NMPA Methods for births occurring from 1 April 2018

Contents

Selection of audit measures

The suitability of a measure for inclusion in a national clinical audit depends on a number of explicit criteria: validity, fairness, sufficient statistical power and adequate technical specification. In addition to these criteria, it is also important for a set of audit measures to be balanced. In other words, the audit should cover various dimensions of care to give a complete overall picture of the service.

Measures were selected for inclusion in the NMPA continuous clinical audit through an iterative process:

1. A long-list of audit measures was prepared by the NMPA project team between July and November 2016, based on:

    - a pre-tender NMPA development and prioritisation project carried out by the National Perinatal Epidemiology Unit in 2014[1]
    - a review of relevant national standards and guidelines
    - consultation with the NMPA partner Colleges
    - a scoping exercise of currently available record-level datasets related to maternity to determine which measures would be possible to derive

2. The long-list was used as a basis for consultation with the NMPA Clinical Reference Group and Women and Families Involvement Group to determine the validity and usefulness of each measure. This process took place between November 2016 and May 2017 and resulted in a short-list of measures that were deemed clinically relevant and of use to our audience of women and families, clinicians, policymakers, commissioners and stakeholder groups.

3. Each short-listed measure was evaluated further by the NMPA project team, taking into account the data the NMPA was able to collect and access in its first year. The team considered the suitability of a measure in terms of:

    - feasibility and data quality

    i. how well can the population of interest be defined with the available data items?

    ii. how well can the important case-mix difference be captured by the available data?

    iii. how well can the procedures or outcomes that define the measure be captured?

    - statistical power

    i. what is the average number of patients within each unit with the procedure or outcome of interest?

    ii. what is the average number of relevant events within each unit?

    iii. what is the chance that a true outlier will be detected (in a unit of average size)?

Sixteen measures met these criteria. The NMPA has also developed a list of audit measures that are currently aspirational because the necessary data items are not routinely collected in datasets. The

NMPA continues to provide feedback and work with the national organisations responsible for managing maternity datasets to determine whether some of these measures may be collectable on a national basis in future years. It is also possible that some of the measures developed as part of the NMPA sprint audits will become part of the set of continuous audit measures.

The NMPA's approach to data collection

The NMPA differs from many other NCAPOP audits in that it brings together available data sources (i.e. those that are already collected either for clinical or hospital administrative purposes) rather than collecting primary data to create a bespoke audit dataset. By using existing datasets and linking these together, it eliminates the burden on clinical staff of data collection for the sole purpose of the NMPA.

Since the Clinical Report 2021 based on births in 2017-18, the NMPA uses national-level centralised maternity datasets for each participating nation. The use of national centralised datasets is advantageous as NHS trusts and boards are mandated to submit data to the centralised dataset according to a specification. This improves alignment of data items and coding before the NMPA receives data. An example specification can be found in the MSDS Technical Output Specification. Such an approach adheres to the principle of 'collect once, use many times' advocated by national data collection strategies. We hope that by using these datasets for national audit and feeding back results to trusts and boards, the NMPA will help to drive up the quality of the data contained within them year-on-year.

The NMPA's Data Specifications can be found on the NMPA Website.

Data sources used in the Annual Clinical Audit Report

Data for births in Wales are provided in the Maternity Indicators dataset (MIds) and the Initial Assessment (IA) dataset, both managed by Digital Health and Care Wales (DHCW), as well as Admitted Patient Care (APC) records from the Patient Episode Database for Wales (PEDW), and some fields from the National Community Child Health Database (NCCHD).

Data for births in England are provided in NHS Digital's Maternity Services Data Set and for the Clinical reports for births in 2017-18 and 2018-19 uses version 1.5 (MSDS v1.5) as well as Hospital Episode Statistics (HES) records. These are linked to the ONS register of live births, stillbirths and mortality register as well as the PDS birth notification dataset which together form the ONS-PDS spine. For births reported from 1 April 2019 onwards MSDS version 2.0 will be used.

A data flow diagram is available on the NMPA website.

Data quality

The NMPA team assesses data quality in four ways:

• Data completeness: for each key data items required by the NMPA, we exclude sites or trusts/boards if the proportion of records missing this information exceeds 30%. A summary of data completeness on some of the key variables is available online in the Data Completeness documents.

• Distribution: for many key variables, we define acceptable ranges within non-missing values. We exclude strongly outlying sites or trusts/boards that have a rate that is either too low or too high to be plausible (i.e. where no clinical reason for this level of variation could be envisaged). For example, a site with an obstetric unit will fail the gestational age check if the proportion of babies born at term (37+0 to 42+6 weeks) is less than 70%.

• Within-dataset consistency: for some variables, it is possible to perform internal consistency checks within the dataset. For example, it would be implausible for a woman who is coded as having her labour start as '*not applicable – delivered prior to labour onset via caesarean section'* to also be coded as having given birth vaginally. We check that these types of implausible records are rare within the dataset. Where consistency issues are detected, the variable identified as being the least plausible is set to missing for that record.

• Between-dataset consistency: for a few key variables, it is also possible to check for consistency between the different source datasets. For example, for England, gestation length at birth is available in the main maternity dataset MSDS as well as the ONS-PDS spine and the HES dataset. This is extremely valuable in the sense that it allows better decision making when cleaning the data, particularly when addressing within-dataset consistency issues on key variables. Where consistency issues are detected, the source identified as being the most plausible and/or reliable is retained for that record.
Additionally, in some rare cases where important data quality issues on a given measure affect an entire site or trust/board, it is sometimes possible to substitute that measure with that of an alternative source dataset for that given site or trust/board. Affected trusts/boards are notified when this is the case.

Assessment criteria were developed based on previous work.[2] The four data quality assessment approaches detailed above each serve a different purpose and, together, improve the likelihood of detecting poor quality data. For example, data quality assessment based on the proportion of missing data alone would not be sufficient, as it could lead to the inclusion of records from hospitals with seemingly complete data but with an observed distribution of data outside the expected range of values. By combining these approaches we can be confident that the published figures are based on data that have met at least a minimum standard of completeness and consistency.

Details of data quality checks performed for each measure as well as data item definition and alternative sources available for each can be found in the online Technical Specifications document.

Data analysis

Minimum requirements for inclusion in the analysis

The analysis in the clinical report is restricted to:

- sites and trusts/boards that passed the NMPA site-level and trust/board-level data quality checks
- birth records within those sites and trusts/boards that contain the required data to construct the measure
- birth records within those sites and trusts/boards that contain the required data to construct case-mix adjusted results (where adjustment is applicable)

The number of sites and trusts/boards for which results are available therefore varies from measure to measure, depending on the specific data requirements.

Construction of audit measures

The statistics in the clinical report are given as the proportion of events occurring within a group of women or babies. The reference group of women or babies (the denominator) changes between audit measures. As a general principle, the denominator for each measure is restricted to women or babies to whom the outcome or intervention of interest is applicable. For example, the measure of the '*proportion of women with a third or fourth degree tear*' is restricted to women who gave birth vaginally. A full description of these groups is contained in the Technical Specifications under each measure.

For measures relating to maternal care, results are presented per woman giving birth. For measures relating to the care of the baby, results are presented per baby born. In order to compare like with like, the majority of measures are restricted to singleton, term births.

Case-mix adjustment

When presenting figures for individual health service providers, it is often appropriate for audit measures to take into account how similar the patient groups are at each service, and how they differ between services. Clinical and demographic characteristics of women can affect both the demands placed on the maternity service and the outcomes of care. In turn, some women and birthing people and babies with more complex needs and at higher risk are referred to specialist services. Accounting for risk factors which are outside the control of care providers is essential before fair and meaningful comparisons across services can be performed.

In the NMPA clinical reports, we control for differences in the case-mix between services by adjusting results using logistic regression models. These models adjust for risk factors that are beyond the control of the maternity services such as age, parity, previous caesarean birth and clinical risk factors that may contribute to variation in performance between organisations. Neonatal factors included in the case mix are birthweight and gestational age.

Since the clinical report for births in 2017-18 the NMPA's case-mix adjustment no longer includes ethnicity or Index of Multiple Deprivation (IMD). This amendment, adopted for future clinical reports, is so that our results accurately show the inequalities in maternity outcomes and do not

mask the variation in outcomes associated with differences in a woman and birthing person's ethnic or socio-demographic background.

Further details, including which case-mix factors were used in each model, are given in the online Technical Specifications.

Reporting

All results for the NMPA's Annual Clinical Audit reports can be found on our website.

Levels of reporting

Current configuration of services has resulted in many NHS trusts and boards providing maternity services at more than one site. Hospitals with both an obstetric unit (OU) and an alongside midwifery unit (AMU) are therefore treated as one site. Site is the lowest level of granularity we are currently able to report on for the clinical measures, because for most sites with a co-located OU and AMU it is not possible to be absolutely certain whether a woman gave birth in the OU or AMU due to inconsistencies in the way the place of birth is recorded and lack of information on transfers in labour. Results by site as well as by trust/board are available on the NMPA website and will allow services to benchmark themselves against other services as well as national averages.

Suppression of small numbers

We are not able to present results where individual women or babies could theoretically become identifiable. Statistical power to detect true differences between sites is also influenced by the number of births occurring at that site. These issues affect the level at which some results can be reported, and particularly affect freestanding midwifery units (FMU), the majority of which have fewer than 500 births annually. For each measure, any site reporting fewer than 5 births that are eligible to be in the denominator are not reported at site level.

Presentation of results on funnel plots

A funnel plot (explained on our website in a short video) is a graphical method for comparing the performance of organisations.[3] The main advantage of this technique is that it takes the size of each organisation into account. This is important because the amount by which the result of an individual provider may vary from the national mean is influenced by random fluctuations that are related to the number of births within the provider.

In the sample funnel plot in Figure 1, results for England are shown as blue circles and for Wales as lilac squares. The dotted lines show the 95% control limits and the dashed (outer) lines the 99.8% control limits. 5% of trusts would be expected to lie outside the dotted lines and 0.2% outside the dashed lines due to chance. Here, many more trusts and boards lie outside these lines. This is an example of overdispersion.

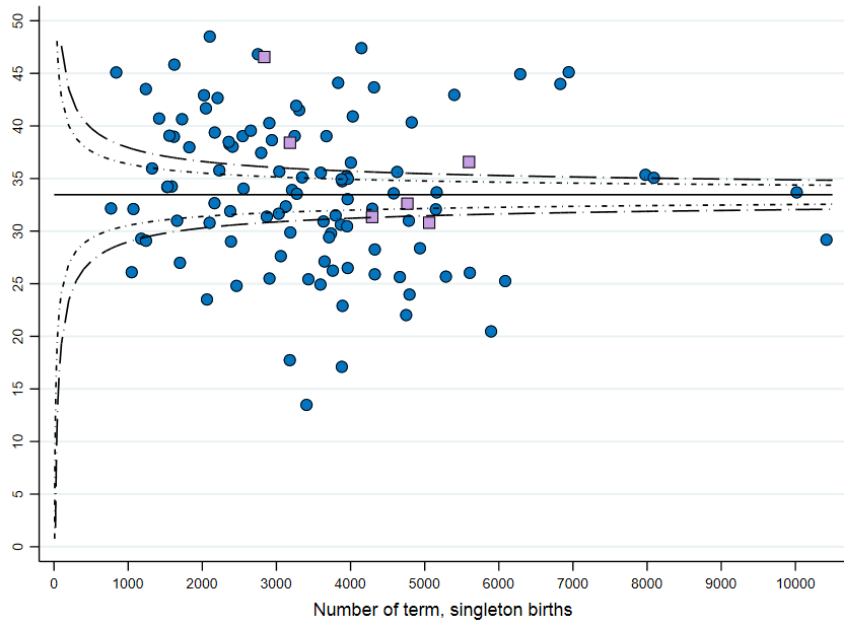

Figure 1: Sample funnel plot

The control limits within funnel plots highlight how much of the variation between providers exceeds that expected to occur due to chance alone. For the NMPA's three 'outlier measures', this is taken as an indication of care quality. For other measures, we use funnel plots only to show where there are substantial systematic (non-random) differences between services.

Several of the funnel plots presented in the clinical report show evidence of a phenomenon known as overdispersion.[4] Overdispersion occurs when a greater level of variability among providers is demonstrated than can be explained by chance and the existence of a few outlying units. Important possible explanations for overdispersion include differences in data quality, the limitations of the risk adjustment methods and 'clinical uncertainty.' This means variation in practice as a result of the absence of clear evidence-based clinical standards and different clinician preferences. We have attempted to limit the impact of differences in case-mix and in data collection and coding practices between sites as well as trusts/boards. However it is likely that some of the systematic variation between providers reflects clinical uncertainty.

Interactive funnel plots are available on the NMPA website, as well as data tables and overviews of all results per trust/board and site.

Presentation of results on scatter plots

For our measures on smoking cessation, feeding of babies and skin-to-skin contact at birth, data quality is sufficiently poor that the comparison of data on funnel plots is considered inappropriate. These results are therefore presented on scatter plots, which do not seek to compare performance between organisations.

The NMPA's outlier reporting

The NMPA's measures selected for outlier reporting were chosen because they represent adverse outcomes for women or babies with potential serious or long-term effects. Trusts/boards with results for these measures that are higher than would be expected by chance alone are notified and

asked to investigate why this might be the case. The measures included in the outlier reporting for the 2016-17 clinical report were:

- Proportion of women who sustained a 3rd or 4th degree perineal tear
- Proportion of women with an obstetric haemorrhage of 1500 ml or more
- Proportion of singleton, term, liveborn infants with a 5-minute Apgar score of less than 7

Outlier indicator reporting was not performed for Clinical reporting 2017-18 due to the lower proportion of English NHS trusts that could be included in the report, resulting in low confidence of an accurate national average for which to base outlier reporting on, and also due to poor timeliness of the data. Outlier reporting is currently also not planned for the 2018-19 clinical report after consulting with clinicians, due to the same issue of poor timeliness. However, it is the aim of the NMPA to continue with outlier reporting when both data quality and timeliness make outlier reporting valuable and meaningful to trusts and boards.

A full description on the NMPA's Outlier Policy is available on the NMPA website.

References

1. Kurinczuk J, Knight M, Draper E. National Maternity and Perinatal Audit Development Project. 2014.

2. Carroll F, Knight HE, Gurol-Urganci I, van der Meulen JH. Patterns of maternity care in English NHS Trusts 2013/14. RCOG. 2016 [https://www.rcog.org.uk/en/guidelines-research-services/audit-quality-improvement/clinical-indicators-programme/].

3. Spiegelhalter DJ. Funnel plots for comparing institutional performance. Stat Med. 2005 Apr 30; 24(8):1185–202. [https://onlinelibrary.wiley.com/doi/10.1002/sim.1970].

4. Spiegelhalter DJ. Handling over-dispersion of performance indicators. BMJ Quality & Safety. 2005; 14(5):347–51. [https://qualitysafety.bmj.com/content/14/5/347].