

## Supplementary Appendix 1

List of obstetric procedures according to the Management System of Procedures, Medications, Orthoses, Prostheses and Special Materials of the Unified Health System Table. <http://sigtap.datasus.gov.br/tabela-unificada/app/sec/inicio.jsp>

#0201010011 Amniocentesis  
 #0211040010 Amnioscopy  
 #0211040061 Antepartum tococardiography  
 #0310010012 Birth assistance without dystocia  
 # 0310010020 Care for newborns at birth  
 #0310010039 Normal delivery  
 # 0310010047 Normal delivery in high-risk pregnancies  
 # 0310010055 Normal delivery at normal delivery centre  
 # 0303100010 Treatment of intercurrents predominantly related to the postpartum period  
 #0303100028 Eclampsia treatment  
 # 0303100036 Treatment of edema, proteinuria and hypertensive disorders in pregnancy, childbirth and the postpartum period.  
 # 0303100044 Treatment of clinical intercurrents during pregnancy  
 #0303100052 Hydatidiform mole treatment  
 #0409060011 Cervical cerclage  
 #0409060054 Uterine curettage in hydatidiform mole  
 # 0409060070 Post-abortion uterus emptying by manual intrauterine aspiration  
 #0411010018 Manual placental abruption  
 # 0411010026 Caesarean delivery in high-risk pregnancies  
 #0411010034 Caesarean delivery  
 #0411010042 Caesarean delivery with tubal ligation  
 #0411010050 Manual reduction of acute postpartum uterine inversion  
 #0411010069 Postpartum episiotomy resuture  
 # 0411010077 Suture of pelvic tract lacerations (in childbirth before admission)  
 #0411010085 Surgical treatment of acute postpartum uterine inversion  
 #0411020013 Post-abortion/puerperal curettage  
 #0411020021 Embryotomy  
 #0411020030 Puerperal hysterectomy  
 #0411020048 Surgical treatment of ectopic pregnancy  
 #0411020056 Treatment of other maternal disorders predominantly related to pregnancy  
 #0417010028 Obstetric analgesia for normal delivery  
 #0417010010 Obstetric anaesthesia for caesarean section  
 # 0417010036 Obstetric anaesthesia for caesarean section in high-risk pregnancy

## Supplementary Appendix 2

C++ deterministic matching routine to identify hospitalizations (AIH records) belonging to the same hospital episode in the AIH database from the SIH-SUS and Python routine to create a hospital episodes database.

Since the operation requires processing large amounts of data, sorting and iterating through it, we opted to work with a SQL database (SQLite, widely available free and open source database) as backend for the routine, which was developed as a command line (character mode) program in C++ for faster execution. The gcc version 11.4.0 toolchain was used. The source code and respective makefile are available at github: [https://github.com/kencamargo/cuidados\\_multiplos/tree/main/processaih](https://github.com/kencamargo/cuidados_multiplos/tree/main/processaih).

The processing was done under Linux, but since both the C++ and Python codes only used standard features of both languages, the programs can be run in any OS that supports a standard C++ toolchain and a Python 3 interpreter.

First, the CSV file is imported into a table, at the same time the necessary indexes are created, in order to speed up processing later. Only the fields that are going to be used are imported. Two additional fields, which will hold the episode code as reference

for the cluster (AIHREF) and a flag marking whether the event initiates or is part of a sequence or transference are added. The fields of the table are all variable length text fields, named as follows:

Field name	Description
CNES	Identifier of the health care facility
CEP	Postal code
MUNIC_RES	Municipality where the patient resides
NASC	Date of birth
DT_INTER	Date of admission
DT_SAIDA	Date of hospital discharge (AIH closure)
COBRANCA	Reason for hospital discharge
N_AIH	Individual identifier of the AIH form
AIHREF	Common identifier for all rows for a full hospital episode (added when imported)
FLAG	Flag marking whether row initiates or is part of a sequence or transfer
DIAG_PRINC	Main diagnosis on admission
PROC_REA	Principal Procedure on admission
US_TOT	Reimbursement value in US dollars
UTI_MES_TO	Days in an intensive care unit
alta	Reason for hospital episode discharge (consolidated)
criterio_primario	Flag marking whether episode constitutes an obstetric event

Second, the table is ordered by CNES, MUNIC\_RES, NASC and DT\_INTER. The table is traversed from beginning to end, and all the rows that have the same value for the first three fields, and have DT\_INTER equal to or at most one day after DT\_SAIDA of the previous rows get the AIHREF. In this step, we reconstitute the sequence of events happening in a single healthcare facility. The resulting table is exported as a CSV file for further processing if necessary.

In the next steps the program searches for all the rows that have COBRANCA showing a transferto another hospital, in order to add rows subsequent to the transfer to the episode group. The following steps are repeated until all the transfers have been processed:

The table is reordered by CEP, MUNIC\_RES, NASC and DT\_INTER. A second table is created with the following fields (all variable size text):

Field name	Description
ORIGAIHREF	Original common identifier
AIHREF	New common identifier
DT_INTER	Date of admission

Every row with a transfer discharge is compared with the next row in the sequence, and if the latter has the same value for the first three fields, has DT\_INTER equal to or at most one day after DT\_SAIDA of the previous row, and a different AIHREF from the row with the transfer discharge generates a row in the second table. For easier comprehension of the following steps, the row that originated the transfer will be referred to as the origin row and the following row in the sequence will be referred to as the destination row.

If the destination row was not previously flagged as being a sequence to a transfer (see below), the second table will have ORIGAIHREF set to the destination row AIHREF, and AIHREF set to the origin row AIHREF. If, however, the destination row is already part of a transference sequence, ORIGAIHREF in the second table will be set to the origin row AIHREF, and AIHREF in the second table will be set to AIHREF of the origin row. This is necessary because some odd transference sequences break already ordered permanence sequences. In both cases DT\_INTER will be taken from the destination row.

In the next step, the first (main) table is ordered by AIHREF and DT\_INTER, and the second table is ordered by DT\_INTER. The second table is then traversed, and for every row ORIGAIHREF and AIHREF are retrieved. All the rows in the first table that have the same AIHREF as the ORIGAIHREF of the second table have that changed to the AIHREF of the second table, thus

grouping all the rows connected by the transfer. Those rows are also flagged as being part of a transfer sequence, changing the corresponding FLAG field.

After successive passes over the last two steps exhaust all the unconnected transfer rows, this part of the algorithm ends. At the end of processing, the resulting table is exported to a comma-separated values (CSV) file, with an added field with the sequential number of the row in its respective cluster.

Further processing takes place with a Python version 3.11 scripts accessing the SQLite database and generating an additional CSV file.

First, the processed table resulting from the processing with the C++ routine is ordered by AIHREF and DT\_INTER. The first row initialises several variables in memory that will contain, respectively:

- AIHREF
- initial date (from DT\_INTER)
- last date (from DT\_SAIDA)
- COBRANCA
- total reimbursement (from US\_TOT)
- total days in ICU (from UTI\_MES\_TO)
- number of episodes
- total number of rows which fit the obstetric admission criterion (from criterio\_primario)

The table is traversed, and for each row:

If AIHREF is equal to the one stored in memory:

- if DT\_INTER is smaller than the initial date in memory, the latter is replaced.
- if DT\_SAIDA is greater than the initial date in memory, the latter is replaced.
- COBRANCA is stored in memory.
- US\_TOT is added to total reimbursement.
- UTI\_MES\_TO is added to total days in ICU.
- number of episodes is incremented.
- criterio\_primario is added to total rows fitting obstetric criterion.

On the other hand, if AIHREF is different, a record is added to the CSV file containing all the variables above, plus the calculated difference in days from initial date to last date, and all the variables are reinitialized.

