Considerations on The National Neonatal Research Database (NNRD) in the UK National Health System

Divya Ramesh

University of Arizona School of Nursing and School of Information

LIS 634/NURS 634 – Data Management in Healthcare Systems

Project 1 – Database Analysis

Professor Sheila Gephart, Ph.D

February 20, 2024

Database Description:

The National Neonatal Research Database (NNRD) [URI[1]: https://web.www.healthdatagateway.org/dataset/67020745-9def-4c6e-b5ac-bb273bd0a20e] is a national resource in the United Kingdom, developed and managed by various interdisciplinary professionals at Imperial College London. It holds the growing Neonatal Data Set that comprises 1) daily episodic care data and 2) two-year neonatal outcomes assessment data (National Health Service, 2023). Developed to support research studies and policy research with community partners, it catalogs "pre-defined clinical data extracted at regular intervals from point-of-care, [and] clinician-entered electronic patient records on all admissions to National Health Service neonatal units in England, Wales, and Scotland" and the Isle of Mann (Battersby et. al., 2018). The National Data Analysis Unit has developed a parent-facing brochure that describes data collection to parents in plainer English: "These are your data. They originate from information entered by doctors and nurses into Badger.net[2]. All 200 neonatal units in England, Wales and Scotland contribute data and form the UK Neonatal Collaborative."

As Modi (2019) writes, the current mission of the NNRD is "to provide a single source of high-quality data that diverse users could use for multiple purposes to improve neonatal care, support neonatal services, facilitate research and develop national policies" (193). Perhaps due to the active migration process of the NNRD into the United Kingdom's (UK) National Health Service system, another instance of it also seems to exist [URI: https://digital.nhs.uk/data-and-information/information-standards/information-standards-and-data-collections-including-extractions/publications-and-notifications/standards-and-collections/dapb1595-neonatal-data-set] with a more updated release and review date of 2022 and slightly-different data specifications than those included at the Imperial College URI.

[1] URI is an abbreviation for Unique Resource Identifier

[2] Badger.net is an electronic maternity healthcare record system used widely within the NHS system in the United Kingdom.

History of Development

When the Department of Health Social Care convened a working group in 2001-2002, concerned about "the common transfer of babies and pregnant women inappropriately in an unplanned fashion and over long distances" self-sufficient neonatal networks reorganized services in response, with £20 million issued annually to support the project starting in 2006 (Marlow & Gil, 2007). As Marlow & Gil (2007) outline, financial support for this effort "devolved" and needed to be "disentangled" due to disbursement disinclination by the Primary Care Trusts and a "lack of engagement, particularly in local district hospitals" who maintained a "degree of rejection of the model." Largely, it was advanced and implemented by neonatal specialist practitioners whose Copernican pull emerged a care network of staff around themselves, until these isolated specialist-backed smaller networks aimed to connect amongst themselves to emerge a network model from the ground up.

Given inconsistent adoption of networked neonatal care model where there was a lack of specialist-advocacy in certain areas across the United Kingdom, it became crucial to measure centers' success over the longer term to justify benefits. So emerged the NNRD, an "attractive" hope that "ongoing national collated data" would further a "noble goal for neonatology generally" and "show a reduction in neonatal mortality and other morbidity measures" over time in the networked care model (Marlow & Gil, 2007).

Broader Database Aims Towards Condensed Data Collection Effort

As the NNRD has grown into the present, so too have its aims. While it still focused on "secur[ing] strong clinician engagement" into a national resource, the NNRD of recent years is also expanding to include parent perspectives: understanding parents' "primary responsibility" in the development and maintenance of clinical database of infant health information while infants cannot yet provide "autonomous consent;" researchers in Manchester are co-designing parent surveys around parent acceptance, motivation, and inclination to provide infant health data for research and including

these in an extensible database as new variables or on paper notes after collection and summarization (Modi et. al., 2019) that as of most recent report are in the process of being stored electronically in the maternity information portion of the database (Information Standards Board for Health and Social Care, 2014).

More generally, this broadened approach to information collection highlights the database's growing "focus on data quality" theoretically by including complementary adults' perspectives in a child's life but so too practically by examining automations and workflows so that data "should be recorded once and not repeatedly, recorded to a high standard, and made available to support multiple outputs" (Modi, 2019). As the Neonatal Medicine Research Group describes on their website, the work focuses on using "'real-world data'" recorded over the course of routine care" towards research that is to be "conducted more efficiently, rapidly and at lower cost than currently possible" (Imperial College London, 2024).

While direct estimates of database maintenance cost have been difficult to locate in the literature, the Neonatal Data Analysis Unit describes an established costing model to self-sustain, stating in a brochure that it "receives no core funding" and a "charge is necessary" on a "project-by-project basis" that covers "data transfer, extraction, cleaning, storage, and any analysis requested" of the dataset (Neonatal Data Analysis Unit, 2016).

Cost Containment-Inspired Aims for Efficiency

Given the public good origins of this enterprise where "changes in medical manpower and the scarcity of neonatal nurses" (Marlow & Gill, 2007) in the first instance emerged a database solution whose present aim is in part to increase efficiency and reduce the medical burden of duplicated neonatal entry (Statnikov et. al., 2017) via a "single national data source to service multiple outputs" (Modi et al., 2019), it is possible that the posted project-by-project example charges by the Neonatal Data Analysis Unit of "£10,000 for complex national quality improvement projects" that are recommended to be "built

into funding applications" equate maintenance costs for various arms of the NNRD without markups or profits built into the posted price.

Description of Data Contents and Database Features

Types of Data and Data Format:

In its initial conception, the NNRD "built on and extended the British Association of Perinatal Medicine 'minimum' data set" and too the National Critical Care Minimum Data Set to ensure that interoperability across data systems remained (Modi et. al., 2019). Modi et. al., (2019) describes the initial data types in the NNRD given these efforts:

> "[The database] comprised basic demographic details (e.g. date of birth, birthweight), clinical interventions captured daily (e.g. respiratory support, type of feeds, surgical procedures, high-cost drugs), clinical outcomes and diagnoses. Each data item is clearly defined in an accompanying metadata set, and mapped to existing national standards as well as ICD codes" (5)

However, upon "'advance notice'" from the Information Standards Board that the Neonatal Data Set and NNRD would evolve into becoming a national standard, 1) database co-design with the public and 2) terminology reviews for concordance with the NHS data dictionary added an additional 25 data items to the Neonatal Data Set around demographics within the NNRD and 28 items—around specialty care—to invite more user engagement and improve the usability of the database. Examples of data types and format for some variables are included in the table below:

Examples from the NNRD Data Dictionary (National Health Service, 2023) and Modi, et. al., (2019):

- Category: Daily General Information - Category Detail: General Information - Item Name: OneToOneNursing - Field ID: R (Required by National Health Service) - Data Dictionary Item Status: PATIENT RECEIVING ONE TO ONE NURSING CARE INDICATOR - NHS Item Name: OneToOneNursing - Format: an1 - Coding: N (no), Y (yes) - Description: Specify if the baby received 1:1 nursing on this day of care - Purpose: Used to determine the level of care	- Category: Daily Ophthalmology - Category Detail: ROP screen - Item Name: ROPScreen - Field ID: R - Data Dictionary Item Status: RETINOPATHY OF PREMATURITY SCREENING PERFORMED INDICATOR - NHS Item Name: Not specified - Format: an1 (alphabetic, numeric with length 1) - Coding: N for no, Y for yes - Description: Specify if the baby has had a screen for ROP on this day of care

	- Purpose: Used in the National Neonatal Audit Programme

Also excerpted from the NNRD Data Dictionary are the “mandations”: (National Health Service, 2023)

M = Mandatory: this data element is mandatory and the technical process (e.g. submission of the data set, production of output etc) cannot be completed without this data element being present
R = Required: NHS business processes cannot be delivered without this data element
O = Optional: the inclusion of this data element is optional as required for local purposes
P = Pilot: this data element is for piloting use only.

A note that having a controlled vocabulary for pilot data elements supports further development and extensibility of the database. As reported by Modi (2019), the NNRD hopes to later include, long term follow up into adulthood, and “physiological monitoring, imaging and genomics data” (196).

Data entered into pilot elements may also inform user behavior and the development of USE/USE FOR and foreign key terms. These could support future goals to link data to other health datasets such as the as the "UK NHS Hospital Episodes Statistics, Office for National Statistics data, and National Maternity Dataset" (Modi, 2019, 196).

Data Source and Frequency of Collection: Parent and clinician entry into an infant’s electronic patient record that is updated quarterly per Modi (2019).

Report Generation: It does not look as though this database generates reports. However, it seems as though the files are downloadable via CSV or other form to be viewable in another application. It seems as though studies that use this database make a request and are provided files of the specific variables they would need for analysis, and the variable groups are provided publicly (for example, dailydata, episodes, masterlist, radiology, etc).

Technical Requirements for Use:

It is difficult to find any hardware or software information around data extracted from this database, despite reviewing the National Health Service (2023) specification documents. Modi et. al.,

(2019) includes a Figure 2 on page 5 of the article that suggests that the Clevermind app interface is used foe ease of user-side data entry.

Examples of Database Use in Literature:

Interestingly, the NNRD includes a 'Collections' section that describes its database implemented within or interoperable with other databases such as a maternity database and the UK neonatal collaborative.

The database in recent years has been used to evaluate the data completeness. van Blankenstein et al., (2022) conduct a retrospective cohort study with NNRD data to a) determine if a "2-year neurodevelopmental assessment record in the NNRD, use of standardised assessment tools, results of functional 2-year neurodevelopmental assessments (visual, auditory, neuromotor, communication, overall development)" is present to then determine rates natures and proportions of disability diagnosis and rates of follow up care. They find that "the proportion of infants admitted to neonatal units in England and Wales with a 2-year neurodevelopmental record has improved over time" while "[r]ates of follow-up data from recent years are comparable to those of bespoke observational studies" which promises potential for the utility of the database. Using the database to better understand the needs of high-need infants seems to be a continued focus. Harnden et. al., (2023), in one example used the dataset to help determine follow up rates for neonates in palliative care.

Strengths and Limitations of the Dataset:

<u>Strengths:</u> Per Modi (2019), "Data held in the NNRD [was] complete and accurate for key data items; for example an evaluation of data completeness and quality of data items held in the NNRD compared to data independently collected as part of a parallel multi-centre randomised controlled trial show excellent agreement." She ascribes this data quality to a data error-checking process that identifies mistakes in the baby's electronic patient record via continual review, communicates these to the clinical

team that then corrects them such that the subsequent data download is then correct upon next verification. Modi (2019) also notes that parents are given choice to participate in the database, although none have yet declined (195-196).

Limitations: However, a limitation is also identified in how information about infant in-delivery-room deaths, stillbirth, and transitional care is excluded from the NNRD. The database also does not contain laboratory or genetic data (Modi, 2019), although space has been made for it in the data dictionary, as described in Modi et. al., (2019) in Appendix 3 for future inclusion.

Also, "reliance on clinician entry at the point of care places constraints upon the inferences that can be made regarding accuracy," (Modi, 2019), and future directions for the database hope to improve usability and automate this process via machine learning technologies (Modi et. al., 2019).

References:

Battersby, C., Statnikov, Y., Santhakumaran, S., Gray, D., Modi, N., Costeloe, K., UK Neonatal Collaborative, & Medicines for Neonates Investigator Group. (2018). The United Kingdom National Neonatal Research Database: A validation study. PLOS ONE, 13(8), https://doi.org/10.1371/journal.pone.0201815

Harnden, F., Lanoue, J., Modi, N., Uthaya, S. N., & Battersby, C. (2023). Data-driven approach to understanding neonatal palliative care needs in England and Wales: a population-based study 2015–2020. Archives of Disease in Childhood-Fetal and Neonatal Edition.

Jawad, S., Modi, N., Prevost, A. T., & Gale, C. (2019). A systematic review identifying common data items in neonatal trials and assessing their completeness in routinely recorded United Kingdom national neonatal data. *Trials*, *20*(1), 1-12.

Marlow, N., & Gill, A. B. (2007). Establishing neonatal networks: The reality. *Archives of Disease in Childhood. Fetal and Neonatal Edition, 92* (2), F137–F142. https://doi.org/10.1136/adc.2005.086413

Modi. (2019). Information technology infrastructure, quality improvement and research: the UK National Neonatal Research Database. *Translational Pediatrics*, *8*(3), 193–198. https://doi.org/10.21037/tp.2019.07.08

Modi, Ashby, D., Battersby, C., Brocklehurst, P., Chivers, Z., Costeloe, K., Draper, E. S., Foster, V., Kemp, J., Majeed, A., Murray, J., Petrou, S., Rogers, K., Santhakumaran, S., Saxena, S., Statnikov, Y., Wong, H., & Young, A. (2019). Developing routinely recorded clinical data from electronic patient records as a national resource to improve neonatal health care: the Medicines for Neonates research programme. *Programme Grants for Applied Research*, *7*(6), 1–396. https://doi.org/10.3310/pgfar07060

National Health Service. (2023). National Neonatal Data Set - Episodic and Daily Care. NHS Data Dictionary. URL: https://www.datadictionary.nhs.uk/data_sets/clinical_data_sets/national_neonatal_data_set/national_neonatal_data_set_-_episodic_and_daily_care.html

National Neonatal Research Database (NNRD). (2010). [Neonatal health data set collected and maintained by the UK National Neonatal Research Database (NNRD) and the Neonatal Data Analysis Unit at Imperial College London.]. Health Data Gateway. https://web.www.healthdatagateway.org/dataset/67020745-9def-4c6e-b5ac-bb273bd0a20e

Neonatal Data Analysis Unit [NDAU]. (2016). *How to use the National Neonatal Research Database (NNRD) for research, audit and quality improvement: Information for clinicians and other health professionals*. Retrieved from https://s3.eu-west-2.amazonaws.com/sr-bliss/documents/Research-and-campaigns/Research/NDAU_leaflet_form.pdf

Statnikov, Y. (2014, January 17). Neonatal Data Set ISB 1595 Specification. Information Standards Board for Health and Social Care. https://digital.nhs.uk/binaries/content/assets/website-assets/data-and-information/information-standards/standards-and-collections/isb-1595-neonatal-data-set/1595322012spec.pdf

Statnikov, Y., Ibrahim, B., & Modi, N. (2017). A systematic review of administrative and clinical databases of infants admitted to neonatal units. *Archives of Disease in Childhood - Fetal and Neonatal Edition, 102*(3), F270. https://fn.bmj.com/content/102/3/F270

van Blankenstein, E., Sodiwala, T., Lanoue, J., Modi, N., Uthaya, S. N., Battersby, C., & UK Neonatal Collaborative. (2022). Two-year neurodevelopmental data for preterm infants born over an 11-year period in England and Wales, 2008–2018: A retrospective study using the National Neonatal Research Database. *Archives of Disease in Childhood - Fetal and Neonatal Edition, 109*(2), 143. https://fn.bmj.com/content/109/2/143.abstract