

National Pregnancy in Diabetes Audit

Methodology Document

Published 21 October 2016

Information and technology
for better health and care

Contents

	1
Introduction	3
Data collection	4
Data quality checks	5
Linkage	6
Variables derived using other datasets and tools	7
Index of Multiple Deprivation	7
Birthweight centiles	7
Analysis	8
Participation	8
Data by diabetes type	8
No comparisons over time	8
Records with missing data	9
Statistical methods used in the report	9
Quartiles	9
Testing for a significant difference between two proportions	9
Testing for a significant difference between two mean values	9
Confidence interval for a proportion	10
Confidence interval for a mean	10
Service level data	10
Disclosure control	11
Appendix 1 Data cleaning	12
Rejection reasons	12
Cleaning rules	12

Introduction

The National Pregnancy in Diabetes (NPID) audit is a continuous data collection measuring the quality of care and outcomes for women with pre-gestational diabetes who are pregnant and aims to support quality improvement.

All joint antenatal diabetes services in England and Wales are eligible to collect and submit audit data from consenting women. The audit also receives data from the Isle of Man.

The NPID audit measures against national standards set out in the NICE (National Institute for Health and Clinical Excellence) guideline NG3 (<https://www.nice.org.uk/guidance/ng3>), previously NICE Clinical Guideline CG63.

The audit seeks to address three key questions:

- Were women adequately prepared for pregnancy?
- Were appropriate steps taken during pregnancy to minimise adverse outcomes to the mother?
- Did any adverse outcomes occur?

The third annual national report on pregnancies ending in 2015 was published on 21 October 2016 together with service level data aggregated for pregnancies ending in 2013, 2014 or 2015.

The audit is part of the National Diabetes Audit (NDA) programme, and is commissioned by the Healthcare Quality Improvement Partnership (HQIP) as part of the National Clinical Audit (NCA) programme.

The NDA is managed by NHS Digital in partnership with Diabetes UK and is supported by Public Health England (PHE).

This document describes how the data is collected, processed and analysed before publication.

Data collection

To minimise the burden of data collection and submission the audit data is partly collected from antenatal diabetes services and partly obtained by linking the directly collected data to:

- National Diabetes Audit (NDA) data
- Hospital Episode Statistics (HES) data
- Patient Episode Database for Wales (PEDW) data.

A full list of data items collected by the audit is available from www.digital.nhs.uk/npid

Antenatal diabetes services provide each woman eligible for inclusion in the audit with a patient information leaflet and discuss the audit with them. If the woman agrees to be included in the audit they are then provided with a consent form.

The service collects and submits NPID audit data for women only who have given written consent for their data to be included in the audit. No record is kept of how many women did not consent so it is not possible to calculate case ascertainment.

All data collected by the NPID audit should already be stored in clinical records; no additional data should need to be collected.

Staff in antenatal diabetes services can use the NPID audit data collection form available from www.digital.nhs.uk/npid if they wish to collate the relevant data prior to entering it electronically.

The data is electronically transferred to NHS Digital by manual data entry into a secure online system called the Clinical Audit Platform. The data is entered by registered users at antenatal diabetes services. The data entry system checks each record for obvious errors and only records that pass these checks can be saved and submitted successfully.

Data can be collected and entered throughout the year. Services are asked to submit all records of pregnancies with a recorded outcome ending in a calendar year by mid-February in the following year. Any records submitted after the deadline will not be included in that year's report but will be added to the cumulative dataset.

For women attending antenatal diabetes services in the North East of England, much of the information collected in the NPID audit was recorded in the Northern Diabetes in Pregnancy (NorDIP) survey managed by the Regional Maternity Service Office on behalf of Public Health England. To reduce the burden of data submission, from mid-2014 to the end of 2015 the NorDIP survey also consented women for NPID and securely transferred data for consenting women to the NPID audit team at NHS Digital.

Data quality checks

Following the submission deadline, each service receives details of data quality checks run on their data, flagging records to review and potentially amend before a final submission deadline in early April.

The number of errors in the data has reduced greatly since internal consistency checks of each record were introduced at the point of data entry in Jan 2015, (for example that the HbA1c measurements within pregnancy are between 0 and 40 weeks prior to the pregnancy end date). However some more complex errors are only detectable when the dataset is checked as a whole (for example, that the same outcome data has been attached to more than one pregnancy).

Following the final submission deadline in April, data quality checks are re-run, and either:

- Where the error would affect a lot of the analysis, the record is rejected. Across the 2013, 2014 and 2015 dataset approximately 100 records were rejected (leaving nearly 7,400 pregnancies).
- Where the other data in the record can still be used for analysis, the data is 'cleaned' by setting the value(s) assumed to be incorrect to missing.

A full list of the rejection reasons and cleaning actions is included in Appendix 1.

Linkage

National Diabetes Audit data

For each pregnancy record in the NPID data ethnicity and Lower Super Output Area LSOA of residence were obtained from the most recent NDA data record in the year of or prior to pregnancy. For example, for 2013 pregnancies, a match was sought in the 2013-2014 NDA, then the 2012-2013 NDA and finally in the 2011-2012 NDA, while for 2015 pregnancies the matching order was 2014-2015, 2013-2014 and 2012-2013.

As mother's diabetes type was added to the NPID data collection as a mandatory data item from 1 January 2015, this was also obtained from NDA data for the 2013 and 2014 NPID audit records, again by linking to the most recent relevant NDA record.

Year of diabetes diagnosis was obtained from the earliest NDA record available for each woman in NPID, excluding diagnosis dates that were after the start of the pregnancy or before the woman's date of birth as recorded in the NPID data.

Hospital Episode Statistics (HES) and Patient Episode Database for Wales (PEDW)

Onset of labour, mode of delivery and parity (number of previous pregnancies) were obtained by linking the NPID audit data to HES and PEDW data.

Because HES data is released by financial year and the NPID audit reports by calendar year, only NPID pregnancies ending in 2013 or 2014 can currently be linked to available HES data.

A hospital episode record was accepted as a match when the pregnancy end date in NPID was up to 7 days earlier than the episode start date or up to 7 days after the episode end date.

The PEDW data was sparsely populated for these variables.

Details of hypoglycaemia and DKA (diabetic ketoacidosis) episodes during pregnancy were also obtained by linking to HES and PEDW data. A hospital episode record was accepted as a match when the episode started no more than 40 weeks before the estimated delivery date and on or before the pregnancy end date.

Where hypoglycaemia and DKA diagnosis codes were recorded on the same episode, this episode was counted as both an episode with hypoglycaemia and an episode with DKA. This is consistent with the method for counting complications used by the NDA.

The diagnosis codes used to determine if there was a Hypoglycaemia or Diabetic Ketoacidosis (DKA) episode during the women's pregnancy are as follows:

Hypoglycaemia

- E16.0 Drug-induced hypoglycaemia without coma in any Diagnosis field and Y42.3 Insulin and oral hypoglycaemia (antidiabetic drugs) in a secondary diagnosis position

OR (in any Diagnosis field)

- E161 Other hypoglycaemia
- E162 hypoglycaemia, unspecified

Diabetic ketoacidosis (DKA)

The following codes in any diagnosis field

- E10.1 Type 1 diabetes mellitus with ketoacidosis
- E11.1 Type 2 diabetes mellitus with ketoacidosis
- E13.1 Other specified diabetes mellitus with ketoacidosis
- E14.1 Unspecified diabetes mellitus with ketoacidosis

Variables derived using other datasets and tools

Index of Multiple Deprivation

Local measures of deprivation are produced by the Office for National Statistics for England and Wales separately, with the measures for each country considering different factors and using different calculation methods.

For the national NPID audit report England and Wales data have been analysed together. In order to assign a deprivation quintile to each pregnancy record based on the LSOA of residence of the mother (where this could be obtained from linking to NDA data), a 'combined' deprivation score was derived and ranked into quintiles. This combined score was based on an equally weighted combination of the individual scores for the employment and income indices using methodology designed by the Office for National Statistics (ONS)¹.

Birthweight centiles

Birthweight centiles are used to adjust the babies' actual birthweight in line with maternal factors such as ethnicity, height and weight as well as gestational age at delivery.

Birthweight centiles were calculated for all singleton babies in the NPID dataset where the gestation at delivery and birthweight was known, using the GROW centile tool².

A baby is described as large for gestational age (LGA) if its birthweight is above the 90th centile based on gestation and maternal characteristics. In the general population, 10 per cent of babies would be expected to be above the 90th centile.

¹ ONS (2013) Using Indices of Deprivation in the United Kingdom p.7
http://www.neighbourhood.statistics.gov.uk/HTMLDocs/images/UK%20wide%20guidance%20paper%20April%202013%20revision_tcm97-129456.pdf

² GROW centile tool: Gardosi J, Francis A. Customised Weight Centile Calculator. GROW v6.7.7.1 (UK), 2015, Gestation Network, www.gestation.net

Analysis

Participation

An antenatal diabetes service is counted as participating in the audit if at least one completed pregnancy record has been submitted by that service for the audit period. A full list of participating services is including in the service level excel file.

Data by diabetes type

The national report mainly shows data for women with Type 1 diabetes and women with Type 2 diabetes. Analysis putting these two groups together as 'All diabetes' can be misleading because the care pathways and physiology are often different for women with Type 1 diabetes to those women with Type 2 diabetes, and the 'All diabetes' figure will be a reflection of the mix of diabetes types seen by the service rather than the overall quality of care.

Notwithstanding this, the service level reports do group together data for women with different diabetes types in order to provide some basic feedback to services that do not see a large volume of patients.

The number of women with MODY (Maturity Onset Diabetes of the Young) or 'Other diabetes' is too small to provide analysis for this group.

No comparisons over time

The 2015 report does not compare NPID audit data over the first three years of the audit for the following reasons:

- The number of services submitting data to the audit and also the total number of records submitted each year has increased quite significantly each year, so variations in the data from year to year may well be due to different services starting to submit data or services submitting data for a fuller range of their patients.
- We do not yet expect to see the effects of service improvement plans put in place following the first year of the audit. The first audit report was published in October 2014, at which time most of the women whose pregnancies ended in 2015 will already have been pregnant. The first full year of pregnancies that could have been helped by service improvement following that first report will be the pregnancies ending in 2016, which will be reported in 2017.

Records with missing data

Each table or chart uses all records for which the relevant data is valid and not missing in order to include as many records as possible.

This means that the denominator for percentages varies between the charts and tables.

For example, pregnancies where gestation is unknown because estimated delivery date has been removed during cleaning will be excluded from analysis where gestation is relevant, such as birthweight centiles, but will be included in other analysis where the gestation at delivery is not needed, such as whether the mother was taking 5mg folic acid prior to pregnancy.

For non-mandatory data items such as HbA1c measurements, the denominator will be much lower reflecting the number of records in which this data has not been entered

Statistical methods used in the report

Quartiles

The national report shows the variation between services for some measures, such as the percentage of women taking 5mg folic acid prior to pregnancy, and summarises the extent of variation using the median and quartile values.

Arranging all the values in order, the median is the middle value. The lower quartile is the value below which the bottom 25 per cent of data values lie and the upper quartile is the value above which the top 25 per cent of data values lie.

Testing for a significant difference between two proportions

Where the report explicitly compares a proportion for two groups, such as the percentage of women with Type 1 diabetes taking 5mg folic acid compared to the percentage of women with Type 2 diabetes taking 5mg folic acid, the difference between the two groups has been identified as significant by statistical testing with a p-value below 0.05 using a two-sample z-test.

Testing for a significant difference between two mean values

Where the report explicitly compares a mean value for two groups, such as the mean BMI for women with first trimester HbA1c <48 mmol/mol and first trimester HbA1c ≥48 mmol/mol, the difference between the two groups has been identified as significant by statistical testing with a p-value below 0.05 using a two-sample t-test.

Confidence interval for a proportion

The 95 per cent confidence intervals for the stillbirth, neonatal death and congenital anomaly rates were calculated using Byar's method as described in 'Analytical Tools for Public Health: Commonly used public health statistics and their confidence intervals'³ as

$$\text{Lower limit} = \frac{O}{n} \times \left(1 - \frac{1}{90} - \frac{1.96}{3\sqrt{O}}\right)^3$$

$$\text{Upper limit} = \frac{(O + 1)}{n} \times \left(1 - \frac{1}{9(O + 1)} - \frac{1.96}{3\sqrt{(O + 1)}}\right)^3$$

where O is the observed number of events and n is the rate denominator.

Confidence interval for a mean

The 95 per cent confidence intervals for the mean first trimester HbA1c values for groups with different pregnancy outcomes were calculated using the CONFIDENCE.T function in Excel.

Service level data

The service level data published with this report includes all data for pregnancies ending in 2013, 2014 or 2015. The number of records for each service will in part depend on when they started participating in the audit, so some services that see a smaller number of patients per year may have more records included than a larger service because they have participated for all 3 audit years while the larger service has only joined during 2015.

The average number of pregnancy records per service is less than 50.

Service level data is only published where a service has submitted 10 or more completed pregnancy records across the 3 years. As the total number of records for each service is relatively small (an average of less than 50 records per service), an apparently large difference in percentages between services may only reflect a small difference in the numerator. The service level report contains a list of participating services.

The England, Wales and Isle of Man figures in the service level report are similarly based on 3 years of data so will differ from the 2015 figures in the national report.

³ <http://www.apho.org.uk/resource/view.aspx?RID=48617>

Disclosure control

In line with the requirements of the *Anonymisation Standard for Publishing Health and Social Care Data* and in agreement with the NHS Digital Disclosure Control Panel, disclosure control to mitigate the risk of patient identification is not required in the national report or for data at Government Office region level but is required at service level.

The nature of the NPID data means that it could be possible to identify both where an individual did or did not meet a condition or have a particular outcome. For example, identifying that a woman was not taking 5mg folic acid could be as sensitive as identifying that they were.

Disclosure control has therefore been applied to each percentage where any of the following apply:

- Numerator 5 or less (including 0).
- (Denominator – numerator) 5 or less (including 0).
- Denominator 5 or less (unlikely as limited to services with 10 or more records but possible if there is missing data).

To provide some feedback to the large number of services that have one or more suppressed values, the suppressed values have been replaced by a 'banding' (<25 per cent, 25-50 per cent, 50-75 per cent or ≥ 75 per cent) which does not reveal the exact value but gives a broad indication of the service's value for the relevant measure.

Where differencing within a region or between the 'All diabetes' and diabetes type data would reveal an exact value concealed by the banding, banding has been applied to further values to prevent this.

Appendix 1 Data cleaning

Rejection reasons

A small number of records with data entry errors were rejected because the error meant that the correct complete record could not be identified or mandatory data needed to allocate a record to the correct service was missing.

The total number of records rejected from the 2013-2015 dataset was approximately 100.

A record with pregnancy outcome data completed was rejected if:

- Two outcomes with the same pregnancy end date were attached to different pregnancy records for the same woman.
- The same baby NHS number was entered in outcome data for more than one pregnancy.
- The same baby NHS number was entered more than once in outcome data for the same pregnancy.
- Two outcomes with substantially different pregnancy end dates were attached to the same pregnancy.
- Booking or delivery hospital was missing – these are mandatory data items needed to allocate each record to a service.

Following requests from submitters, the data entry system was altered from January 2015 to allow the 'Alive at 28 days' question to be completed separately from the other outcome data. However, 'Alive at 28 days' was not completed for a number of otherwise complete and valid records which were also rejected.

Cleaning rules

Dates

The table below lists the cleaning rules applies to the NPID audit data prior to analysis in order to use as much data as possible from each record where errors remained after the data quality review by services.

Where the dates within a record were inconsistent, the pregnancy end date and pregnancy outcome have been assumed to be correct, and dates that were inconsistent with this have been set to missing.

The order of date checking means that where an estimated delivery date has been set to missing, checks which compare other dates to estimated delivery date will result in the removal of those dates.

Table: Data cleaning rules for dates used for 2015 NPID audit annual report

Data set to missing	Reason(s)
Mother's date of birth	Before 01/01/1961 or after 01/01/2001
Estimated delivery date (EDD)	<ol style="list-style-type: none"> 1) More than 40 weeks later than pregnancy end date (negative gestation) 2) Gestation by comparing with pregnancy end date 43 weeks or more 3) Gestation by comparing with pregnancy end date less than 4 weeks 4) Stillbirth at < 24 weeks 5) Miscarriage at >=24 weeks 6) Live birth at <20 weeks 7) Live birth at <24 weeks with weight >1kg
Date of first contact with antenatal diabetes team	<ol style="list-style-type: none"> 1) More than 40 weeks before estimated delivery date 2) After pregnancy end date 3) More than 3 weeks after estimated delivery date
First HbA1c measurement in pregnancy and date of this measurement	<ol style="list-style-type: none"> 1) More than 40 weeks before estimated delivery date 2) After pregnancy end date 3) More than 40 weeks before pregnancy end date
Last HbA1c measurement in pregnancy and date of this measurement	<ol style="list-style-type: none"> 1) More than 40 weeks before estimated delivery date 2) After pregnancy end date 3) More than 40 weeks before pregnancy end date

Diabetes type

It was intended that mother's diabetes type would be obtained by linking the NPID audit data to NDA data. However, fluctuating GP participation in the NDA and the application of patient opt-outs to the NDA data meant that this linkage was unlikely to be complete. As maternal diabetes type is an important data item for the NPID audit analysis, it was added to the NPID online data entry system for pregnancies ending on or after 1 January 2015. Whilst some services have retrospectively populated this data for some 2013 and 2014 pregnancies, linkage to the NDA is still required to populate this data for as many records as possible from the earlier years.

As the NPID audit data accumulates over time, it includes second and subsequent pregnancies for many women. We would expect the woman's diabetes type to remain the same between pregnancies. However, data quality checks revealed some possible data entry errors with different diabetes types being recorded for the same woman.

There may also be inconsistencies in diabetes type recording within the NDA data however the approach to linkage means that only one value is chosen without considering conflicts with values from other NDA years.

To maintain consistency in the dataset, the following cleaning rules have been applied to diabetes type data:

Diabetes type data		
Entered in NPID	Obtained from NDA	Used for analysis
One value that is not '99 – Not specified' entered for either single pregnancy or same value for multiple pregnancies	Not needed	NPID value
One value that is not '99 – Not specified' entered (same value for one or more pregnancies) and one or more pregnancies with '99 – Not specified' or missing	Not needed	Value that is not 99 or missing applied to all pregnancies for this woman
'99 – Not specified' or missing for one or more pregnancies and no other pregnancies with a value that is not 99 or missing	Diabetes type that is not 99 or missing	NDA value
'99 – Not specified' or missing for one or more pregnancies and no other pregnancies with a value that is not 99 or missing	No data in relevant NDA audit years for this pregnancy, but data in a different NDA year for another pregnancy for this woman.	Value obtained from NDA for any pregnancy applied to all pregnancies
'99 – Not specified' for one or more pregnancies and no other pregnancies with a value that is not 99 or missing	No matches to NDA for any pregnancy	NPID value