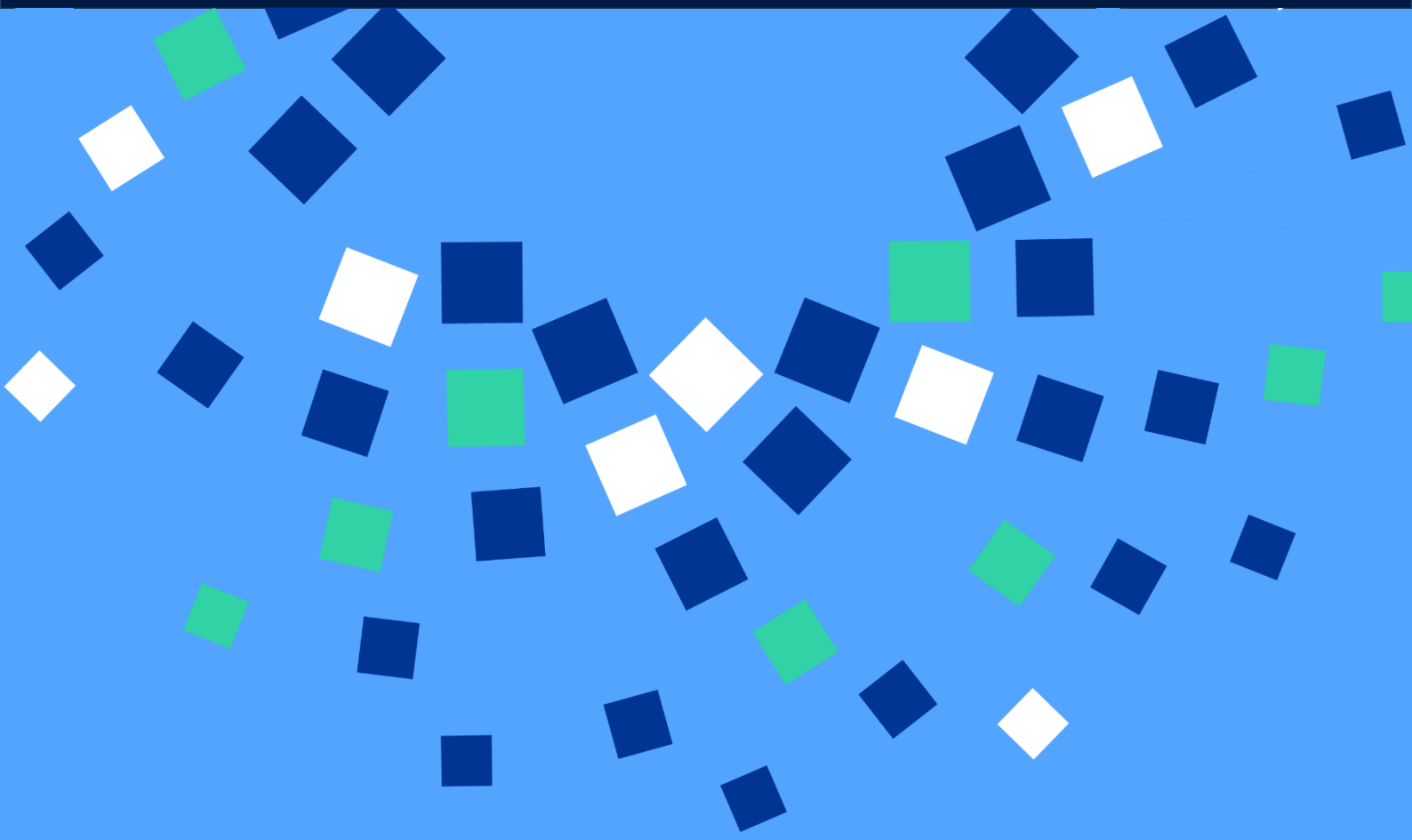
YDG CYMRU | ADR WALES

Data Explained

Maternity Indicators Data Set (MIDS)

Authors: Sarah Aldridge, Victoria Best, Roberta Chiovoloni, Ashley Akbari
Date: October 2026

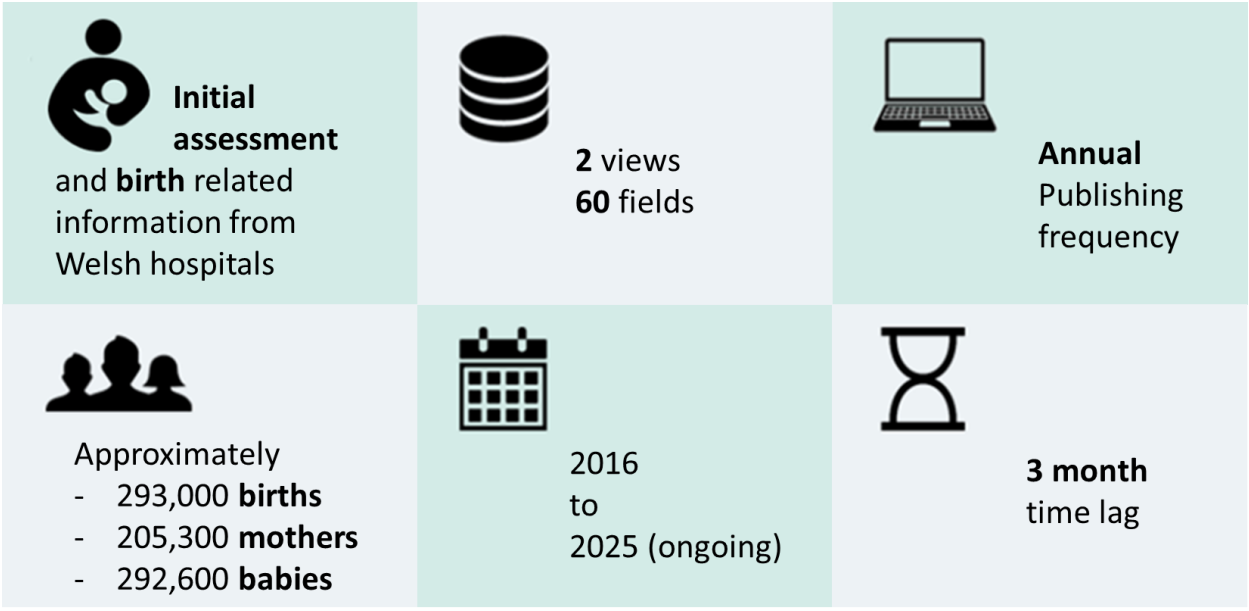


Summary

This Data Explained summarises experiences and learning from working with the Maternity Indicators Data Set (MIDS). This publication is intended to help guide future researchers using this data and to provide feedback on future data development and documentation.

The administrative data discussed in this Data Explained was made securely available through an approved SAIL project via the independent Information Governance Review Panel (IGRP) – project 1934. The data used in this research project comes from Digital Health and Care Wales (DHCW) and was accessed through the Secure Anonymised Information Linkage (SAIL) Databank. The data were not originally collected for research, and gaps and inconsistencies in their recording are expected; some are detailed below.

Overview



Introduction

The Maternity Indicators Data Set (MIDS) was established in 2016 as a national initiative to enable the consistent collection and reporting of data on maternity services across NHS Wales. Its purpose was to support the monitoring of nationally agreed outcome indicators and performance measures relating to maternity care – including antenatal, birth and neonatal stages – as set out in A Strategic Vision for Maternity Services in Wales (Welsh Government 2011). MIDS was designed to address existing gaps in data quality and coverage, particularly where existing national data failed to meet the requirements for comprehensive maternity monitoring.

At the time, the two main data sources - the Admitted Patient Care Data set (APCds) and the National Community Child Health (NCCH) database - were unable to support the full range of maternity indicators required (Welsh Government 2017). The APCds did not include sufficient antenatal data or information on home births, and was populated mainly from patient administration systems rather than dedicated maternity systems. Similarly, the NCCH began at birth and

excluded earlier pregnancy data. In response, a national programme was launched to establish a baseline for data quality across health boards and ensure that local maternity systems could generate data suitable for national reporting.

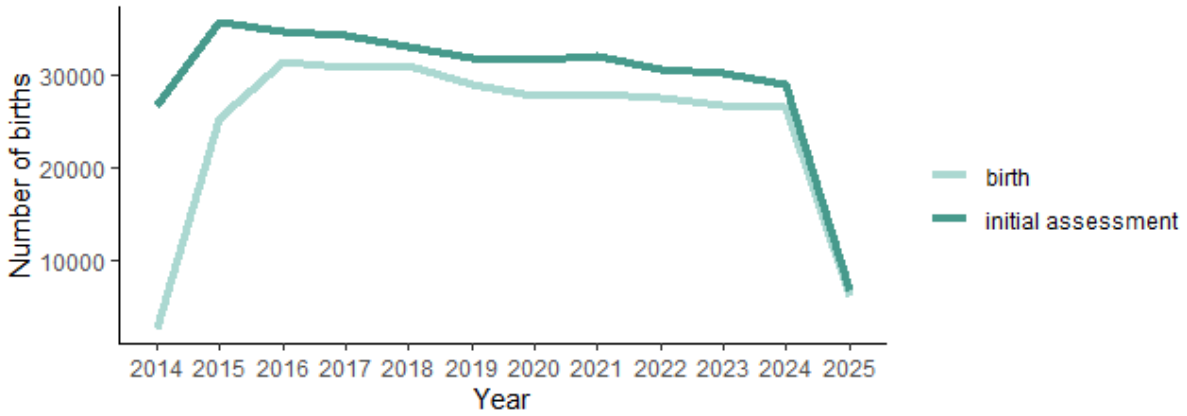


Figure 1 – The number of births recorded in MIDS per year.

How is the data collected?

The MIDS captures data on the mother at the initial assessment with a midwife, and on the mother and baby (or babies) at birth in Wales only (i.e. initial assessments of birth in English providers aren't included). Each health board (HB) must make available data relating to the initial assessments and/or birth events that they managed. Velindre NHS Trust is excluded from this requirement, as it does not provide any maternity services. When the initial assessment and birth events occur in different HBs, data are linked nationally via Digital Health and Care Wales (DHCW), formerly known as the NHS Wales Informatics Service (NWIS) (NHS Wales, n.d.).

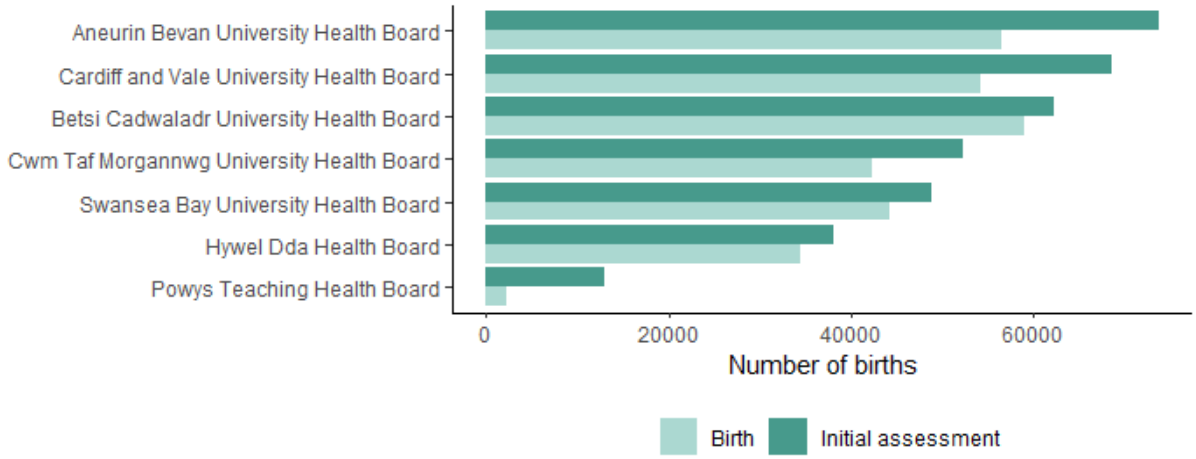


Figure 2 – The number of births and initial assessments recorded in each health board in MIDS.

MIDS allows a researcher to analyse characteristics of the mother's pregnancy and birth process. The process for producing this data extract is complex, largely because multiple initial assessment records can exist, and birth records are not always complete. To merge initial assessment data with birth record data, DHCW performs several validation checks, including removing mothers and babies with missing NHS numbers, deduplicating, consolidating, and validating initial assessments, and applying other cleaning measures. This complex process removes many records, so statistics produced from the MIDS have limitations.

Data structure and linkage

MIDS is formed from two individual tables that provide information from a mother's initial assessment and the birth. While MIDS was established in 2016, there is data available from 2014 for births and 2015 for initial assessments.

SAIL uses an Anonymous Linkage Field (ALF) for individual-level linkage across all other data sources. MIDS is linkable to these via the mother's (*MOTHER_ALF_PE)* and the child's *(CHILD_ALF_PE)* individual linkage fields. The birth table contains the maternal ALF only, while the birth table contains both the mother's and child's ALFs. In the birth table, linkage quality for the maternal ALF is high, with >99% of births reported having a corresponding maternal ALF with high linkage quality (*MOTHER_ALF_STS_CD* of 1). Less than 1% of births have either a missing child ALF, maternal ALF, or both fields are the same.

Linkage quality	Proportion
Valid child ALF and maternal ALF	99.4%
No child ALF	<0.1%
No Maternal ALF	0.3%
Maternal ALF = child ALF	0.3%

Table 1 - A summary of the linkage in MIDS.

Initial assessments table: This table collects a snapshot of data typically around 8-9 weeks into gestation. This initial data capture includes information relevant to the mother and the pregnancy, including the mother's age, weight, height, presence of mental health conditions, demographic factors, service provider, week of gestation, and number of previous pregnancies and caesareans. This table is at the individual-level, however not all those who go on to record a birth in MIDS are present, while other mothers have multiple initial assessments.

Birth table: The birth table includes data relevant at the time of birth, collecting information about the birth itself, such as the date, mode of birth, labour augmentation, and the healthcare provider. For the mother, demographic factors (smoker status and weight) are reported; for the baby, linkage field sex and weight. This table is at the individual level for each birth; in multiple gestation, these are treated as multiple birth events, with details for each child and birth recorded.

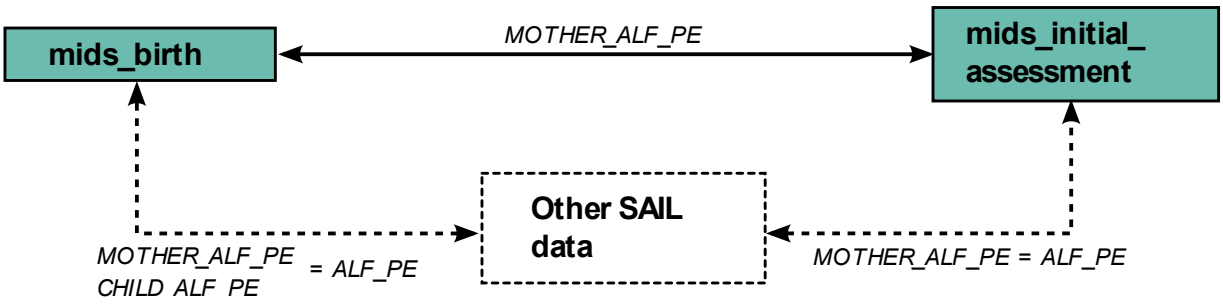


Figure 3 - Linkage diagram for MIDS views.

Key variables

Birth outcomes are recorded in the *BIRTH_OUTCOME_CD* as 1 for live births and 0 for stillbirths. However, a stillbirth is only registerable in MIDS as a birth if it occurs after a gestation of 24 weeks. Gestation is measured in complete weeks; e.g., 23 weeks and 6 days would be recorded as 23 weeks.

Variables including *BIRTH_MODE*, *LABOUR_ONSET_MODE_CD*, *LABOUR_ONSET_FOETAL_LIE_CD*, *LABOUR_ONSET_FOETAL_PRESENTATION_CD*, *LABOUR_AUGMENTATION_CD*, *LABOUR_EPISIOTOMY_CD*, *LABOUR_PERINEAL_STS_CD* have data dictionaries that differ from those of other maternity data sources. As such, we highly recommend that users consult the MIDS data dictionaries when using MIDS to ensure accurate data interpretation.

It's also worth noting that some fields should be incompatible with others according to their definitions, so some caution should be taken when combining these variables. A transverse foetal lie (*LABOUR_ONSET_FOETAL_LIE_CD* = 1) is incompatible with a breech or cephalic foetal presentation (*LABOUR_ONSET_FOETAL_PRESENTATION_CD* = 1 or 2); however there are nearly 700 births with these combinations of codes present in MIDS.

What can the data be used for?

MIDS contains detailed, individual-level data relating to both the mother and their children for births registered in Wales to mothers who are resident in Wales. This makes it a valuable resource for studying a wide range of maternal and child-related health research. It can be used to explore perinatal and postnatal health, monitor service use, assess health inequalities, and evaluate the impact of public health interventions. The ability to link MIDS with other data sources within SAIL further enhances its potential for research into intergenerational health and long-term outcomes in Wales.

Using MIDS

Users might consider comparing their MIDS analysis with analyses from other maternity data sources. MIDS is limited to births registered in Wales to Welsh residents, and won't capture births in English hospitals or those of mothers who aren't Welsh residents. The Office for National Statistics (ONS) routinely produces registered births and infant mortality statistics, and these should be used as the main source for the number of births in Wales.

In addition to conflicting birth characteristic definitions and data dictionaries compared with other maternity data sources, there are inconsistencies in how characteristics are reported over time or – quite notably - by health board. Some of these may be due to the facilities available in those health boards; for example, Cardiff and Vale and Betsi Cadwaladr University health boards having higher rates of "Unknown" breastfeeding intentions compared to other health boards could be a result of the neonatal intensive care units present at those hospitals, which would interrupt the normal progression onto breastfeeding after birth (Figure 4). Variation could also be a reflection of differences in reporting standards in different health boards, such as those seen in the reporting of foetal lie at the onset of labour (almost completely absent in Swansea Bay University and Hywel Dda Health boards (Figure 5)), labour augmentation (Figure 6) and epidural (Figure 7). This extends to non-birth related information such as ethnicity (Figure 8), which is affected by reporting consistency (reported in 93% of births in Cardiff and Vale University Health Board, and only 49% of births in Hywel Dda Health Board) and smoker status (Figure 9), which is affected by reporting consistency and the availability of validation equipment (reported in 82% of births in Swansea Bay University Health Board, while Cardiff and Vale and Cwm Taf Morgannwg University Health Boards have a smoker status for >99% of births).

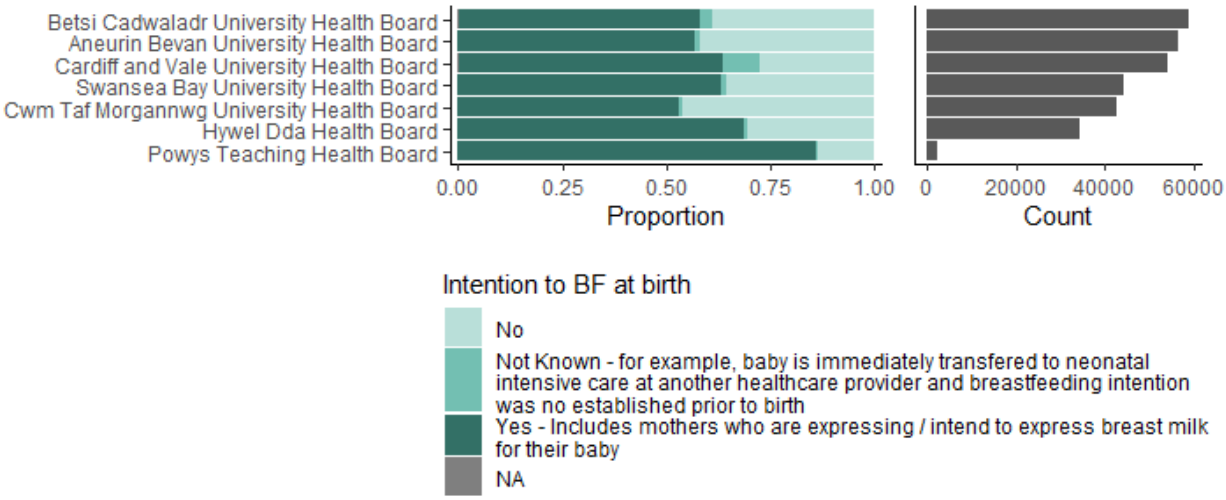


Figure 4 - Counts for breastfeeding intention by health board present in MIDS.

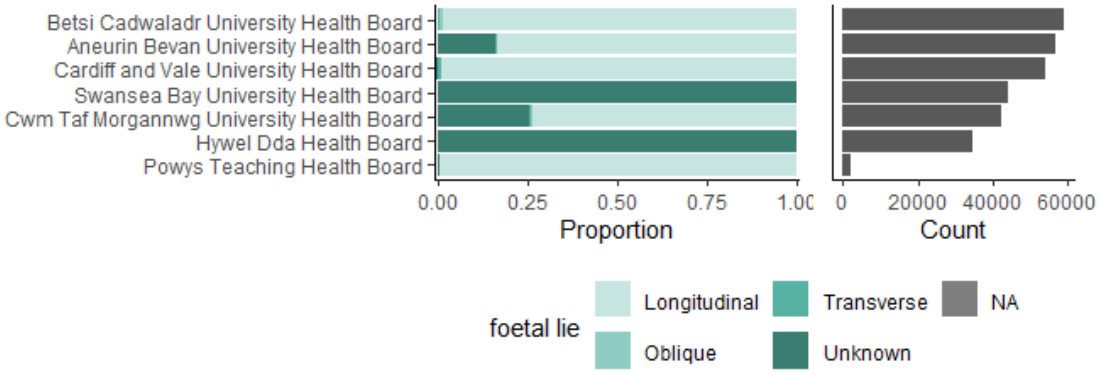


Figure 5 - Counts for foetal lie at the onset of labour by health board present in MIDS.

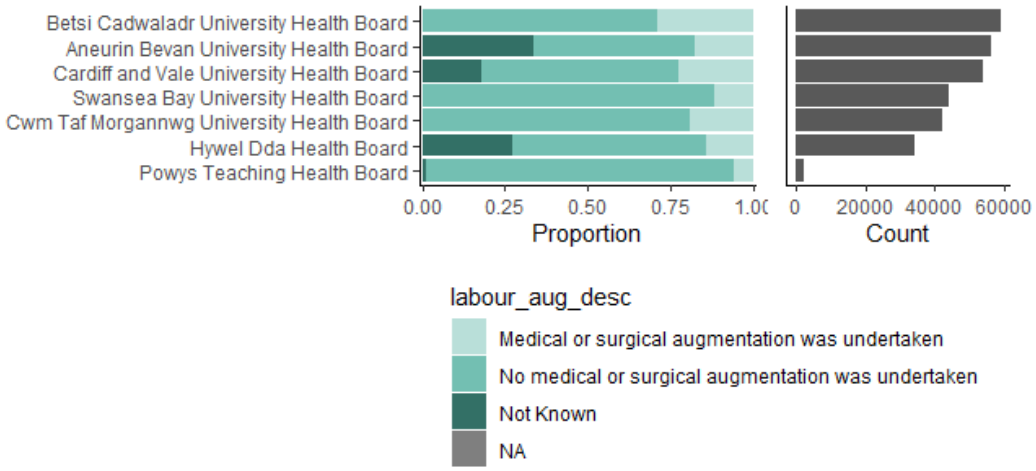


Figure 6 - Counts for augmentation of labour by health board present in MIDS.

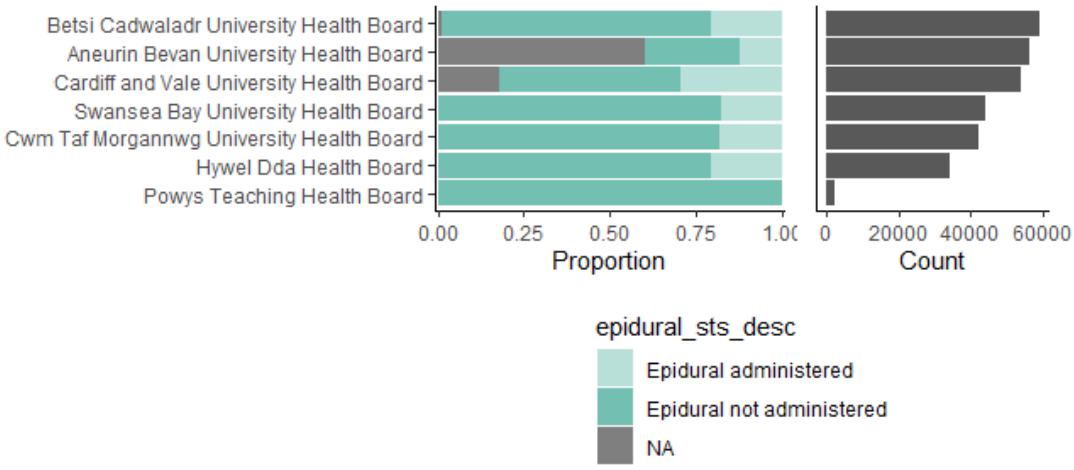


Figure 7 - Counts of epidural status by health board present in MIDS.

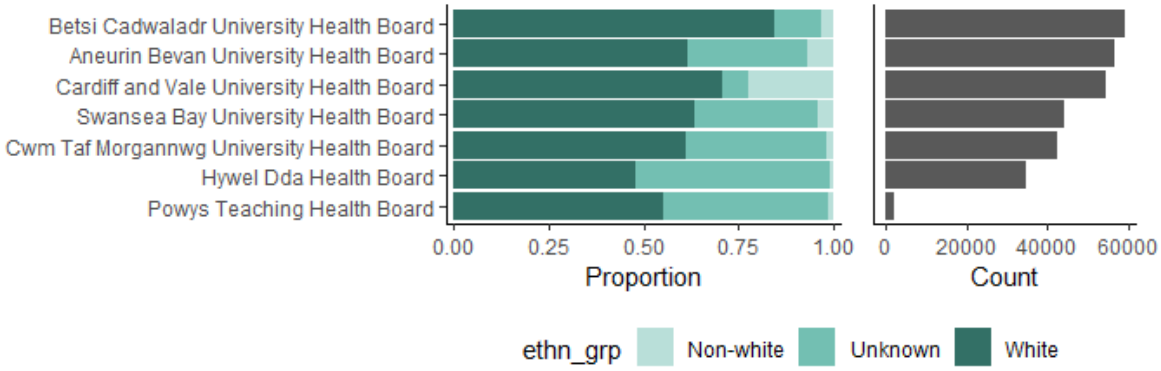


Figure 8 - Counts for ethnic group by health board present in MIDS.

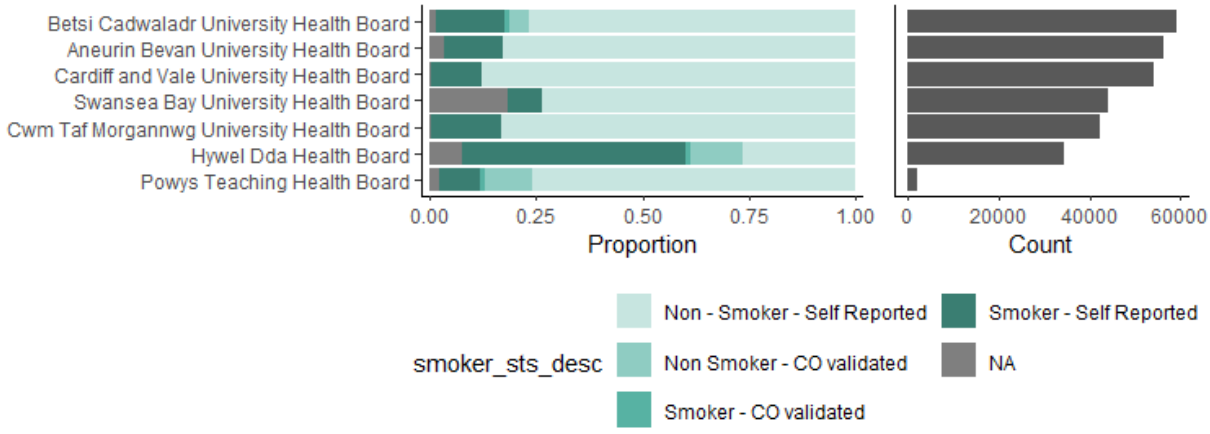


Figure 9 - Proportions and counts for the mother's smoker status per pregnancy by health board present in MIDS.

Strengths

MIDS captures detailed information about both mothers and their children at the individual-level for all births recorded in Wales to mothers registered with a Welsh general practice. Both mother and child linkage fields are available in >99% of births, providing population-wide coverage that can be linked to other SAIL data sources. Merging the initial

assessment and birth records allows users to analyse data across the full maternity and birth process. It also enables monitoring of areas of interest, such as smoking during pregnancy or Body Mass Index (BMI).

Data limitations encountered

MIDS is a relatively young data source and provides data only from 2014 onwards. The maternal ALF is missing from 0.3% of births, and the child ALF is missing from <0.1% of births. In 0.3% of birth records, the mother and child ALF are the same. Linkage to other SAIL data may enable users to determine whether this ALF should be allocated to the mother or the child, but users may want to remove these births during their cohort selection.

MIDS does not include any data for birth records for Welsh mothers who gave birth in England, or any other country outside Wales, though a birth record may still be present. Data from initial assessments performed on Welsh mothers in English hospitals are also not imported. Additionally, initial assessments performed in a different HB to the birth are removed during the merging process. Manual data entry introduces a range of potential errors and inconsistencies within the MIDS. These include missing data, entry mistakes and variation in data formats - such as differences in recorded units or measurement scales. Some issues, such as incorrect units, can often be identified and corrected through validation processes, while others, such as mismatched mother-child identifiers, are more difficult to detect. There is also notable variation among HBs in data completeness and quality. These differences may reflect local guidance on data entry, variation in clinical practice, or disparities in available equipment - such as whether carbon monoxide monitors are used to validate smoke status. Some health boards are also unable to distinguish between home and hospital births (Welsh Government 2025). Validation rules are not always applied consistently: for example, some health boards record ‘no previous Caesareans’ as “0” while others record it as “99”. In some cases, fields do not align with the MIDS NHS data dictionary. While data completeness is relatively high for most data items, researchers may need to account for it longitudinally or regionally.

References

NHS Wales. n.d. 'NHS Wales Data Dictionary'. NHS Wales Data Dictionary. Accessed 18 December 2025. https://www.datadictionary.wales.nhs.uk/index.html#!WordDocuments/scope16.htm.

Welsh Government. 2011. *A Strategic Vision for Maternity Services in Wales*. No. WG12896. https://www.gov.wales/sites/default/files/publications/2019-03/a-strategic-vision-for-maternity-services-in-wales.pdf.

Welsh Government. 2017. *Maternity Statistics Wales, 2015-16 (Experimental Statistics)*.

Welsh Government. 2025. *Maternity and Birth Statistics: Quality Report*. https://www.gov.wales/sites/default/files/pdf-versions/2025/8/2/1754383675/maternity-and-birth-statistics-quality-report.pdf.

Glossary

Abbreviation	Description
ADR UK	Administrative Data Research UK
ALF	Anonymous Linkage Field
APCds	Admitted Patient Care dataset
DHCW	Digital Health and Care Wales
HB	Health Board
IGRP	Information Governance Review Panel
NCCH	National Community Child Health
NWIS	NHS Wales Informatics Service
ONS	Office for National Statistics
SAIL	Secure Anonymised Information Linkage (SAIL) Databank

Disclaimer

This work was produced using administrative data accessed through the SAIL Databank. The use of the data in this work does not imply the endorsement of the SAIL Databank or data owners in relation to the interpretation or analysis.

This work uses research data which may not exactly reproduce National Statistics aggregates.

Acknowledgements

This study uses anonymised data held in the Secure Anonymised Information Linkage (SAIL) Databank, which is part of the national population data research infrastructure for Wales. We would like to acknowledge all the data providers who make anonymised data available for research. This study was carried out as part of the ADR Wales programme of work. ADR Wales unites research expertise from Swansea University Medical School and WISERD (Wales Institute of Social and Economic Research and Data) at Cardiff University with analysts from Welsh Government.

This Data Insight has been produced by the ADR Wales Ymateb research team. It provides a snapshot of informative research currently underway at ADR Wales but is not intended to provide a complete picture of work undertaken within this field or the ADR Wales programme of work. The information presented in this Data Insight has been reviewed by ADR Wales colleagues with expertise within this thematic area and accepted to be accurate at the point of publication. Views expressed in this Data Insight are those of the researchers and not necessarily those of ADR Wales partner organisations.

The work is supported by ADR UK (Administrative Data Research UK). ADR UK is a partnership transforming the way researchers access the UK's wealth of public sector data, to enable better informed policy decisions that improve people's lives. ADR UK is an Economic and Social Research Council (ESRC) investment (part of UK Research and Innovation). [Grant number: ES/W012227/1].

Publication contact

Please contact the SAIL Databank via the Helpdesk.