# OBUS – Fetal Presentation Data

## Overview

The fetal presentation (FP) model was trained and evaluated on the highly curated NEJM subset of the FAMLI2_enrolled dataset.

## Data Window

Babies start moving in the uterus between 16 and 22 weeks of pregnancy. In many cases, they will continue to change position until around 36 weeks. It is only after this gestational age that fetal presentation becomes relevant to monitoring pregnancy risk. However, the FAMLI dataset contains only a limited number of exams at 36 weeks GA or later, which would be insufficient for training and evaluating a deep learning model. Therefore, it was decided to broaden the data window and set the minimum GA threshold at 28 weeks (196 days). The statistics that are provided in the next section are based on this GA threshold.

## Label definition

The fetal lie/fetal presentation metadata for NEJM is annotated in the FAMLI2_enrolled case report form spreadsheet. There are five values of the us_lie parameter as shown in Table 1.

| **`us_lie` value** | **Interpretation** |
|---|---|
| -88 | N/A (unknown or unrecorded) |
| 1 | cephalic |
| 2 | breech |
| 3 | transverse |
| 4 | variable or N/A |
| 5 | oblique |

*Table 1. Unique values of* `us_lie` *and their interpretation*

The incidence of non-cephalic presentations in the dataset is much lower than cephalic. Based on these low data volumes it would be difficult to train accurate multi-class models. Therefore, the breech, transverse, and oblique classes were combined into a non-cephalic category. The “N/A” and “variable or N/A” exams are ignored for the purposes of the FP model development (not used for training or evaluation because the fetal presentation is

unknown). Notwithstanding this grouping, there is still imbalance in the dataset between cephalic and the grouped non-cephalic classes as noted below in Table 2.

## Data Distribution

The NEJM dataset included five subsets that were defined by the UNC team. The statistics of the data subsets used for fetal presentation are provided in Table 2. In the testing set, only vertical sweeps with tags M, NM, L0, L1, L2, NL, R0, R1, R2, NR, L15, ML, L45, R15, MR, R45 and exams with GA greater than the 196-day threshold are included in the tally. For the training and tuning sets, the same GA threshold is applied, but the sweep tags RTA, RTB, RTC, FA1, and FA2 are also included. Although these are not technically vertical sweeps, they are blind and add discriminative information as well as increase the amount of training data. They are not included in the testing split because they are not part of a standard use-case acquisition protocol. The splits are defined at the patient level.

The extreme imbalance between cephalic and non-cephalic labels can be seen from Table 2. Methods for compensating for the class imbalance are outlined in the README of the repository [1].

| Subset | Patients | Exams | Videos | Purpose |
|---|---|---|---|---|
| Training | 1,778 | 2,415 | 20,215 | Training model weights via backpropagation |
| - cephalic | 1,510 | 2,114 | 17,748 | |
| - non-cephalic | 268 | 301 | 2,467 | |
| Tuning | 440 | 588 | 4,962 | Optimizing hyperparameters |
| - cephalic | 376 | 510 | 4,287 | |
| - non-cephalic | 64 | 78 | 675 | |
| Testing | 616 | 785 | 4,757 | Estimating performance |
| - cephalic | 530 | 695 | 4,225 | |
| - non-cephalic | 86 | 90 | 532 | |

*Table 2. NEJM subsets by patients, exams, and videos beyond the 196 day GA threshold*

## References

[1] "GHL OBUS GitHub Repository," 2025. [Online]. Available: https://github.com/Global-Health-Labs/OBUS-GHL.