**Pneumo-DEEP file notes**

To protect the anonymity of study participants some epidemiological data has been obscured in the publicly available files. These entries will be filled with ‘A.O.R’ (available on request) in the associated tables. The original metadata files are available on request via the MORU Tropical Health Network Data Access Committee <https://www.tropmedres.ac/units/moru-bangkok/bioethics-engagement/data-sharing>.

**Pneumo-DEEP\_infant\_house\_metadata.csv**

An overall participant metadata summary: one participant per row

|  |  |
| --- | --- |
| **variable** | **definition** |
| codenum | Unique participant ID (shared between mother and infant) |
| group | NP swab processing group |
| twin\_code | Associated twin ID |
| immunology | Original cohort grouping (Immunology / Routine) |
| mother\_ethnic\_group | Ethnic group of mother |
| mother\_smoke | Mother smoked at time of enrolment (Yes / No) |
| infant\_dob | Infant date of birth |
| infant\_gender | Infant gender (FEMALE / MALE) |
| infant\_bw | Infant birth weight (g) |
| inf\_gestation | Infant gestation (weeks) |
| home\_delivery | Mother gave birth at home (Yes / No) |
| type\_delivery | Delivery mode (Normal (NVD), Breech, etc) |
| house\_latitude | House co-ordinate (lat) |
| house\_longitude | House co-ordinate (long) |
| house\_altitude | House altitude (m) |
| house\_rooms | Number of rooms in house |
| house\_fire | Fire inside the house (Yes / No) |
| house\_fuel | Type of fuel used in the house |
| house\_volume | Volume of house (m^3) |
| number\_household | Number of people living in the house at time of enrolment |
| num\_child\_less5y | Number of children <5y in the house |
| num\_child\_5\_14y | Number of older children (5-14y) in the house |
| num\_adult\_15\_59y | Number of adults in the house |
| num\_adult\_60y | Number of older adults (>=60y) in the house |
| study\_outcome | End point of participation (completed, lost, died, etc) |

**Pneumo-DEEP\_swab\_summary.csv**

NP swab summary for all ***processed*** swabs (i.e. excludes those not cultured / processed by latex sweep): one swab per row

|  |  |
| --- | --- |
| **variable** | **definition** |
| codenum | Unique participant ID (shared between mother and infant) |
| category | Participant category (Mother / Infant) |
| dob | Infant date of birth |
| age\_d | Infant age (days) at swab collection |
| Age\_v | Planned swab visit number (rounded to month) |
| specdate | Swab collection date |
| spectype | Explicit specimen type (should all be NP swab) |
| specnum | Unique specimen ID |
| acute | Swab collected at pneumonia diagnosis (Yes / No / Acute/Routine [pneumonia diagnosed on the same day as a planned swab collection] |
| whocult | Swab processed by WHO culture (yes / No): to identify those processed by latex sweep but also by WHO culture (i.e. may have an isolate sequenced already) |
| pnc | Pneumococci isolates |
| pncserotypes | Number fo pneumococcal serotypes detected |
| serotype1 | Pneumococcal serotype #1 |
| serotype2 | Pneumococcal serotype #2 |
| serotype3 | Pneumococcal serotype #3 |
| serotype4 | Pneumococcal serotype #4 |
| group | NP swab processing group |

**Note regarding serotypes:**

In the sweep work, 6A and 6C could not be differentiated so are recorded as 6A/C. For consistency, the WHO swab results are also reported in this way in the swab file.

Serotypes 15B and 15C spontaneously interconvert over time so recorded as 15B/C in the swab file.

**Pneumo-DEEP\_pnc\_isolate\_summary.csv**

Isolate summary for all pneumococci identified during WHO culture: one isolate per row

|  |  |
| --- | --- |
| **variable** | **definition** |
| codenum | Unique participant ID (shared between mother and infant) |
| specnum | Unique specimen ID |
| ceftriaxone | Ceftriaxone susceptibility (susceptible, intermediate, resistant)\* |
| chloramphenicol | Chloramphenicol susceptibility (susceptible, intermediate, resistant) |
| clindamycin | Clindamycin susceptibility (susceptible, intermediate, resistant) |
| co.trimoxazole | Co-trimoxazole susceptibility (susceptible, intermediate, resistant) |
| erythromycin | Erythromycin susceptibility (susceptible, intermediate, resistant) |
| penicillin | Penicillin susceptibility (susceptible, intermediate, resistant) |
| tetracycline | Tetracycline susceptibility (susceptible, intermediate, resistant) |
| penicillin\_mic | Penicillin MIC (mg/L) – if done |
| ceftriaxone\_mic | Ceftriaxone MIC (mg/L) – if done |
| serotype | Phenotypic serotype |
| sequence\_type | ST derived from PCR / capillary sequencing |
| lane\_id | Sequencing ID #1 |
| sanger\_id | Sequencing ID #2 |
| sequence\_type\_wgs | ST derived from WGS data |

**\*** Ceftriaonxe susceptibility has been inferred as “susceptible” where penicillin was susceptible, with the result derived from the MIC result in all other cases (as per CLSI guideline)

**Pneumo-DEEP\_illness\_abx\_summary.csv**

Clinic visits for infant illness plus associated antibiotic prescription data: one illness per row.

|  |  |
| --- | --- |
| **variable** | **definition** |
| illness\_id | Unique illness ID |
| codenum | Unique participant ID (shared between mother and infant). |
| date\_start | Antibiotic start date |
| drug1 | Antibiotic #1 |
| drug2 | Antibiotic #2 |
| drug3 | Antibiotic #3 |
| drug4 | Antibiotic #4 |
| clinic\_date | Date of first presentation to the clinic |
| diagnosis\_clinical | Clinical diagnosis category |
| cxr | Chest x-ray result (if done; reported by WHO criteria) |
| nps\_illness | NP swab collected during illness (Yes / No) |
| nps\_specdate | NP swab specimen date |
| nps\_specnum | NP swab specimen number |
| npa\_illness | NP aspirate collected during illness (Yes / No)\* |
| npa\_specdate | NP aspirate specimen date |
| npa\_specnum | NP aspirate specimen number |
| influenza\_a\_pcr | Influenza A PCR result (positive / negative) |
| influenza\_swa\_pcr | Influenza A (pdmH1N1) PCR result (positive / negative) |
| influenza\_a\_h\_type | Influenza A H1 / H3 PCR result (if done) |
| influenza\_b\_pcr | Influenza B PCR result (positive / negative) |
| rsv\_pcr | RSV PCR result (positive / negative) |
| hmpv\_pcr | Human metapneumovirus PCR result (positive / negative) |
| adv\_pcr | Adenovirus PCR result (positive / negative) |
| rnp\_cdc | Internal control [RNaseP] PCR result (positive / negative) |

**\*** NP aspirates were collected in children with clinical pneumonia and bronchiolitis

**Pneumo-DEEP\_carriage\_summary\_RESTRICTED.csv**

* Pneumo-DEEP\_carriage\_summary\_RESTRICTED.csv
  + this file only contains information on the infants included in the carriage duration GWAS (Lees et al., 2017)

**File organisation / relationships**

