

International Journal of Population Data Science

Journal Website: www.ijpds.org



Swansea University
Prifysgol Abertawe

Data Resource Profile: Maternity Services Dataset (MSDS)

Catherine Stewart^{1,*}, Jennifer Hall¹, and Katie Harron²

Submission History

Submitted:	21/05/2025
Accepted:	20/10/2025
Published:	02/03/2026

¹Reproductive Health, Institute for Women's Health, University College London

²Population, Policy & Practice, Institute of Child Health, University College London

Abstract

Introduction

The Maternity Services Dataset (MSDS) is a comprehensive administrative dataset, capturing detailed information on NHS-funded maternity services in England. It encompasses the full maternity care pathway. There are two versions of the dataset: MSDS V1.5, introduced in May 2015, and the updated MSDS V2.0, which replaced it in April 2019.

Methods

MSDS is a secondary-use dataset, repurposing clinical and operational data primarily for monitoring, commissioning, and auditing purposes. Data is submitted monthly by trusts, enabling NHS England to generate a national picture of maternity activity and publish official statistics. Patient identifiers enable linkage with other NHS datasets, such as Hospital Episode Statistics (HES).

Results

The dataset includes demographic, clinical, and organisational information for each pregnant person and their baby, such as appointment dates, diagnoses, delivery details, birthweight, and neonatal admissions. Provider and geographical data are also recorded. Data quality varies due to the classification of variables as mandatory, required, or optional. Most variables are required, meaning submission is expected when applicable, but not enforced, leading to inconsistencies in data completeness across trusts. Only a small number of fields are mandatory, meaning that data must be present for a successful submission.

Conclusions

The MSDS is used for national and local health service planning, including major audits such as the National Maternity and Perinatal Audit. Although it holds potential for research, uptake has been limited to date due to its relative novelty. As coverage improves and the number of available years increases, the MSDS will become an increasingly robust resource for conducting research. It contains a wealth of information from the maternity services pathway - much of which is not available elsewhere - making it a highly valuable dataset for pregnancy-related research. Access to the MSDS is provided by NHS England. Data are requested through the online Data Access Request Service.

Keywords

Maternity Services Dataset; Administrative Data; Maternity Services; NHS

Key Features

- MSDS contains data on all NHS funded maternity services in England.
- There are two versions of MSDS: V1.5 (which runs from May 2015 to March 2019) and V2.0 (which went live in April 2019). While data are available from financial year 2015/2016, there is limited usability until financial year 2018/2019 due to coverage.
- Detailed demographic, clinical and organisational information is available for each pregnant person (throughout their maternity services pathway) as well as demographic and clinical information for the baby.
- Data are linkable to other datasets held by NHS England, including Hospital Episode Statistics (HES) Admitted Patient Care (APC), HES A&E, Emergency Care Data Set (ECDS), HES Outpatient, Mental Health Services Data Set (MHSDS) and Community Services Data Set (CSDS).
- Access to MSDS is provided by NHS England. Data are requested through the online Data Access Request Service.

*Corresponding Author:

Email Address: catherine.stewart.20@ucl.ac.uk (Catherine Stewart)

Background

The Maternity Services Dataset (MSDS) contains information on activity carried out by NHS-funded Maternity Services in England, from the first booking appointment until the person who was pregnant (and baby) are discharged from maternity services, which occurs when they leave hospital after birth, or after the pregnancy has ended. The MSDS also includes information about stillbirths and some limited data on miscarriages and terminations. MSDS data are collated and curated by NHS England. MSDS is a patient level data set that captures data on key information from each stage of the maternity care pathway including pregnant person's demographics, antenatal appointments, hospital admissions and re-admissions, screening tests, labour and delivery as well as baby's demographics, neonatal unit admissions, diagnoses and screening tests.

There are currently two versions of MSDS: version 1.5 and version 2.0. V1.5 went live in May 2015 and was subsequently replaced by V2.0 in April 2019. Although MSDS V1.5 is available from financial year 2015/16, there is limited use of the data until the financial year 2018/19, due to lack of coverage which is defined as the proportion of providers submitting data, reflecting the dataset's national representativeness. In 2016/17, only 56% of all deliveries (both live births and stillbirths) reported in Hospital Episode Statistics (HES) Admitted Patient Care (APC) were also reported in MSDS V1.5, in 2017/18, this figure increased to 78% and by 2018/19, approximately 91% of deliveries reported in HES APC were also reported in MSDS V1.5. During the transition from MSDS V1.5 to V2.0 (2019/20 and 2020/21), there was a slight dip in coverage. However, since 2021/22, coverage has been steadily increasing, and for MSDS V2.0 2023/24 there was only a 4% difference in the number of deliveries reported in HES APC and MSDS (HES APC reported 545,149 deliveries, while MSDS reported 524,745 deliveries) [1]. It is interesting to note that the MSDS may contain some information on some pregnancies and/or deliveries that are not captured in HES APC, for example, some pregnancies not delivered in a hospital or some pregnancies that ended in a miscarriage (only miscarriages that occur after a booking appointment and that are reported to the hospital).

The transition from MSDS V1.5 to MSDS V2.0 came as a result of the 2016 National Maternity Review [2]. The update utilises a new structure and content (including new variables, classifications and codes) and therefore direct comparison between the two versions is not possible.

The data within MSDS is primarily collected and used for clinical purposes (direct patient care). However, MSDS is a secondary use, administrative dataset that re-uses clinical and operational data for purposes other than direct patient care, notably monitoring and commissioning. Data is submitted to NHS England monthly and is used to create a national picture of maternity services activity, with official statistics being published by NHS Digital monthly [3].

MSDS structure

Within MSDS data files are structured according to financial years (April to March).

MSDS V1.5 structure

For MSDS V1.5 all data are presented within one table, with each row comprising all the maternity services provided (at each stage of the maternity care pathway, from booking to delivery and discharge) to a pregnant person and their baby by a single provider (NHS Trust). If a pregnant person changes provider a new row will be created; if a pregnant person has a multiple pregnancy, there will be separate rows for each baby (each row will contain the same information about the pregnant person) and if a pregnant person has two (or more) pregnancies within the same financial year, there will be multiple rows.

MSDS V2.0 structure

For MSDS V2.0 data is split into 11 tables: 'demographics, booking and pregnancy'; 'maternity care plan'; 'diagnoses and history'; 'findings and observations'; 'care activities'; 'coded scored assessments'; 'labour activities'; 'baby demographics'; 'baby activities'; 'baby diagnoses'; and 'hospital provider spell'.

Data is structured differently for each table. For example, within the 'demographics, booking and pregnancy' table, each row contains information about the pregnant person, and an overview of the maternity services provided (at different stages of the maternity care pathway) to a pregnant person by a single provider (NHS Trust). If a pregnant person changes provider a new row will be created; if a pregnant person has two pregnancies within the same financial year, there will be two rows. In the 'baby demographics' table, each row provides information about an individual baby. Therefore, for multiple pregnancies, there will be separate rows for each baby. The 'findings and observations' table contains information about all findings or observations recorded for an individual throughout the maternity care pathway. Within this table, each row contains information about a single finding or observation, so an individual pregnant person will have numerous rows within this table, regardless of number of children, providers or pregnancies (this is the same for the 'baby activities' table). As a result of the variation in structure, the different tables have different numbers of rows and require careful cleaning before linkage of tables can occur.

Linkage within MSDS

A local patient identifier is a mandatory variable that must be submitted for each individual, alongside their appointment information. This allows an individual's data from multiple maternity care appointments to be linked together. Within V1.5 these patient identifiers (MSDS_ID_Mother and MSDS_ID_Baby) have been used to create a single observation (within a single table) for each person along the whole maternity services pathway (provided by a single provider). In V2.0 these patient identifiers (Person_ID_Mother and Person_ID_Baby) allow MSDS users to link data within and across the 12 different tables, providing a comprehensive view of the entire maternity services pathway. These patient identifiers not only allow an individual's data to be linked within a single MSDS year, but

also allows longitudinal linkage between different MSDS years. However, these unique patient identifiers (in V1.5 and V2.0) are created using different methodologies (MSDS V1.5 derives the identifiers from a combination of identifiable data items, while MSDS V2.0 uses the Master Person Service (MPS)), and therefore are not consistent across the two datasets meaning longitudinal linkage of individuals across V1.5 and V2.0 is not currently possible. However, according to NHS England, there are plans to retrospectively apply the MPS to MSDS V1.5 data, which would make it possible to track mothers who have babies over the lifespan of both datasets in the future [4].

Linkage to other datasets

MSDS data (both V1.5 and V2.0) can be linked to other datasets held by NHS England, including Hospital Episode Statistics (HES) Admitted Patient Care, HES Accident and Emergency, Emergency Care Data Set (ECDS), HES Outpatient, Mental Health Services Data Set (MHSDS) and Community Services Data Set (CSDS). Users can link these datasets using a unique identifier provided by NHS England (currently the Token Person ID), allowing direct linkage of patient records across the different data sets.

MSDS data can also be linked to external datasets, for example, the National Pupil Database (NPD), as done in ECHILD [5] (using deterministic linkage, and Token Person IDs) or the National Congenital Anomaly and Rare Disease Registration Service data as done by the EXPRESS Study (using NHS numbers) [6, 7]. There are also plans to add the MSDS to OpenSAFELY [8].

Measures

MSDS provides detailed information about all activities carried out by NHS-funded Maternity Services, including the booking appointment, antenatal scans, delivery and discharge. Detailed demographic, clinical and organisational information for each pregnant person as well as demographic and clinical information for their baby is provided. There are dates for antenatal appointments, delivery and discharge, as well as information about diagnoses and findings recorded throughout pregnancy (e.g. information about pre-existing health conditions and diagnoses made during pregnancy). Data relating to the baby includes delivery method, presentation, gestational length, sex, weight and neonatal unit admission. Information on the care provider and geographical variables (based on a patient's postcode) are also available. Socioeconomic status is measured by the Index of Multiple Deprivation 2015 (IMD), a small area-based indicator constructed from several different measures of deprivation. Individual-level measures of socioeconomic status, such as educational attainment and income are not available, however information about employment status (and employment status of partner) is available (although may be subject to missing data). Detailed information on the variables available, the specific cleaning rules applied, and the coding used are available in the Maternity Services Data Set (MSDS) Technical Output Specification [9, 10].

Within MSDS, some of the variables are classed as mandatory (meaning submission cannot be completed without

these data being present), some are required (these variables 'must' be reported where they apply, however failure to submit these items will not result in the rejection of the record) and some are optional (data items may be submitted on an optional basis at the submitter's discretion). Most variables are classed as 'required', with only a few variables being mandatory (or optional). As a result, the completeness of variables is inconsistent and there is fairly significant variation between the data submitted by each trust. Figures 1 and 2 are heatmaps, showing the variation in completeness of a select number of variables – these variables were chosen as they are all collected as part of the booking appointment, and therefore missingness should not be impacted by pregnancy outcome. Finally, there is also some variation in the classification of variables in MSDS V1.5 and V2.0.

A full list of the mandatory and required variables can be found within the NHS Data Model and Dictionary resource [11]. A simplified list of MSDS V1.5 and V2.0 variable names, descriptions and missingness, can be found in the supplementary data.

Users should also check the NHS Maternity Services Monthly Statistics [3] (on the NHS England website) for any known issues with the data published for the year(s) of interest. Any issues with the data will be detailed here, for example if any trusts have missing MSDS submissions or whether any trusts had fewer records than expected.

Coding within MSDS

Within MSDS V1.5 the main code used is NHS National Code, which includes diagnoses variables. Within MSDS V2.0 the main codes used are SNOMED Codes and International Classification of Diseases (ICD) 10 codes (Read Coded Clinical Terms Version 2 and Read Coded Clinical Terms Version 3 can also be used but are rare). It is important to note, that within V2.0 findings and observations identified during the pregnancy will be recorded using SNOMED Codes (or ICD10 codes), this includes information about a pregnant person's height, weight, smoking status and diagnoses. This information can be found in the 'Findings and Observations' table, the 'Care Activities' table or the 'Diagnoses and History' table. To use this data for analysis, users must derive variables (e.g. Smoking Status, BMI or previous diagnoses) from these codes.

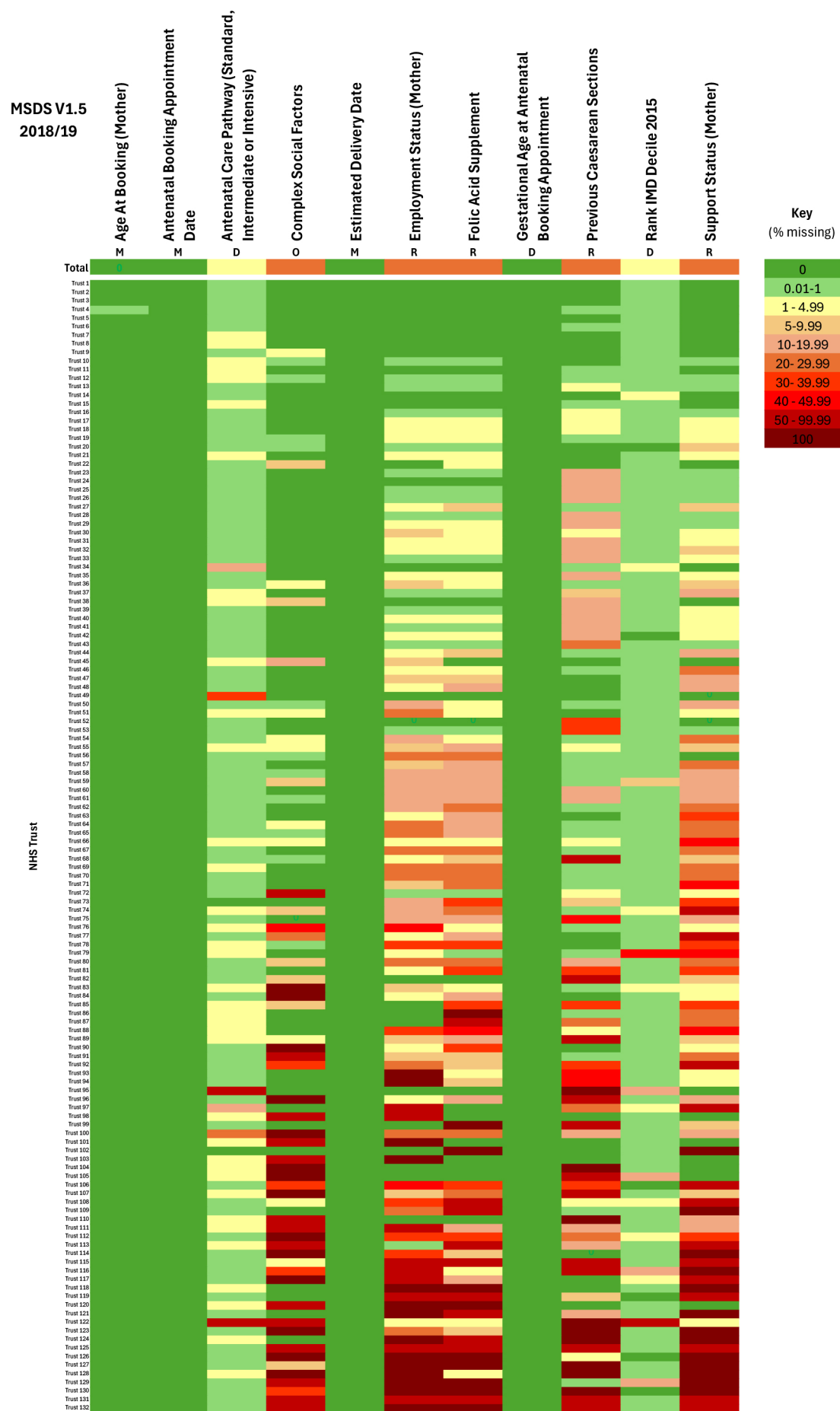
Resources

Resources for MSDS V1.5 and V2.0 are available online, through NHS England [9, 10]. These resources include a Data Model (a visual representation of the dataset), a Technical Output Specification (an excel spreadsheet that defines each data item within the data set), as well as other resources aimed at those submitting the data. NHS Data Model and Dictionary resource contains information about all mandatory and required variables for MSDS V2.0.

Discussion

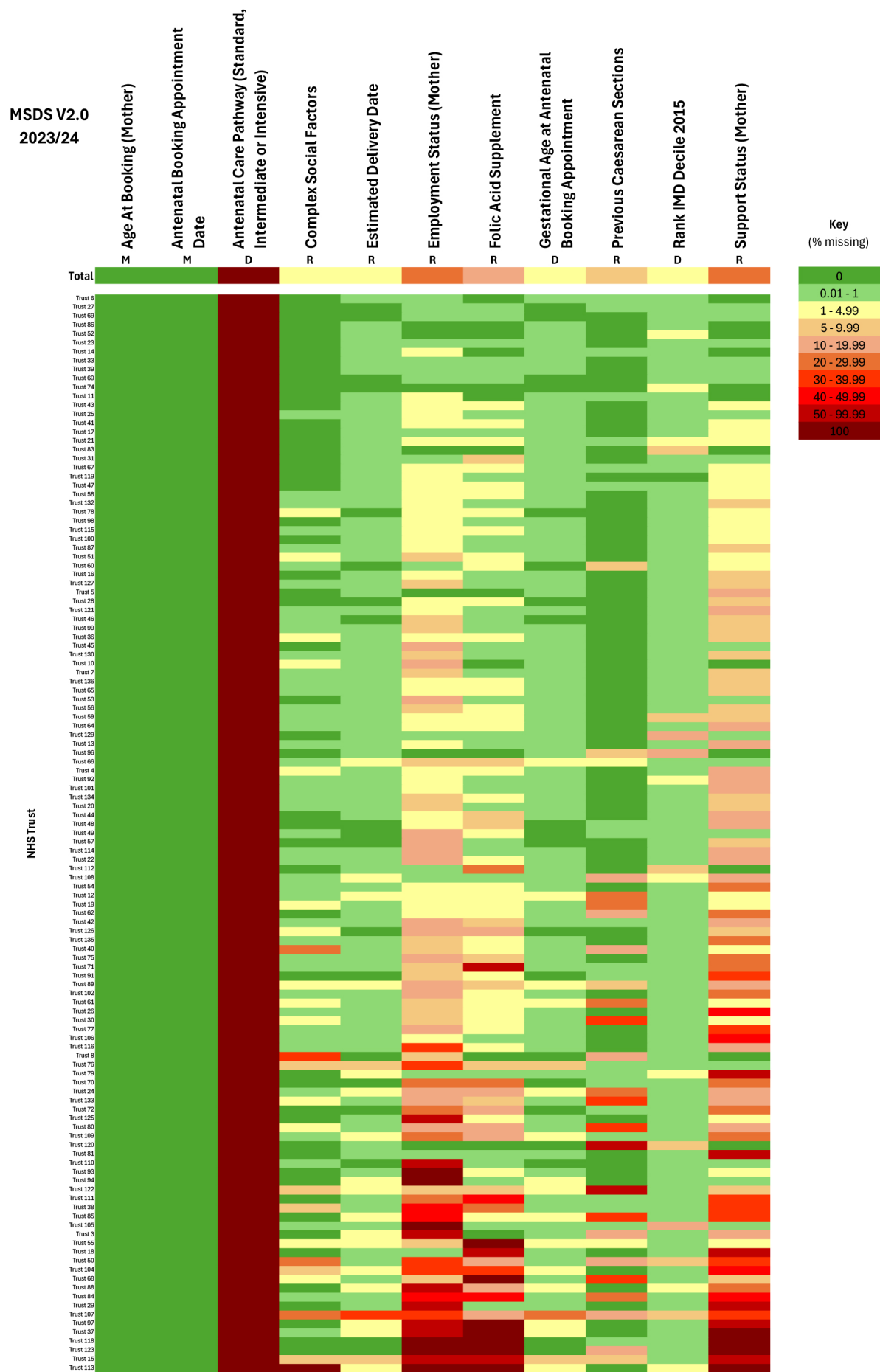
MSDS has been used for local and national monitoring, commissioning and clinical audit; for example, the National Maternity and Perinatal Audit [12] used MSDS data for a

Figure 1: Heatmap of variable completeness between NHS Trusts MSDS V1.5 2018/19



M = mandatory variable, R = required variable, O = optional variable D= derived variable.

Figure 2: Heatmap of variable completeness between NHS Trusts MSDS V2.0 2023/24



M = mandatory variable, R = required variable, O = optional variable D= derived variable.

large-scale audit of the NHS maternity services. MSDS can also be used for research, however as the dataset is relatively new, it has not yet been used extensively. We carried out a search on Medline on 13th October 2025 using the search term 'Maternity Services Data Set' and identified six relevant publications [6, 7, 13–16]. However, as coverage improves (and the number of years available increases) MSDS will become an increasingly robust resource for conducting research.

Strengths and weaknesses

One of the key strengths of the MSDS is its national coverage. It is a large dataset, which allows rare pregnancy outcomes (such as stillbirth) to be investigated. MSDS also contains a high level of detail, collecting data at each stage of the maternity service care pathway, and includes a lot of information which is not available in any other dataset, including HES. Therefore, the scope of research using MSDS is wide. Another strength is the ability to longitudinally link patients using Person IDs, both within MSDS (although limited to linkage within versions and constrained by low coverage in early years) and with other NHS England datasets (including HES); allowing follow up of patients, enabling researchers to investigate long-term outcomes. As coverage and available years increase, MSDS will become a more robust research resource.

One of the main weaknesses of MSDS is its maturing status. While all NHS-funded maternity service providers are meant to submit MSDS data, some providers are not making full submissions and coverage of the data is not universal. For example, all deliveries reported within HES are not currently reported within MSDS (although this figure is improving as the dataset matures). Therefore, users should be aware of potential biases of the data, such as geographical or provider-level bias, and some caution should be taken when interpreting the data at a national level. Detailed MSDS Data Quality Statements and CSV files presenting data quality are available with each annual NHS Maternity Statistics publication [17]. Another weaknesses of MSDS is missing data. Since not all variables are mandatory, completeness varies both between variables and across NHS trusts (see Figure 1 and Figure 2). For example, although both *Complex Social Factors* and *Mother's Employment Status* are required variables, *Complex Social Factors* is only 4.7% missing, while *Mother's Employment Status* is 22.2% missing in MSDS V2.0 2023/24. Across NHS trusts, completeness varies further: in some trusts, *Complex Social Factors* is 100% complete, while in others it is entirely missing. Users should therefore be cautious of the potential biases introduced by these differences in data completeness. Finally, because of the transition from MSDS V1.5 to MSDS V2.0, direct comparison between V1.5 years and V2.0 years is not possible, which impacts the ability to longitudinally link patients within MSDS (however, this may change in the future, due to plans to retrospectively apply the MPS to MSDS V1.5 data).

Data access

Access to MSDS data is provided by NHS England. Data are requested through the online Data Access Request Service [18]. While there are no direct charges for the data, there

are charges to cover the cost of processing and delivering the data. Ethical approval would usually be required prior to data access.

Conclusions

The MSDS offers rich, detailed clinical data from the maternity services pathway, much of which is unavailable in other datasets. It also has strong linkage potential, as the use of NHS unique identifiers enables researchers to link the MSDS with other NHS datasets (e.g. HES) with relative ease. However, several challenges remain, including incomplete submissions, variable data quality, and incompatibility between versions 1.5 and 2.0, which can limit longitudinal analyses. As data coverage improves and more years of data become available, the MSDS (V2.0) will become an increasingly robust resource for conducting research.

Supplementary data

A full list of the variables available within the MSDS, and the level of missingness of each variable can be found in the supplementary data.

Funding

CS is funded by Medical Research Council (MR/W006774/1). ECHILD is funded by Administrative Data Research UK (ADR UK), an investment by the Economic and Social Research Council (part of UK Research and Innovation) through the following grants: ES/V000977/1, ES/X000427/1 & ES/X003663/1.

Funding

This research benefits from and contributes to the NIHR Children and Families Policy Research Unit (CPRU) (NIHR206114) but was not commissioned by the National Institute for Health and Care Research (NIHR) Policy Research Programme. The views expressed are those of the author(s) and not necessarily those of the NIHR or the Department of Health and Social Care. CS is funded by Medical Research Council (MR/W006774/1).

Ethics

This work did not require formal ethical approval as it involved the secondary, exploratory use of existing data only. All data were accessed and used in accordance with an existing data sharing agreement, and were handled in compliance with applicable data protection and governance requirements

Conflict of interest

No conflicts of interest to declare.

AI Disclosure Statement

The authors declare that no generative AI tools were used in the preparation of this manuscript.

References

- England N. Comparison of data common to HES and MSDS. 2024; Available from: <https://digital.nhs.uk/data-and-information/publications/statistical/nhs-maternity-statistics/2023-24/hes-msds-comparison>
- England N. National Maternity Review Report – Better Births. 2016; Available from: <https://www.england.nhs.uk/mat-transformation/implementing-better-births/mat-review/>
- England N. Maternity Services Monthly Statistics. 2025; Available from: <https://digital.nhs.uk/data-and-information/publications/statistical/maternity-services-monthly-statistics>
- England N. Maternity Services Monthly Statistics, Final July 2024, Provisional August 2024, official statistics. 2024; Available from: <https://digital.nhs.uk/data-and-information/publications/statistical/maternity-services-monthly-statistics/final-july-2024-provisional-august-2024-official-statistics/methodology>
- Mc Grath-Lone L, Libuy N, Harron K, Jay MA, Wijlaars L, Etoori D, et al. Data Resource Profile: The Education and Child Health Insights from Linked Data (ECHILD) Database. International Journal of Epidemiology. 2022;51(1):17-f. <https://doi.org/10.1093/ije/dyab149>
- Hill MA-O, Ellard S, Fisher J, Fulop N, Knight MA-O, Kroese M, et al. Optimising Exome Prenatal Sequencing Services (EXPRESS): a study protocol to evaluate rapid prenatal exome sequencing in the NHS Genomic Medicine Service. NIHR Open Res. 2022 (2633-4402 (Electronic)). <https://doi.org/10.3310/nihropenres.13247.2>
- Ramakrishnan R, Mallinson C, Hardy S, Broughan J, Blyth M, Melis G, et al. Implementation of a national rapid prenatal exome sequencing service in England: evaluation of service outcomes and factors associated with regional variation. Frontiers in Genetics. 2024;Volume 15 - 2024. <https://doi.org/10.3389/fgene.2024.1485306>
- Higgins R. Collaborate with us to advance pregnancy and maternity research in OpenSAFELY. 2024; Available from: <https://www.bennett.ox.ac.uk/blog/2024/11/collaborate-with-us-to-advance-pregnancy-and-maternity-research-in-opensafely/>
- England N. MSDS v1.5 archived guidance documents. 2022; Available from: <https://digital.nhs.uk/data-and-information/data-collections-and-data-sets/data-sets/maternity-services-data-set/archived-guidance-documents>
- England N. Implementing the Maternity Services Data Set (MSDS) v2.0 tools and guidance. 2024; Available from: <https://digital.nhs.uk/data-and-information/data-collections-and-data-sets/data-sets/maternity-services-data-set/tools-and-guidance>
- England N. NHS Data Model and Dictionary – MSDS. Available from: https://www.datadictionary.nhs.uk/data_sets/clinical_data_sets/maternity_services_data_set.html?hl=msds
- Audit NMaP. National Maternity and Perinatal Audit Clinical Report 2022. 2022.
- Schoenaker DAJM, Stephenson J, Smith H, Thurland K, Duncan H, Godfrey KM, et al. Women's preconception health in England: a report card based on cross-sectional analysis of national maternity services data from 2018/2019. BJOG Int J Obstet Gynaecol. 2023 2023/09/01;130(10):1187-95. <https://doi.org/10.1111/1471-0528.17436>
- Mensah AA-O, Stowe J, Jardine JA-O, Kirsebom FCM, Clare TA-O, Kall M, et al. COVID-19 Vaccine Safety in Pregnancy, A Nested Case-Control Study in Births From April 2021 to March 2022, England. BJOG. 2024 (1471-0528 (Electronic)). <https://doi.org/10.1111/1471-0528.17949>
- Roebuck C, Sandall J, West R, Atherden C, Parkyn K, Johnson O. Impact of midwife continuity of carer on stillbirth rate and first feed in England. Communications Medicine. 2025 2025/08/07;5(1):339. <https://doi.org/10.1038/s43856-025-01025-z>
- Gillibrand S, Parkyn K, Hall C, Kletter M, Harkness E, Munford LA, et al. Releasing time to deliver care: a mixed methods evaluation of the implementation of enhanced midwifery continuity of carer. BMJ Open. 2025;15(7):e095509. <https://doi.org/10.1136/bmjopen-2024-095509>
- England N. NHS Maternity Statistics. 2024; Available from: <https://digital.nhs.uk/data-and-information/publications/statistical/nhs-maternity-statistics>
- England N. Data Access Request Service (DARS). 2025; Available from: <https://digital.nhs.uk/services/data-access-request-service-dars>

