# Code example: HLHS

For HLHS, running the code in HLHS\_main\_file.R will result in the following.

### Step 1 HLHS inclusion and exclusion module: identify patients with evidence of HLHS

#### Input dataset

NCHDAdata (Initial NCHDA cohort in record-level)

#### Derived variables

HLHSpat (patient-level): marker of HLHS patients, 0: No /1: Yes

#### Output dataset

HLHSdata\_raw (HLHS cohort in record-level, raw data)

### Step 2 HLHS pathway module: assign each patient's procedure records to a category and diagnostic subgroup

#### Input dataset

HLHSdata\_raw from step 1

#### Derived variables

* CatProc (record-level): Category of each procedure

A generic label for CatProc was used for all CHDs:

*0: pre-pathway*

*1: palliative stage 1 (pathway)*

*2: single ventricle palliative stage 2 (pathway)*

*3: single ventricle palliative stage 3 (pathway)*

*4: reparative procedure (pathway, empty in HLHS)*

*5: other pathway (Hybrid procedure in HLHS)*

*6: heart transplantation*

*7: reintervention/off-pathway*

*8: excluded (non-contributory record)*

Of note, procedure records with CatProc 1,2,3,4, 5 and 6 were considered as pathway.

* patAge\* (patient-level): age at pathway. For example, patAge1 and patAge2 are age at palliate stage 1 and 2, respectively.
* ProcSeq (patient-level): procedure sequence (pathway and heart transplantation). For example “1236” means patient had palliative stage 1, 2, 3 and heart transplantation.
* diagsubgroup (patient-level): diagnosis subgroup. HLHS has only 1 group as HLHS.

#### Output dataset

HLHSdata\_raw (derived variables added, record-level data)

### Step 3 HLHS further processing module: identify patients who met the specific violation rule and who met the flag to center rule.

#### Input dataset

HLHSdata\_raw from step 2

#### Sub module:

* Generic violation module
* HLHS specific violation module
* Generic suspected missing/miscoded data module
* HLHS specific suspected missing/miscoded data module
* Generic minor data error data module
* HLHS specific minor data error data module

#### Derived variables

* pat\_violation/ pat\_violationinf (patient-level): marker of patients who met the violation rule (0: No /1: Yes) and the reason. They will be removed from the study cohort.
* Flagcenter\_missingdata/ Flagcenter\_missingdatainf (patient-level): marker of patients suspicious missing or unusual procedure sequences (0: No /1: Yes) and the reason. They will remain in the study cohort. In future routine monitoring all such patients will be flagged with the treating centres for correction.
* Flagcenter\_minor/ Flagcenter\_ minorinf (patient-level): marker of patients with minor data errors (0: No /1: Yes) and the reason. They will remain in the study cohort. In future routine monitoring all such patients will be flagged with the treating centres for correction.

#### Output dataset

HLHSdata\_final (derived variables added, and patients who met the violation rule has been removed. record-level data)

### Step 4 Descriptive analysis module (generic): assign pathway management route, comorbidity and compute number of interventions, pathways and reinterventions for each patient.

#### Input dataset

HLHSdata\_final from step 3

#### Derived variables

* patpremat (patient-level): marker of pre-term birth (0: No /1: Yes)
* patCongComorb (patient-level): marker of congenital comorbidity (0: No /1: Yes)
* patDowns (patient-level): marker of Down syndrome (0: No /1: Yes)
* pathas\*(patient-level): marker of patients who had certain pathway procedure in records (0: No /1: Yes)
  + pathasStage1: patients had palliative stage 1
  + pathasSVstage2: patients had palliative stage 2
  + pathasSVstage3: patients had palliative stage 3
  + pathasBVpathway: patients had reparative procedure (none in HLHS)
* num\* (patient-level): number of interventions for each patient
  + numPrePath: number of pre-pathway
  + numallReint: number of reintervention (any type)
  + numHSReint: number of reintervention (surgery or hybrid)
  + numCEreint: number of reintervention (catheter or electrophysiology )
  + numIntby1: number of all cardiac interventions (all procedures except for those marked as excluded in var. CatProc) by age 1-year
  + numPathwayby1: number of pathway (procedure marked as 1,2,3, 4, and 5 in var CatProc) by age 1-year
* Status\_\*(patient-level): reintervention status
  + Status\_allReint: status of reintervention (any type)

Labels:

0: censored without reintervention

1: occurrence of reintervention

2: Death or heart transplantation without reintervention

* + Status\_HSReint: status of reintervention (surgery or hybrid, labels see above)
  + Status\_HSReint: status of reintervention (catheter or electrophysiology, labels see above)
* pathwaytype (patient-level): whether the patient was managed under a single ventricle pathway or a biventricular reparative pathway. Of note, HLHS is exclusively single ventricle.

#### Output dataset

HLHSdata\_final (derived variables added, record-level data)

### Step 5 Save patient level data module (generic): create a patient-level dataset that includes all the patient-level variables derived from previous steps.

#### Input dataset

HLHSdata\_final from step 4

#### Output dataset

HLHSpatientleveldata (patient-level data)

### Step 6 compute metrics module (generic): compute survival and reintervention metrics.

#### Input dataset

HLHSpatientleveldata from step 5

#### Output aggregated results

* survival metrics: survival rates at 1-year, 5-year and 10-years (using Kaplan-Meier estimator)
* reintervention metrics: cumulative incidence rates of reintervention rates at 1-year, 5-year and 10-years (using competing risk analysis, accounting for death without reintervention as competing events).