The **NNRD-AI** provides enhanced **flexibility and accessibility** for neonatal critical care admissions data

**NNRD-AI: a national neonatal research database for rapid insights with machine learning and artificial intelligence**

**Background:** The National Neonatal Research Database (NNRD) is a relational database that holds patient-level clinical admissions data drawn from Electronic Patient Records from all 181 NHS neonatal units in England, Wales, Scotland, and Isle of Man. Using the NNRD requires domain knowledge to correctly manipulate and interpret items. The NNRD-AI addresses these complexities by providing a curated dataset that is easy to use without prior domain knowledge and will advance the range of applications.

### Pipeline Structure

**NNRD**
- Episodes
- Daily Data
- Abdominal X-Rays
- Retinopathy of Prematurity Screening
- Sepsis Screening
- Brain Imaging
- External Reference Data

**Intermediary Processing**
- COVID-19 Infection
- Curated Clinical Outcomes
- Drug Variables
- Maternal Conditions
- Sociodemographic Proxies
- Static Items
- Time Conversions
- Transfers
- UK WHO Growth Variables

**NNRD-AI**
- `nnrd_ai_static`
- `nnrd_ai_daily`
- `nnrd_ai_agg`

**Methods:**
- The NNRD data is fed into the Python pipeline
- Curation includes:
  - Removal of improbable values
  - Outlier management
  - Type conversion
  - Growth standard deviation score calculation
  - Simplification of categorical variables
  - Conversion of nominal data to binary

**Results:**
- There are over 1.3 million babies from 181 units included in the NNRD-AI
- The three main outputs are:
  - `nnrd_ai_static`, baby-level table with static and clinical outcomes
  - `nnrd_ai_daily`, curated daily summary table
  - `nnrd_ai_agg`, baby-level table with daily summary items aggregated at different points across stay

### Example Application

One of the NNRD-AI intermediaries is a curated transfers table. This table includes information on admission and discharge times, discharge and receiving unit designation, and columns to indicate if the transfer occurred within the first 24-, 48- or 72-hours following delivery.

All clinical definitions and variables are validated against existing standard operating procedures used for the NNRD.

1. Neonatal Data Analysis Unit, Imperial College London
2. Institute for Translational Medicine and Therapeutics Data Science Group, Imperial College London

**Julia Lanoue**, Kayleigh Ougham, Neena Modi, Sam Greenbury

Link to Health Data Research UK Gateway for full data specification and request form