National Neonatal Audit Programme (NNAP) Methodology and Statistical Analysis Plan

Published October 2025

V3.1 Published by RCPCH October 2025.

The Royal College of Paediatrics and Child Health is a registered charity in England and Wales (1057744) and in Scotland (SCO38299)

Cite as: National Neonatal Audit Programme Methodology and Statistical Analysis Plan v3.1, National Neonatal Audit Programme, October 2025. RCPCH: London.

Contents

Introduction

The purpose of this document is to provide a detailed methodological overview of the data and measure development steps and analyses contained in the National Neonatal Audit Programme (NNAP) annual reports.

The document is not designed to relate specifically to any one report year but will be updated annually to ensure that it is in line with the latest NNAP methodology.

The document does not contain measure-specific methodologies (e.g. inclusion criteria, attribution, derivation tables), as these are covered in detail in the year-specific measures guides, which can be found at https://www.rcpch.ac.uk/work-we-do/quality-improvement-patient-safety/national-neonatal-audit-programme-nnap/measures.

The NNAP Annual Report contains key messages and recommendations by theme, case studies, support and resources for healthcare improvement, information for parents and carers, and the future direction of the Programme. The latest version is available to download at: www.rcpch.ac.uk/nnap.

Full year-specific results at unit level, interactive reporting tools and unit posters are available on NNAP Online at: https://nnap.rcpch.ac.uk/.

From 2023, the NNAP also publishes quarterly updated results in the public domain via the frequent reporting dashboard, available at: https://www.rcpch.ac.uk/resources/nnap-data-dashboard

1. Data flow

The RCPCH receives the data contained in the NNAP dataset directly from System C, provider of the BadgerNet clinical system. The NNAP team processes data that is solely required for care quality and service improvement in relation to the aims and scope of the NNAP.

Figure 1 - Data flow describes this data flow and the feedback loop which disseminates results and recommendations to neonatal units, networks, and the wider system to inform and promote quality improvement.

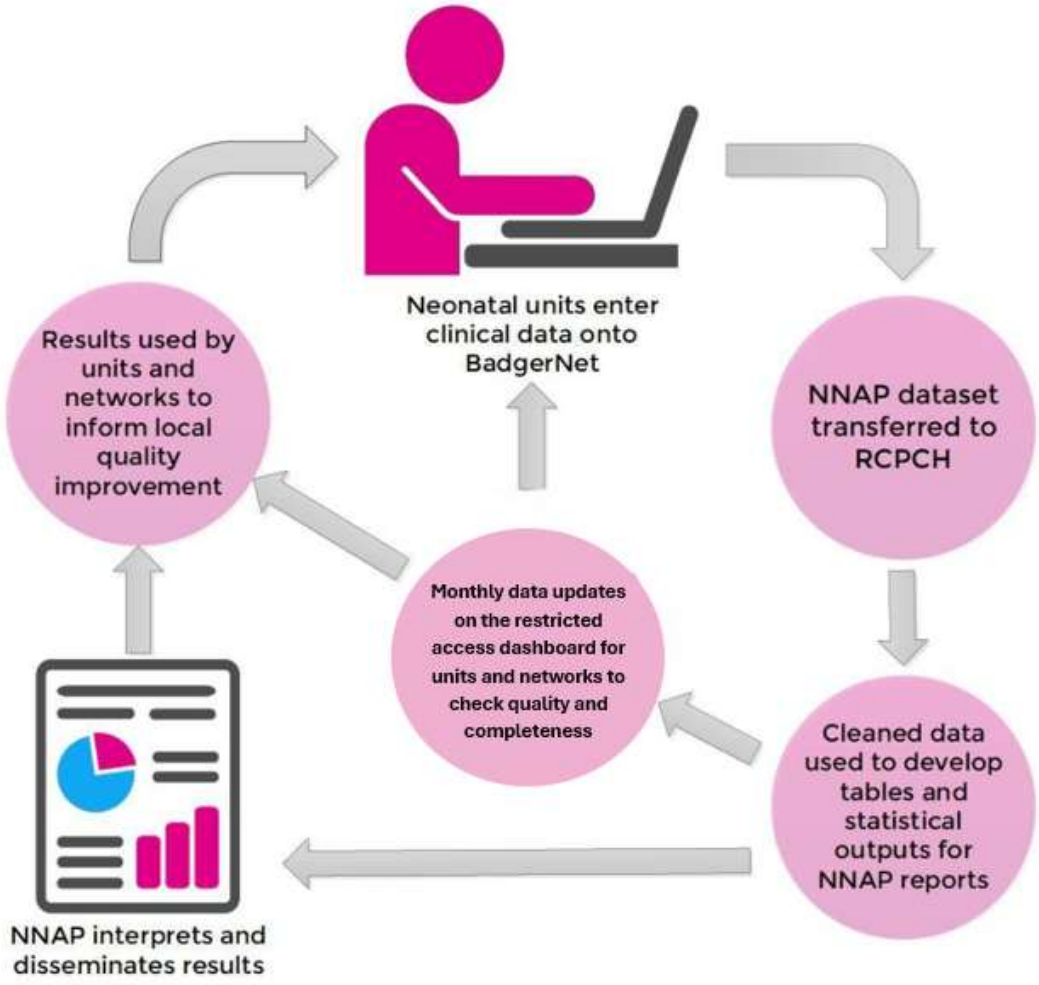

Figure 1 - Data flow

A full data flow diagram detailing data flows, data controllership, legal bases, storage, processing and outputs is available at: https://www.rcpch.ac.uk/nnap-data-flow-methodology

a. Live database linkage

Neonatal unit staff enter data into the BadgerNet platform (or systems that can interface with the BadgerNet database) as part of routine care whilst a baby is present on the neonatal unit and this data is stored on the BadgerNet database by System C.

Agreed data items contained within the NNAP dataset are synchronised to an NNAP database hosted by the RCPCH. This is done via SQL synchronisation between the System C Azure NNAP Database patient data source and the RCPCH Azure database server over a TLS1.2 encrypted SQL<-> Synchronisation platform. This is created and operated by Microsoft on the Azure platform and is designed for secure data transmission and storage requirements.

As of 2022, the MESH mechanism to remove data about patients who have opted out at a national level via the National data opt-out (England only) is no longer applied before the dataset is transferred to the RCPCH, although patients can still directly opt out of inclusion in the NNAP.

b. Static database copies

Before any analysis is conducted on the data contained in the Live database hosted on the RCPCH Azure environment, static, pseudonymised versions of the database need to be copied on the Azure environment. This process allows a snapshot of the database to be taken at an agreed time, ensuring that analysis conducted on the data is replicable as the underlying data in that static version of the database is not updated and therefore does not change. At this point the database is also pseudonymised by the database administrator.

These static pseudonymised copies of the database are taken monthly, with a copy for each annual report saved for future reference.

c. Pseudonymisation process

The RCPCH has approval to receive patient identifiable data without consent for England and Wales under Section 251 of the National Health Service Act 2006, and for Scotland via approval from the Health and Social Care Public Benefit and Privacy Panel, however patient identifiable data is pseudonymised before analysis is conducted by the NNAP team to reduce the risk of data breaches.

When static database copies are made the database administrator uses SQL Server Management Studio to connect to the database and convert the identifiable and sensitive fields to pseudonymised versions of those fields. This is done by running a T-SQL script within SQL Server Management Studio which recreates the relevant data tables

whilst implementing a combination of date transformations and hashing. The original tables are then removed from the database. A full list of fields that are pseudonymised can be seen in Table 1 - Pseudonymised fields.

d. Data analysis

Once the static pseudonymised copy of the database has been prepared, this database is queried using the statistical software Stata and R, and analysis conducted on it. After each session the memory of the software is cleared, and no patient-level data is retained on the RCPCH local servers.

The NNAP dataset is defined in the NNAP data dictionary, available at: https://www.rcpch.ac.uk/work-we-do/quality-improvement-patient-safety/national-neonatal-audit-programme-nnap/measures

Table 1 - Pseudonymised fields

Table Name	Field Name	Field Description	Converted to
NNU2YearFollowup	AssessmentDate	the date and time of the follow up assessment	minutes from birth
NNU2YearFollowup	bailey_notes	free text field used in two year follow up questionnaire	generic text field
NNU2YearFollowup	date_death	the date and time of death	minutes from birth
NNU2YearFollowup	gastro_special_diet_text	free text field used in two year follow up questionnaire	generic text field
NNU2YearFollowup	griffiths_notes	free text field used in two year follow up questionnaire	generic text field
NNU2YearFollowup	growth_length_measuredHowText	free text field used in two year follow up questionnaire	generic text field
NNU2YearFollowup	growth_measurements_date	the date and time that growth measurements were taken	minutes from birth
NNU2YearFollowup	neurological_diagnosis_text	free text field used in two year follow up questionnaire	generic text field
NNU2YearFollowup	schedulegrowing_notes	free text field used in two year follow up questionnaire	generic text field
NNUCranialUSS	NoteTime	the date and time of the cranial ultrasound scan	minutes from birth, month, year
NNUDaySum	CareDate	the date of the care day	minutes from birth, month, year
NNUEpisodes	AdmitTemperatureTime	the date and time that the temperature was taken	minutes from birth, month, year
NNUEpisodes	AdmitTime	admission date and time of the episode	minutes from birth, month, year
NNUEpisodes	BirthDateMother	the date of birth for the mother	minutes from birth, month, year
NNUEpisodes	BirthTime	birth date and time of the baby	minutes from birth, month, year, month 44 weeks PMA, year 44 weeks PMA
NNUEpisodes	DateTimeOfDeath	the date and time of death	minutes from birth, month, year
NNUEpisodes	DischTime	discharge date of the episode	minutes from birth, month, year
NNUEpisodes	FirstConsultationWithParents	the date and time of the first consultation with parents	minutes from birth, month, year
NNUEpisodes	NationalID_CHI	the national ID for Scotland	MD5#
NNUEpisodes	NationalID_HNC	the national ID for Northern Ireland	MD5#
NNUEpisodes	NationalID_NHS	the national ID for England and Wales	MD5#
NNUEpisodes	NationalIDBaby	NHS/CHI number for the baby	MD5#
NNUEpisodes	NationalIDMother	NHS/CHI number for the mother	MD5#
NNUEpisodes	PostCodeMother	The postcode of the mother at time of birth	LSOA
NNUEpisodes	SteroidsLastDose	the date and time that the last steroids dose was given	minutes from birth, month, year
NNUEpisodes	SteroidsFirstDose	the date and time that the first steroids dose was given	minutes from birth, month, year
NNUROP	NoteTime	the date and time of the ROP screening	minutes from birth, month, year
NNUSepsisScreen	NoteTime	the date and time of the sepsis screening	minutes from birth, month, year
NNUUNICEF	firstexpressiontime	the date and time of first milk expression	minutes from birth, month, year
NNUUNICEF	firstrecievedexpressedmilk	the date and time the baby first received expressed milk	minutes from birth, month, year

2. Case ascertainment and unit participation

In usual practice, every baby admitted to a participating neonatal unit entered on the BadgerNet patient record system is eligible for inclusion in NNAP. The audit therefore achieves 100% case ascertainment in the participating organisations, unless a parent or carer has chosen to opt out of having their baby's information submitted to the audit. Babies receiving special care alongside their mother in transitional care areas or postnatal wards can also be entered onto BadgerNet, but it is known that some units do not enter data for such babies. For this reason, NNAP's measures do not concentrate on care outside neonatal units.

All neonatal units in England and Wales associated with a delivery unit are eligible to take part, including special care baby units (SCUs), local neonatal units (LNUs) and neonatal intensive care units (NICUs). As of April 1st 2022, the NNAP has permission from the Privacy and Public Benefit Panel for Health and Social Care (HSC-PBPP) to include neonatal units in Scotland. As of April 1st 2023, the NNAP also has permission to include neonatal units in the Isle of Man.

Where there is a change in unit name, unit level or network configuration, the NNAP will apply the status as at the end of the data reporting year. For example, if the configuration of a network changes on 1 April 2021, 2021 data will be presented as per the network configuration on 31 December 2021. Details on unit participation can be found in the appendix of the most recent audit report.

Name changes and closures of units, networks and trusts are managed both on an *ad hoc* basis (units and trusts can inform the audit of changes by emailing nnap@rcpch.ac.uk at any time), and through an annual review process in which emails are sent to all participating neonatal networks and units requesting notification of any changes to unit level, name, closures or other changes.

3. Data quality and completeness

Although the NNAP uses routinely collected data from the BadgerNet database, this data can be and is still subject to data quality and completeness issues that affect the quality and reliability of the audit findings. Data quality issues include:

1) Data fields that are optional on BadgerNet but are required for audit analysis not being routinely completed by units.
2) Data fields that are optional on BadgerNet but are required for audit analysis being omitted by staff members on an *ad hoc* basis.
3) Data fields that are mandatory on BadgerNet but are not completed by units that use direct interfaces with the BadgerNet database, and thus bypass the BadgerNet mandatory field requirements.
4) Database or BadgerNet errors that cause contradictory or missing results.
5) Missing data resulting from the database pseudonymisation process.

Issues 4) and 5) listed above are mitigated by data cleaning and validation checks on the database. Due to the high degree of control that RCPCH and Clevermed have over the processes related to these steps, such issues are typically resolved quickly and completely. They therefore rarely have a negative impact on data quality and completeness.

a. Restricted access dashboard

Issues 1), 2) and 3) are mitigated through the sharing of a Restricted Access Dashboard with units, which contain summary statistics for each measure and a list of pseudonymised patient identifiers corresponding to patients that are missing data or that may require review, for each measure.

Restricted Access Dashboard updates occur monthly, and units are expected to use the data contained in it to identify and resolve data quality and completeness issues generally (but not exclusively) related to missing data.

The Restricted Access Dashboard is hosted on Power BI, from within the RCPCH Microsoft tenancy, to which only the unit's clinical lead, the NNAP Data Analyst and Manager, and selected staff members at the unit, have access.

Units can update their data records to reduce missing data at any point over the year and their data will be updated in the following month. Units are given until 31st March of each year to ensure that their data for the prior year is correct and complete.

The pseudonymised static version of the database that is used for the generation of annual report figures is copied after 31 March to ensure all units have had an opportunity to clean their data.

Units can also access and review their data in real-time using the BadgerNet system reporting tools.

Unit feedback and communications

The Restricted Access Dashboard also provides assurance around the implementation of the measures' criteria on the data, with units feeding back to the audit wherever divergence between the measures guide and the attribution, inclusion or classification of patients is observed by units. Feedback provided by units is vital to ensure that analyses are correct and appropriate, and for the continuous improvement of the NNAP methodology and reporting.

Whenever new comments are raised by units, they are recorded in the NNAP query log, which is maintained by the NNAP Data Analyst and periodically reviewed. The NNAP team aims to respond to all queries within 3 working days, and where this is not possible holding emails are sent to inform the sender of the expected delay.

4. Data cleaning and validation

a. Validation

Every time a monthly or annual analysis is undertaken a new static copy of the live database is taken and pseudonymised for this purpose. All pseudonymised static database copies undergo the following validation processes:

- The number of records in the episodic table is compared to the number contained in the last pseudonymised episodic table.
- Missing data percentages for each field in the episodic table are compared to the last pseudonymised episodic table.
- The total number of tables in the database is checked against the last pseudonymised database.
- The number of episodes removed by the data cleaning processes (below) is checked against the last pseudonymised database.
- Import and cleaning code is run on the new database to ensure that it runs without errors.

In addition to these validation checks, the following checks are conducted every time a new cut of the live database is taken and pseudonymised for annual report analysis:

- Total rows for each table are checked against a list of expected total rows for each table provided by Clevermed Ltd.
- Checks are conducted between the episodic table and all other tables to ensure that all records in the ad hoc tables and the daily summary table have matching episodes in the episodic table.
- A review of the NNAP location lookup file is conducted to ensure that all expected units are presented in the data, and that all records match an expected provider unit.
- A draft run of the measure output proportions and N is compared to compiled results from the previous year, and large changes are reviewed.

b. Data cleaning

Once validated, a data cleaning process is applied to the database tables before creating the NNAP dataset used to produce the data included in this report. The pseudonymised database is cleaned using the following steps:

1. Episodes with no recorded birth year are removed from the analysis.
2. Episodes with no recorded gestational age are removed from the analysis.
3. Episodes with no recorded admission time are removed from the analysis.

4. Episode numbers with negative values are removed from the analysis.
5. Duplicated episodes are removed from the analysis, prioritising those with a greater number of completed fields.
6. Babies who have no episodes with days on a neonatal unit are removed from the analysis.
7. Babies with a gestational age of less than 22 weeks at birth are removed from the analysis.

Figure 2 - Episodes eligible for NNAP analysis describes episodes removed from the dataset during the data cleaning process for the 2021 data year.

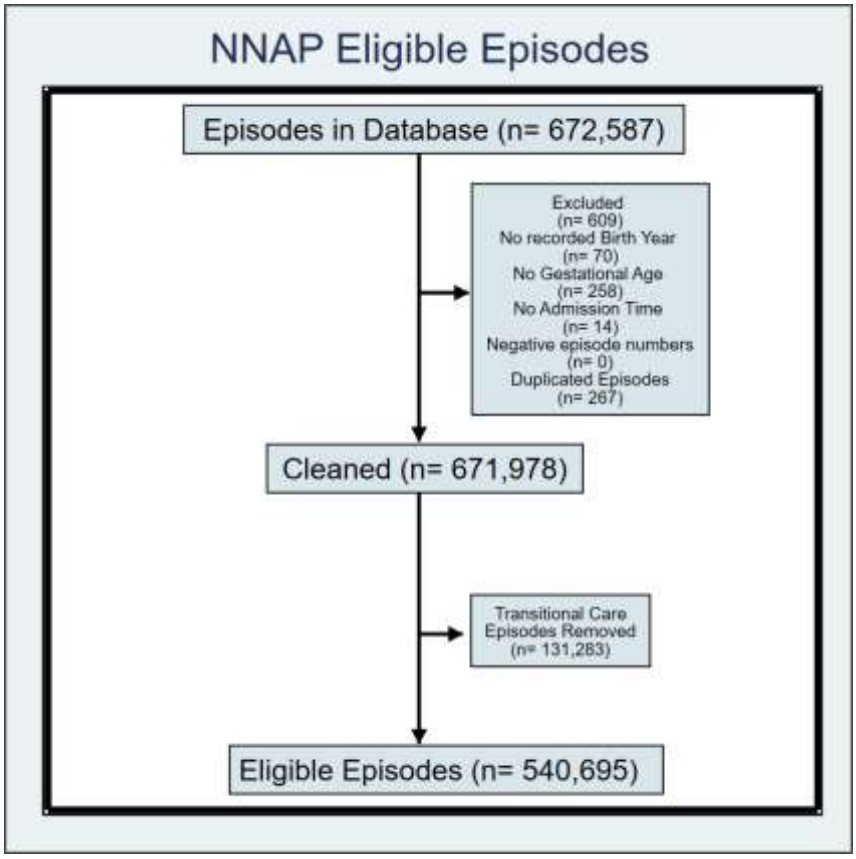

Figure 2 - Episodes eligible for NNAP analysis

Eligible episodes for each audit year include all neonatal admissions to NNAP units that were born between Jan 01 three years prior to the current data year, and 31 December of the current data year. For example, the 2021 data year used babies born between 01 January 2018 – 31 December 2021. It is important to note that not all babies in the eligible episodes are included in the measure analysis, which is dictated by measure-specific inclusion criteria. For details on inclusion criteria for each measure, please review the most recent version of the NNAP measures guide.

c. Location lookup matching

Alongside the data cleaning steps, it is necessary to match data from the pseudonymised database to NNAP unit names, codes, unit levels and networks. This is because:

- BadgerNet does not contain accurate data on neonatal unit levels (NICU, LNU, SCU).
- BadgerNet does not contain accurate data on which units are associated with which Operational Delivery Networks (ODNs).
- BadgerNet contains divergent names for units, with multiple NHS location codes and unit names being associated with one single unit.
- BadgerNet does not know which units are included in the NNAP and which are excluded.

For these reasons it is necessary to match the data in the database with NNAP unit names, unit codes, ODNs and unit levels. This is done by linking the episodic table of the database to a location lookup that converts episodes' unit of birth and provider unit to distinct NNAP unit codes and ODNs. Matching of these codes is made first on BadgerNet's NHS ODS code field, then for unmatched units that process is repeated on BadgerNet's unit name field. Units that cannot be matched are assigned to the OTHER category.

As the parameters of the live database, set by System C and RCPCH, only allow the sharing of episodes directly associated with NNAP units, all provider codes should match known NNAP unit codes. A place of birth code may relate to a different hospital than that of care provider, therefore not all place of birth codes match known NNAP units.

To reduce this number of unmatched place of birth codes in the dataset, babies whose place of birth is listed as Home or Transit have their network of birth updated to the provider network of their earliest episode.

5. Outlier identification and management

Units and networks with extreme indicator statistics are identified as "outliers", using the national proportion as the comparison standard. The NNAP manages outlier status in line with its policy *Detection and management of outlier status*, available at: https://www.rcpch.ac.uk/nnap-data-flow-methodology#outlier-management-policy which is aligned with the HQIP guidance on the identification and management of outliers in England and Wales. Identification of "outliers", and sharing with CQC, is required of all NCAPOP projects. All neonatal services identified as outliers for one or more NNAP measure are notified according to the policy prior to publication of the results.

The definition of an extreme result is made in terms of the probability of obtaining the observed result by chance, if in fact performance at the given unit or network was identical to the national average. This follows similar logic to a statistical hypothesis test and is visualised through a funnel plot. The format of the funnel plot depends on whether casemix adjustment is carried out (see Section 8). Casemix-adjusted indicators produce risk differences between observed and expected proportions of babies, which is termed a "treatment effect". Unadjusted indicators produce a simple count of babies.

a. Unadjusted outlier identification

Funnel plots for *unadjusted indicators* show the count of babies with the process/outcome (the numerator) on the vertical axis and the count of babies eligible for the indicator (the denominator) on the horizontal axis. The added funnel lines are produced using a binomial formula. This linearly interpolates between anticipated counts of babies[i], to obtain consistently increasing or decreasing funnel curves. The funnel curves are then smoothed further for clarity, using local-linear regression via the Stata command 'npregress kernel'.

b. Adjusted outlier identification

Funnel plots for adjusted indicators relate the treatment effect and its confidence interval to a probability of being below the reference line of zero treatment effect (the national average, by definition). This probability is then converted to an equivalent z-score from a normal distribution for graphics. The z-score is shown on the vertical axis, and because the size of the unit/network is already accounted for, any other variable can be shown on

the horizontal axis or in the size of the markers for each unit/network. Please see Appendix 1 detailed description of the for more information.

c. Outlier categorisation

Upon completion of the funnel plot analysis, outliers are attributed to one of 5 categories, which correspond to:

1) If value <Lower 99.8% control limit then Alarm; or else
2) If value <Lower 95% control limit then Alert; or else
3) If value between Lower and Upper 95% control limits then Within expected range; or else
4) If value >Upper 95% control limit then Excellent; or else
5) If value >Upper 99.8% control limit then Outstanding

The 95% control limits, whether a funnel shape or parallel lines, relate to z-scores of ±1.96. The 99.8% control limits relate to z-scores of ±3.09.

The measures on which the full outlier management process is implemented are selected by the NNAP Methodology & Dataset Group with sign off by the NNAP Project Board; these can vary from year to year.

6. Public data sharing

The results of the NNAP are shared with the public through NNAP online, the Annual report and appendices, and through the NNAP Data Dashboard.

Details of the annual report and appendices are contained online and can be found here: https://www.rcpch.ac.uk/work-we-do/clinical-audits/nnap

The results of annual measure calculations and the associated statistical analyses are uploaded to the online portal NNAP Online. NNAP Online contains summary statistics for all units across all measures, as well as plots and charts related to the audit. The site publishes the data on the same day that the report is released.

Annual results for selected measures are used within reporting by the Care Quality Commission (CQC).

More frequent data reporting is available via the NNAP Data Dashboard, which shows time series charts for each of the 10 NNAP reporting metrics. These totals are displayed as rolling annual proportions and are updated quarterly.

The dashboard is created in Power BI and uses the most up to date data available, with all quarters since the last release of the Annual Report listed as provisional. The NNAP Data Dashboard and more information about it can be found here: https://www.rcpch.ac.uk/resources/nnap-data-dashboard

7. Managing small numbers in the NNAP

The NNAP considers the risk of disclosure on a measure-by-measure basis from a variety of methods resulting from the publication of results from small numbers of cases. Given the frequent occurrence of small numbers in unit level, annualised reporting, applying blanket masking to all cells would significantly reduce the utility of published NNAP results for improvement purposes. To minimise the risk of disclosure, the NNAP does not publish demographic data about the cohort of babies included in the audit, which would have the potential to be used alongside published data for the audit measures to aid identification of a patient. In addition to this, small number suppression is applied to measures considered to be sensitive; reporting that a clinical activity took place, such as consultation with parents within 24 hours, is not considered to be sensitive.

Table 2 - Measures subject to small number suppression shows a table of measures to which small number suppression is applied in the final results.

Table 2 - Measures subject to small number suppression

NNAP measure/data item	Application of masking to unit level data.
Parents on ward rounds	Mask sub-measure data when main measure data is <2 or sub-measure proportion is >99% or <1%.
Bloodstream infection	Mask unit results where the denominator is <3 and the infection rate is not 0%.
BPD	Mask unit results where the denominator is <3 (for 3 year rolling results and individual year results) and the BPD or death rate is not 0%.
Necrotising enterocolitis	In outputs, combine the columns "Death prior to discharge with no NEC" and "no NEC". Combine the missing data columns "Alive at discharge" and "Died before discharge". Mask unit rates where the denominator is <3 and the NEC rate is not 0%.
Breastmilk on the first two days of life	Mask unit results where the rate of feeding with any mother's milk is either 0% or 100%, or where the denominator is <3.
Breastmilk feeding at discharge home	Mask unit results where the rate of feeding with any mother's milk is either 0% or 100%, or where the denominator is <3.
Breastmilk feeding at 14 days of life	Mask unit results where the rate of feeding with any mother's milk is either 0% or 100%, or where the denominator is <3.
Brain Injury	Mask unit results where the proportion of IVH 3/4, cPVL or PHVD is 100%, or where the denominator is <3.

8. Accounting for case-mix differences

The aim of case-mix adjustment in the NNAP is to estimate how the risk of an outcome occurring for each unit or network would compare to the national risk, if the unit/network's case-mix reflected the national distribution of patient characteristics.

The key summary measure is the treatment effect, which is the difference between the observed proportion of babies with an outcome and the proportion expected based on the national case-mix.

From 2024, NNAP implements a form of direct standardisation for case-mix adjustment. Predicted probabilities of the outcome are generated for each baby using a logistic regression fitted to the national dataset. These predictions account for the national case-mix and are summed within each unit or network to produce the expected number of outcomes.

Observed and expected counts are then divided by the number of babies in the unit/network to obtain the observed and expected proportions, which are used to calculate the treatment effect.

The process is typically applied to the following measures: bronchopulmonary dysplasia, mortality, necrotising enterocolitis, bloodstream infection, and non-invasive ventilation, though future changes to this list of measures are anticipated in coming years. There are some differences in the case-mix variables that are used for these outcome indicators. These are described in detail in Appendix 1 detailed description of the .

a. Uncertainty in treatment effects

The uncertainty in both the observed and expected proportions is calculated by a statistical procedure called the bootstrap. This provides more accurate confidence intervals as outcomes approach 0%, which are used in the adjusted caterpillar plots (see Section 9a) and funnel plots (see Section 5).

The bootstrap method applied in this setting involves repeated resampling (with replacement) of the babies in each stratum at the unit/network of interest. This produces a collection of 10,000 estimates of the observed count, reflecting the uncertainty arising from the fact that the data comes from a limited number of eligible babies.

When a stratum contains 0 or *n* outcomes out of *n* babies in the observed data at the unit/network of interest, the bootstrap will not produce a valid probability distribution. For units with observed proportions of 0% or 100%, we draw proportions from the corresponding beta distribution, with 1 and *n*+1 parameters, multiply by *n* and round the answer to obtain a *parametric bootstrap* sample.

A second bootstrapping process creates a collection of 10,000 expected counts, reflecting uncertainty in the logistic regression. The 10,000 bootstrap expected estimates are randomly paired with the 10,000 bootstrap observed estimates (because these random processes are independent, except for negligible correlation arising from the unit/network being in the country).

In each pair of observed and expected bootstrap, we compute the treatment effect, providing a distribution of the uncertainty in the unit/network's treatment effect. A 95% (or other centile) confidence interval can be obtained from this by finding the 2.5 centile and the 97.5 centile in the 10,000 bootstrap estimates. The reason for the large number of bootstrap replicates is to obtain some stability and certainty in these "tails" of the distribution.

All case-mix adjustment procedures are programmed in Stata software.

9. Figures

b. Caterpillar plots

Caterpillar plots are generated for units and networks in all measures to help identify variations in the level of compliance between units, networks and unit levels.

In unadjusted caterpillar plots, each unit/network is represented by a vertical segment that extends from the lower to the upper confidence limit for the compliance proportion of the unit. These confidence limits are estimated as 2 standard errors above and below the unit/network proportion. The proportions (marked by discs) of compliance and estimated confidence limits are evaluated without shrinkage or imputation. Units are sorted according to their proportion for the given measure.

Catepillar plots are constructed in the same manner for care-mix adjusted measures, except that units/networks are sorted in the ascending order of their treatment effect scores, and confidence intervals are obtained via bootstrapping.

From 2019, distinct colours are used to indicate the level of the unit. Figure 3 - Antenatal Steroids - Units (2022 data year) displays the unit-level caterpillar plot for antenatal steroids for the 2021 data year. The colours used are indicated in the legend (SCU — level 1, LNU — level 2, and NICU — level 3). For the networks, caterpillar plots are constructed similarly, except that the abbreviated names of the networks are printed at the horizontal axis.

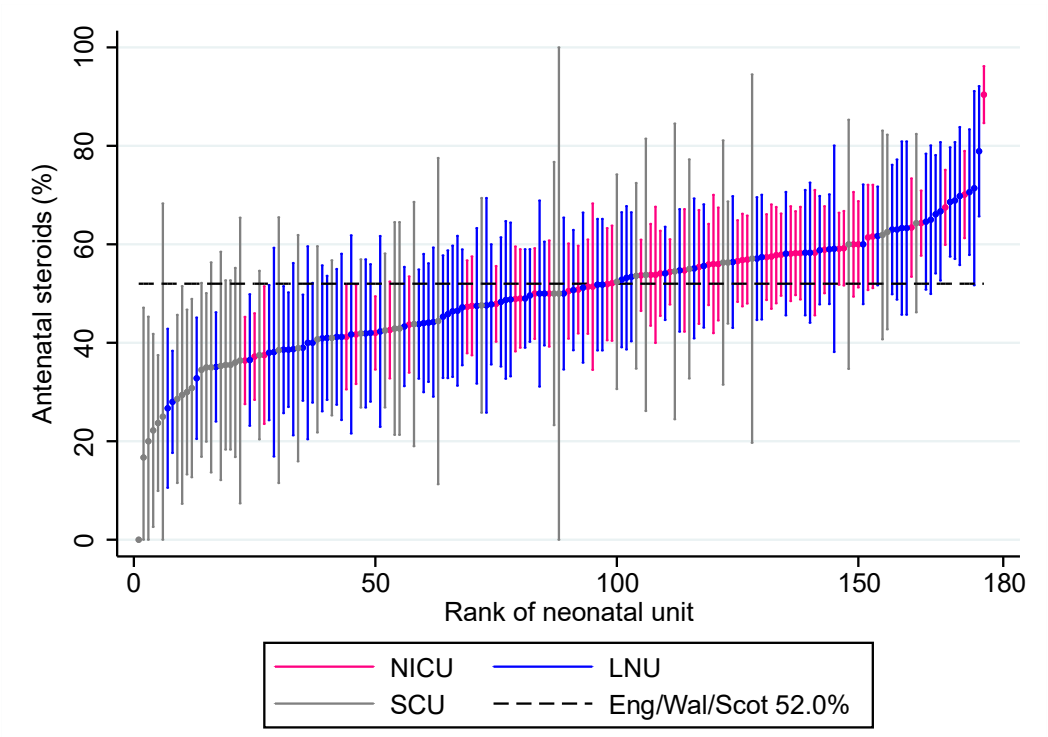

Figure 3 - Antenatal Steroids - Units (2022 data year)

c. Spine plots

A spine plot is a diagram in which the rates of several audit measures, for a unit or network, are presented in a single panel. It is a device to present the results of the audit in a compact form that facilitates comparisons of units, among themselves and with their networks, as well as networks with the whole domain. The segment extends from the lowest to the highest rate for the audit measure among the units in the diagram. Rates are scaled so that their network-level values are vertically aligned (marked by the dashes), and they fit the width available in the panel. The estimated rate for the unit is marked by a black disc, and the range in which the rate for the unit and the measure would be within the funnel limits is indicated by a grey strip. The network level rates are printed in a separate panel.

The black disc for a unit or network that is excellent or outstanding on a given audit item is located to the right of the grey strip and close to the right-hand limit of the segment, which corresponds to the highest rate among the plotted units or networks. The other extreme corresponds to poor performance. In a set of spine plots, we can identify units that have no weaknesses — all their discs are at the right-hand extremes; units that have no strengths — all their discs are at the left-hand extremes; units that have some strengths but also some weaknesses have their discs at either extreme of the segments. Units that are average or mediocre for all items have their discs close to the corresponding network-level rates and within the grey strips.

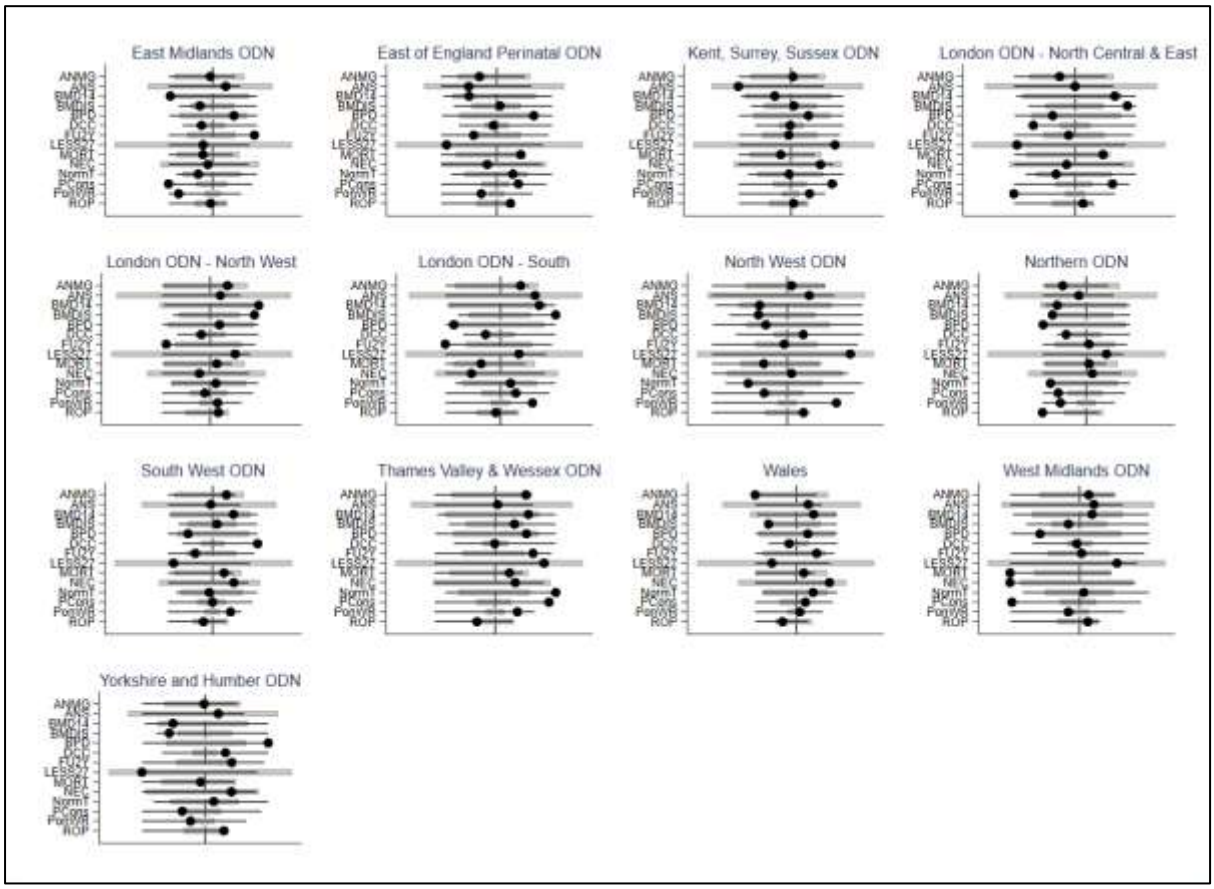

Figure 4 - Network spine plots (2020 data year)

d. Graphical summaries of breastfeeding

No single day on which breastfeeding status might be recorded can satisfactorily summarise a baby's breastfeeding over the entire stay in neonatal care. Moreover, no information about breastfeeding after discharge is available in the database. The rates of breastfeeding in units and networks on day 14 and on the day of discharge, presented in the standard tabular and graphical layouts, are supplemented, starting with the NNAP

2021 data year, by a graphical display in which breastfeeding throughout the stay, including the day of discharge, is represented.

The outcome variable is defined as a dichotomy for each day of stay in a neonatal unit. Breastfed on the day is coded as unity (1), and not breastfed as zero (0). Thus, a baby that stayed in care for 14 days has a breastfeeding record comprising 14 digits, each of them either 0 or 1. An example of such a record is 0011111 1100?00, with question mark '?' indicating missing entry (on day 12). For the definition of the population (the eligible babies and attribution), see Sections 3 and 4. In particular, the data span three years according to the date of birth.

Breastmilk feeding in a unit is graphically described by a colour-plot. Figure 5 – Thames Valley and Wessex breastfeeding (data year 2022) presents a set of examples, for units in Yorkshire and Humber Neonatal ODN. The axes of the plot are postnatal day (horizontal) and number of babies (vertical). An (imaginary) vertical line drawn on any given postnatal day comprises up to six segments of distinct colours, listed from bottom to top:

- Black at the bottom accounts for babies not fed at all (nil by mouth).
- Grey colour is for babies who were in a neonatal unit and were fed enterally but not breastfed on the day.
- Green colour is for babies who were breastfed on the day.
- Light green colour is for babies who have already been discharged and were breastfed on the day of discharge.
- Light grey represents babies who have already been discharged and were not breastfed on the day of discharge.
- Off-white colour at the top represents babies whose breastfeeding status is not known.

Plots are drawn for the first 90 postnatal days. They are drawn only for units with at least 70 babies in the data. If the entry for the day of discharge is missing, it is replaced by the breastfeeding status on the immediately preceding day. If the status is missing on both days, 'not breastfed' is imputed, unless the status is missing for many babies. Grey colour is reserved for missing breastfeeding status at discharge. Sub-sequences 0?0 and 1?1 are replaced by 000 and 111 respectively. That is, if breastfeeding status is missing for an isolated day and the states on the immediately preceding and following days coincide, then this status is imputed for the day with the missing record.
The colour-plots can be interpreted only by comparing the sizes of the areas (and their shapes) for the units (within a network) or of a unit with its network. It is therefore essential to present the colour-plots in sets of a selected network on a single page.

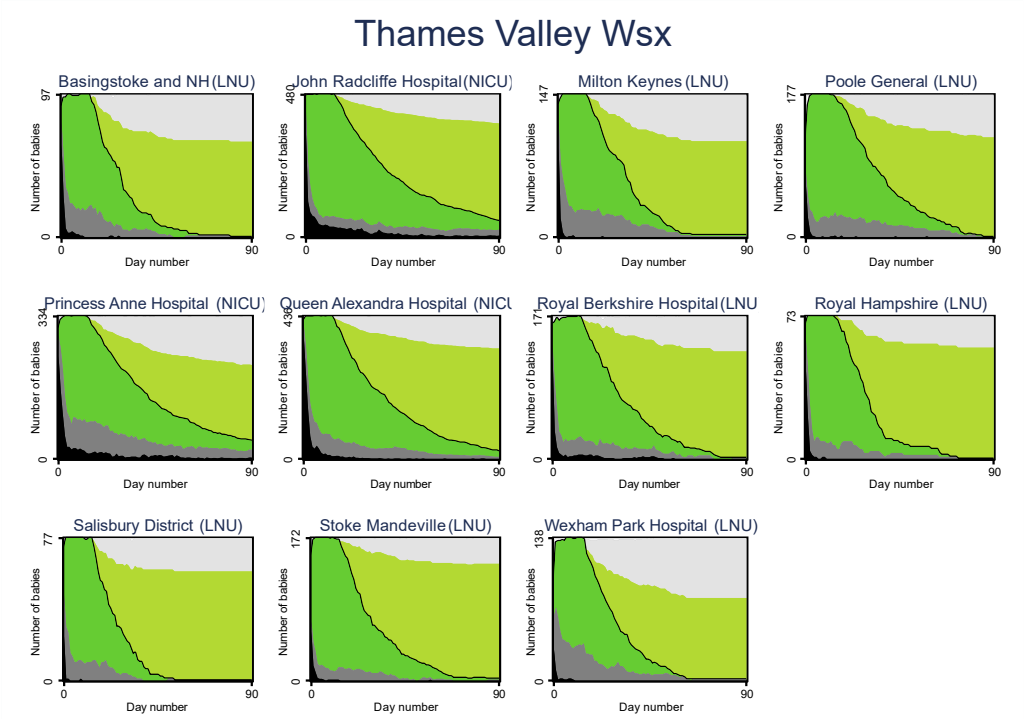

Figure 5 – Thames Valley and Wessex breastfeeding (data year 2022)

e. Nurse staffing diagrams

Nurse staffing plots designed to highlight annual shortfalls were added to the outputs for the 2021 data year. A unit's workloads, shortfalls and excesses for a year, or another period, are presented in a plot of the workloads and nurses available over time. The workloads are drawn by a solid black line. It has to be smoothed to avoid sudden changes and to better represent the overall trend. Orange patches under the workload curve represent shifts with shortfalls and light blue patched above the workload curve are for excesses. These patches also involve smoothing, so that periods covered by the orange colour may contain some shifts with excess of nurses, but shortfalls dominate. The year represented on the horizontal axis is split into months, indicated by vertical dashes.

Smoothing is implemented using the normal kernel with standard deviation set to 7, selected by trial and error. In this method, the value of the smoothed curve at any particular point (shift No.), referred to as the focal point, is calculated as the weighted average of the values (of the workload or the number of nurses present) for all the shifts, with weights equal to the density of the normal distribution (the normal curve) centred around the focal point, with standard deviation 7. Blue patches for shifts with (dominance of) excess of nurses are of no interest to the audit, but may be useful for the unit in analysing past scheduling and identifying where improvement could be made.

The dots at the bottom of the plot represent individual shifts, with the blue dots indicating sufficiently staffed shifts and orange dots indicating shifts with staffing shortfalls.

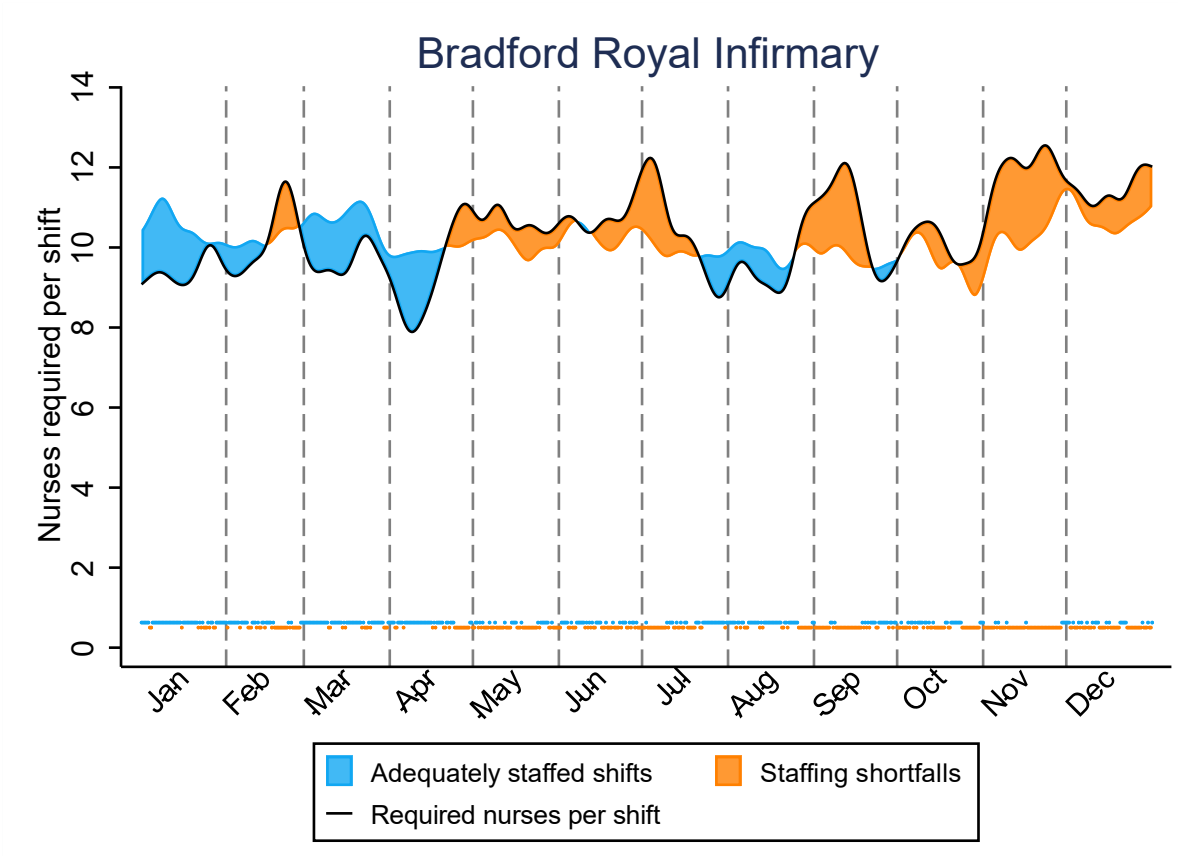

Figure 6 Bradford nurse staffing plots (2022 data year)

f. Other figures

In addition to the figures listed above, the audit regularly contains additional graphs and plots related to specifics in any given year. These include but are not limited to bar charts, longitudinal stick plots and scatter plots of measure proportions by unit.

Such figures are produced on an *ad hoc* basis with the goal of focusing attention on areas of interest pertaining to a specific audit year.

Funnel plots are described in Section 5.

Appendix 1 detailed description of the case-mix adjustment process

The observed and expected proportions of babies for each case-mix-adjusted indicator are intended to be balanced, which is to say that the case-mix should be comparable. This is done by defining strata, which are combinations of categories of case-mix variables. Continuous variables, such as birthweight, are divided into bands for this purpose, with approval of the NNAP the Methodology and Dataset Group.

The method used is sometimes known in the causal inference literature as "parametric g-formula"[ii]. It is closely related to the balancing approach used in previous years.

It involves training a logistic regression model on the complete country's data, using the strata as predictors and the outcome in question as the dependent variable. A predicted count of outcomes is generated for each unit and stratum, as the predicted risk, multiplied by the observed denominator. The counts for each unit are added to produce a case-mix adjusted expected count.

Uncertainty is quantified using the bootstrap, which is a computer-intensive and flexible approach, providing improved confidence intervals as the outcomes become rare and approach 0%. Conversion of the bootstrap results to caterpillar plots and outlier identification plots is described in sections 5, 8 and 9 above. In units where no, or all babies had the outcome in question, a bootstrap of the observed proportion does not provide a meaningful confidence interval. This is because in units with no, or all outcomes, the bootstrap cannot produce any in the resampling and thus the confidence interval is unreliable. In such units we replace the bootstrap output as follows:

1. Sample p from the posterior probability of the true unit proportion: a beta distribution with parameters (1, n+1) where n is the unit's denominator
2. Predict a count by multiplying p by n, and rounding to the nearest integer
3. Record this and repeat as many times as there are bootstrap replicates

a. Case-mix variables for each indicator

Each measure's analysis uses different case-mix adjustment variables, reflecting differences in the assumed causal relationships identified by the MADG. The number of case-mix variables selected also considers practical concerns, particularly the likelihood of

the outcome occurring. When outcomes are rare, statistical models may fail to converge, and using fewer case-mix variables can help mitigate this risk.

Table 3 - Case-mix variables for each adjusted measure

Measure	Type	Variables
NEC	Unit	GA; Small for GA
NEC	Network	GA; Small for GA
BPD	Unit	GA; Ethnicity; Deprivation; Sex; Birthweight; Problems pregnancy; Age mother; Medical problems mother; Mother smoke; Previous preg; Multiple
BPD	Network	GA; Ethnicity; Deprivation; Sex; Birthweight; Problems pregnancy; Age mother; Medical problems mother; Mother smoke; Previous preg; Multiple
Mortality	Network	GA; Ethnicity; Deprivation; Sex; Birthweight; Problems pregnancy; Age mother; Medical problems mother; Mother smoke; Previous preg; Multiple
Non-invasive ventilation	Unit	GA
Non-invasive ventilation	Network	GA
Bloodstream infection	Unit	GA; Unit level (inc. surgical/non-surgical NICU)
Bloodstream infection	Network	GA; Small for GA

b. Estimation of uncertainty

In the NNAP case-mix adjustment process, uncertainty is quantified using bootstrapping, which involves re-running the case-mix adjustment analysis 10,000 times on randomly selected subsets of the national dataset.

The treatment effect estimates from these bootstrap replications are then used to calculate (for each unit or network) the proportion of estimates below a reference line of zero (the national average treatment effect, by definition). This proportion represents a metric of how much evidence there is for a unit/network being different to the target of zero.

This proportion is then converted to an equivalent z-score from a normal distribution. The z-score is shown on the vertical axis of the outlier identification chart. Because the size of the unit/network is already accounted for in the z-score, the funnel plot's limits along the horizontal axis are flat lines.

c. History of changes to case-mix adjustment

Between 2018 and 2019, the statistical analysis process used by the audit to account for case-mix differences between units and networks was stratified matching. Due to complexities involved in the data flow change the report on the 2020 data temporarily used multilevel logistic regression to account for case-mix differences.

Since the report on 2021 data, the NNAP returned to the causal question using a more efficient and robust approach, called balancing, to account for case-mix differences. In balancing, each unit and network is compared to the national dataset. This was implemented in the R programming language in the reports for 2021-2023 data, weighting each baby's data so that the national dataset matches as closely as possible the profile of matching variables of the unit or network being compared.

The R code for weighting was developed by an external sub-contractor and implemented using some methods that were not peer-reviewed. From the 2024 report onward, all analysis of case-mix adjustment was brought in-house to use standard methods. To do this, the adjustment code was written anew in the same Stata software used by NNAP analysts. The current method for adjustment is different in some ways but has been shown to have very similar results to the weighting program in R.

The Stata code has been designed to be maintained by RCPCH staff with the intention of providing stability in the analytic methods for several years to come.

Appendix 2 Acronyms and Capitalisations

BPD	Bronchopulmonary dysplasia
BSI	Bloodstream Infection
CHI Number	Unique numbers allocated to registered users of public health services in Scotland.
CQC	Care Quality Commission
GA	Gestational Age
HNC Number	Unique numbers allocated to registered users of public health services in Northern Ireland.
HQIP	The Healthcare Quality Improvement Partnership (HQIP) aims to promote quality improvement in patient outcomes, and in particular, to increase the impact that clinical audit, outcome review programmes and registries have on healthcare quality in England and Wales. HQIP is led by a consortium of the Academy of Medical Royal Colleges, the Royal College of Nursing and National Voices. https://www.hqip.org.uk/
IVH	Intraventricular haemorrhage
LNU	Local neonatal units (LNUs) provide neonatal care for their own catchment population, except for the sickest babies. They provide all categories of neonatal care, but they transfer babies who require complex or longer-term intensive care to a NICU, as they are not staffed to provide longer-term intensive care. Most babies over 27 weeks gestational age will usually receive their full care, including short periods of intensive care, within their LNU. Some networks have agreed variations on this policy, due to local requirements. Some LNUs provide high dependency care and short periods of intensive care for their network population. LNUs may receive transfers from other neonatal services in the network, if these fall within their agreed work pattern.
LSOA	Lower Super Output Area: an area code for a location in England.
MADG	NNAP Methodology and Dataset Group.

MESH	Messaging Exchange for Social care and Health – the service used to enable users to submit lists of NHS numbers and receive lists back with the NHS numbers removed for those patients that have opted out of third party use of their data.
NEC	Necrotising enterocolitis
NCAPOP	National Clinical Audit and Patient Outcomes Programme
NHS Number	Unique numbers allocated to registered users of the three public health services in England, Wales and the Isle of Man.
NICU	Neonatal intensive care units (NICUs) are sited alongside specialist obstetric and feto-maternal medicine services and provide the whole range of medical neonatal care for their local population, along with additional care for babies and their families referred from the neonatal network. Many NICUs are co-located with neonatal surgery services and other specialised services. Medical staff in a NICU should have no clinical responsibilities outside the neonatal and maternity services.
NNAP	National Neonatal Audit Programme
ODN	Operational delivery network: In England, managed clinical networks for the coordination of neonatal critical care.
ODS	a unique code created by the Organisation Data Service within NHS Digital, and used to identify organisations across health and social care.
PHVD	Post haemorrhagic ventricular dilation
RCPCH	The Royal College of Paediatrics and Child Health (RCPCH) was founded in 1996 and now has over 17,000 members across the world. The RCPCH plays a major role in postgraduate medical education, professional standards, research and policy. https://www.rcpch.ac.uk
ROP	Retinopathy of prematurity
SCU	Special care units (SCUs) provide special care for their own local population. Depending on arrangements within their neonatal network, they may also provide some high dependency services. In addition, SCUs provide a stabilisation facility for babies who need to be transferred to a neonatal intensive care unit (NICU) for intensive or high dependency care, and they also receive transfers from other network units for continuing special care.

[i] Mood AM and Graybill FA. *Introduction to the Theory of Statistics*, 2nd ed. New York: McGraw–Hill (1963), p.408.

[ii] Hernán M and Robins J. *Causal inference: what if?* CRC Press (2024), section 17.5.