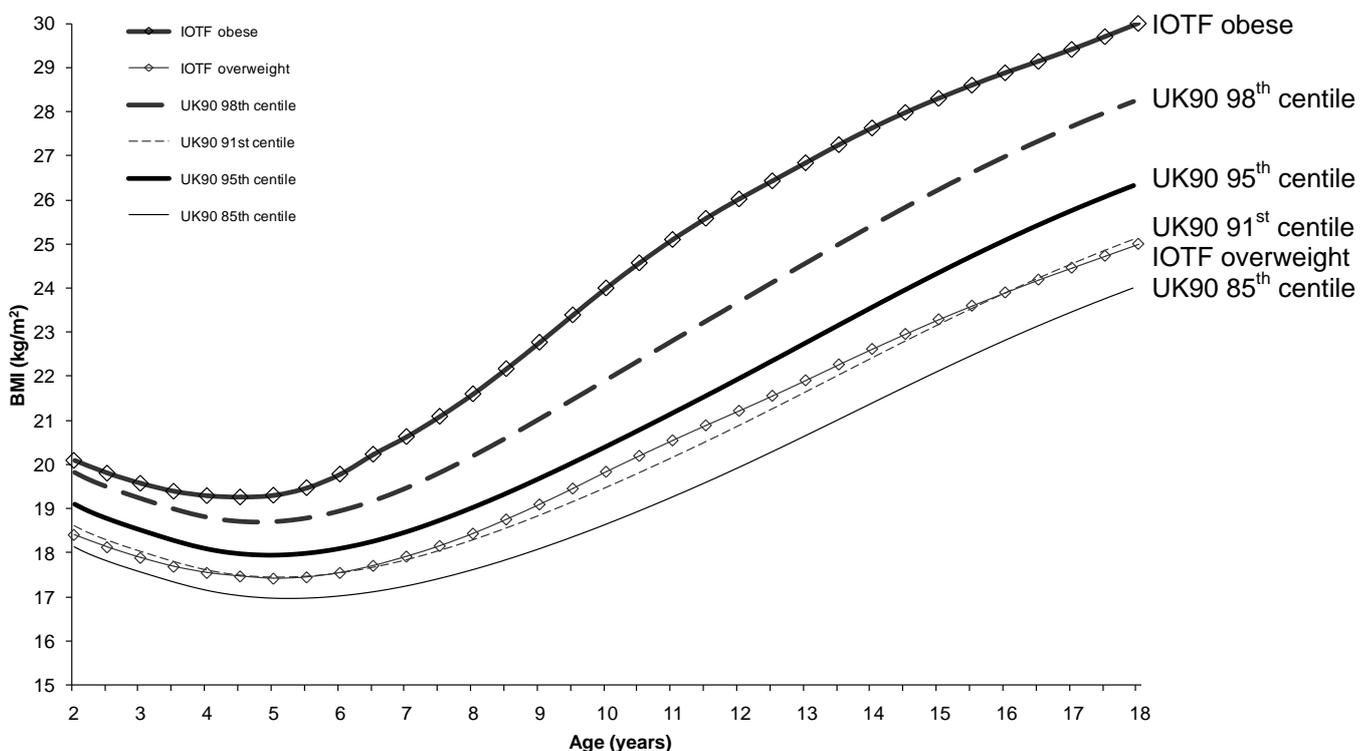
Other thresholds for defining children's BMI status

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 modelling¹⁰ and for child obesity prevalence figures from the Millennium Cohort Study.¹¹

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).

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. More information is available in the [Simple Guide to Classifying Body Mass Index in Children](#).

Figure 3: Obesity and overweight thresholds for boys – UK90 and IOTF



Appendix 5: Calculating confidence limits

We recommend that 95% confidence intervals are calculated with the method described by Wilson¹² and Newcombe¹³ which is a good approximation of the exact method.

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 = $q = (1 - p)$

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

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

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

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

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.

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.

The approach outlined by Altman et al. in *Statistics with Confidence (edition 2)*¹⁴ 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.

This method is also provided as 'method 10' in the Newcombe paper 'Interval estimation for the difference between independent proportions: comparison of eleven methods'.¹⁵

References

1. The Local Authority (Public Health, Health and Wellbeing Boards and Health Scrutiny) Regulations 2013.
2. 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.
3. Cole TJ, Freeman JV, Preece MA. Bodymass index reference curves for the UK. *Arch Dis Child* 1995; 73: 25–29.
4. 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.
5. 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.
6. Porter M, Greene T, Taylor A. Childhood obesity in Hull: paired analysis. Hull: Hull PCT, December 2007.
7. King D. Child Growth Briefing Note; National Child Measurement Programme 2001/02 to 2008/09 Summary Report. Southampton: Southampton PCT, 2010.
8. 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).
9. Dinsdale H, Rutter H. National Child Measurement Programme: Detailed Analysis of the 2006/07 National Dataset. London: Department of Health, June 2008.
10. 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.
11. Hanson K, Joshi H (eds). Millennium Cohort Study Third Survey: A User's Guide to Initial Findings. London: Institute of Education, October 2008.
12. Wilson EB. Probable inference, the law of succession, and statistical inference. *J Am Stat Assoc* 1927; 22:209–12.
13. Newcombe RG. Two-sided confidence intervals for the single proportion: comparison of seven methods. *Stat Med* 1998; 17:857–72.
14. Altman DG, Machin D, Bryant TN, Gardner MJ. *Statistics with Confidence*, 2nd edn. London; BMJ books, 2000: 49.
15. Newcombe RG. Interval estimation for the difference between independent proportions: comparison of eleven methods. *Stat Med* 1998; 17:873–90.

Reader information

Title	NCMP Guidance for analysis and data sharing
Author(s)	Caroline Hancock, Vicky Copley
Reviewer(s)	Paul Niblett, NHS Digital Robert Dobson, NHS Digital Sam Widdowfield, NHS Digital Shireen Mathrani, Public Health England Craig Timpson, Public Health England Alison Gahagan, Public Health England Jade Clark, Public Health England
Publication date	September 2018
Target audience	Local authority and Public Health England public health intelligence analysts
Description	Information and guidance for data sharing and analysis by local authorities and Public Health England of the National Child Measurement Programme datasets.
How to cite	Hancock C, Copley V. NCMP Guidance for analysis and data sharing. London: Public Health England, Obesity Risk Factors Intelligence, 2018.
Contact	ncmp@phe.gov.uk
Electronic location	www.gov.uk/government/publications/national-child-measurement-programme-data-sharing-and-analysis
Copyright	© Crown copyright 2018