

# **HICDEP Standard Specification**

**Version: 1.80**

**Date: 2014-10-2**

## Table of Contents

<b>About HICDEP</b>	<b>5</b>
<b>Change log</b>	<b>6</b>
Version 1.80	6
Version 1.70	7
Version 1.60	8
Version 1.50	8
Version 1.30	9
Version 1.25	9
Version 1.21	9
Version 1.2	10
Version 1.1	10
Version 1.00	10
Version 0.50 and 0.90	10
Version 0.38	10
<b>HICDEP 1.80</b>	<b>11</b>
General data format	11
Overview of data tables	11
Diagram	12
Structure of data	13
From flat files towards a normalized structure	13
Technical considerations	13
Coding Conventions	14
Date codes	14
Day unknown	14
Month and day unknown	14
Year unknown	14
Specification of precision	14
?ICD-10 codes	15
?ATC codes	15
Other codes	16
Quality Assurance	16
<b>tbIART - Antiretroviral treatment</b>	<b>17</b>
Core fields	17
Additional fields	17
QA Checks	17
<b>tbIBAS - Basic clinical, background and demographic information</b>	<b>19</b>
Core fields	19
Additional fields	19
QA Checks	20
<b>tbICANC - Diagnosis of Cancer</b>	<b>22</b>
Core fields	22
Additional fields	22
QA Checks	22
<b>tbICEP - Clinical Events and Procedures</b>	<b>23</b>
Core fields	23
Additional fields	23
QA Checks	24
<b>tbIDELIVERY_CHILD - Delivery information related to the child(ren)</b>	<b>25</b>
Core fields	25
QA Checks	25
<b>tbIDELIVERY_MUM - Delivery information related to the mother</b>	<b>27</b>
Core fields	27
QA Checks	27
<b>tbIDIS - Opportunistic infections</b>	<b>29</b>
Core fields	29
Additional fields	29
QA Checks	29
<b>tbILAB - Laboratory values</b>	<b>31</b>
Core fields	31

Additional fields	31
QA Checks	31
<b>tbILAB_BP - Laboratory values - Blood pressure</b>	<b>33</b>
Core fields	33
QA Checks	33
<b>tbILAB_CD4 - Laboratory values</b>	<b>34</b>
Core fields	34
Additional fields	34
QA Checks	34
<b>tbILAB_RES - Resistance testing</b>	<b>36</b>
Core fields	36
Additional fields	36
QA Checks	37
<b>tbILAB_RES_LVL_1 - Nucleotide sequences (PRO, RT, GP41, GP120)</b>	<b>39</b>
Core fields	39
Additional fields	39
QA Checks	39
<b>tbILAB_RES_LVL_2 - Mutations</b>	<b>41</b>
Core fields	41
QA Checks	41
<b>tbILAB_RES_LVL_3 - Resistance test result</b>	<b>43</b>
Core fields	43
Additional fields	43
QA Checks	43
<b>tbILAB_RNA - Laboratory values</b>	<b>45</b>
Core fields	45
Additional fields	45
QA Checks	45
<b>tbILAB_VIRO - Laboratory values - viro-/serology</b>	<b>47</b>
Core fields	47
Additional fields	47
QA Checks	47
<b>tbILTFU - Death and drop-out</b>	<b>49</b>
Core fields	49
Additional fields	50
QA Checks	50
<b>tbIMED - Other medication</b>	<b>52</b>
Core fields	52
Additional fields	52
QA Checks	52
<b>tbINEWBORN - Newborn</b>	<b>54</b>
Core fields	54
QA Checks	54
<b>tbINEWBORN_ABNORM - Abnormalities</b>	<b>56</b>
Core fields	56
QA Checks	56
<b>tbIOVERLAP - Cross-cohort identification</b>	<b>57</b>
Core fields	57
QA Checks	57
<b>tbIPREG - Pregnancy</b>	<b>58</b>
Core fields	58
QA Checks	59
<b>tbIPREG_OBS - Obstetrical problems during pregnancy</b>	<b>61</b>
Core fields	61
QA Checks	61
<b>tbIPREG_OUT - Pregnancy outcome</b>	<b>63</b>
Core fields	63
QA Checks	63
<b>tbIREFILL - Prescription refill data</b>	<b>65</b>
Core fields	65
QA Checks	65
<b>tbISAMPLES - Blood Samples</b>	<b>66</b>

Core fields	66
Additional fields	66
QA Checks	66
<b>tblVIS - Basic follow-up/visit related data</b>	<b>68</b>
Core fields	68
Additional fields	68
QA Checks	69
<b>Considerations for data management</b>	<b>71</b>
Uniqueness of PATIENT ID	71
Duplicate records	71
<b>Considerations for using the format to create a database</b>	<b>72</b>
Administrative fields	72
Further normalisation	73
Lookup tables	73
Performance	73
<b>QA checks</b>	<b>74</b>
Error Codes	74
QA checks listed nowhere else	74
All QA checks	75

## About HICDEP

HIV cohort collaborations have made substantial contributions to the knowledge of HIV epidemiology and management over the last years. So far, most collaborations have incorporated slightly different protocols for data exchange causing unnecessary workload for the people in charge of data extraction.

We were therefore asked to put together this draft consensus protocol for discussion at the 7th International Workshop on HIV Observational Databases, March 29th-30th 2003, Fiuggi, Italy. It is based on our experience with data-exchange protocols for [D:A:D](#), the [ART Cohort-Collaboration](#), the PLATO Collaboration and several previous studies on the safety of stopping OI prophylaxis.

This protocol is based on a relational structure (with some very minor deviations) and currently incorporates 26 data tables and numerous lookup-tables for the codes. It is evident that - depending on the questions addressed - only subsets of tables and fields will have to be extracted for data exchange.

We have not elaborated on database systems (e.g. SQL-Server, Oracle, Access) and their respective file formats as there are excellent tools for transferring data between most of the popular packages (e.g. [StatTransfer](#)). The suggested data structure should work with most formats and software packages.

Please keep in mind that the primary purpose of this document is to provide you with formats for data-exchange but not for an operational database used for data-management on a day-to-day basis. Some considerations with that respect can be found in the appendices.

We plan to update this document on a regular basis and the most recent versions will be made available on the [HICDEP website](#). HICDEP is a format under constant improvement (currently under the EuroCoord Project, Grant Agreement No. 260694) and additions are made almost every year. Please refer to the [ChangeLog](#) for the most current updates and always use the tables available on the HICDEP website for most current coding lists for ART and MED drugs.

Members of EuroCoord WP4, September 2011

## Change log

### Version 1.80

#### [tbIART](#)

##### [ART\\_ID](#)

- Added:
  - J05AR11 (Lamivudine, tenofovir disoproxil and efavirenz)
  - J05AR12 (Lamivudine and tenofovir disoproxil)
  - J05AR13 (Lamivudine, abacavir and dolutegravir)
  - J05AR14 (Darunavir and cobicistat)
  - J05AX-CAB (Cabotegravir (GSK-744))

#### [tbIBAS](#)

##### [ETHNIC](#)

- Revised and added codes for different ethnicities
- Changed limitations
- Changed description

##### [ORIGIN](#)

- Changed description to be more specific

#### [tbICANC](#)

- Added table

#### [tbILTFU](#)

##### [DEATH\\_R1](#)

- Added sub-types for 22 and 23
- Added additional codes table

#### [tbIMED](#)

##### [MED\\_ID](#)

- Added:
  - J05AE11 (Telaprevir (INCIVEK, INCIVO))
  - J05AE13 (Faldaprevir)
  - J05AE14 (Simeprevir)
  - J05AE15 (Asunaprevir)
  - J05AX14 (Daclatasvir)
  - J05AX15 (Sofosbuvir)
  - J05AR-A450OM (ABT-450/r/Ombitasvir)
  - J05AR-DAAS (Daclatasvir/Asunaprevir)
  - J05AR-LESO (Ledipasvir/Sofosbuvir)
  - J05AX-DBV (Dasabuvir)

#### [tbIVIS](#)

##### [LAB\\_U](#)

- Added:
  - 14 (mg/24h)
  - 15 (mg/mmol)
  - 16 (fl)
  - 17 ( $\mu\text{g/mL} = \text{mg/L}$ )
  - 18 ( $\mu\text{kat/L}$ )

##### [LAB\\_ID](#)

- Added:
  - ACRA (Albumin Creatinin Ratio)
  - PCRA (Protein Creatinin Ratio)
  - PROT (Protein)
  - PSA (Prostate-specific antigen)

- PTH (Parathyroid Hormone)

## Version 1.70

- [tblAE](#) has been renamed to [tblCEP](#)

### [tblCEP](#)

- Added field [CEP\\_V](#)

#### [CEP\\_ID](#)

- Added: HOSP, ICU, JAUN, LIVD, LIVT, USAB, ASP, BART, CHAG, NOCA, PCE, PMAR, REQU.
- Added: BACT, ENDO, MENI, OSTI, PERI, PNEU, PYEL, LEIS, MCDI (Previously in [tblDIS](#)).
- Removed: ANG, BYP, END as they are already in the field CEP\_SPEC.
- Removed COR as now specified with SUD\_DEATH\_Y and EXP\_DEATH\_Y in [tblLTFU](#).

#### [CEP\\_SPEC](#)

- Changed description of NADM - CERV to contain "grade 2 or higher"
- Added values: LIVB - XX, HEP - XX, NADM - HENE, NADM - BRAIN

### [tblBAS](#)

- Added field [RECart\\_D](#) (Date ART started)
- Added field [LTART\\_D](#) (Date last assessed for ART)
- [ETHNIC](#): Added Limitations section in description
- Added field [EDU\\_LVL](#) (Last completed education Level)
- Added field [HIV\\_POS\\_D](#) (Date of first positive HIV test)

### [tblDIS](#)

- [DIS\\_ID](#): Removed HG, BACT, ENDO, MENI, OSTI, PERI, PNEU, PYEL, LEIS, MCDI (now in [tblCEP](#)).

### [tblVIS](#)

- Added optional field [EMPLOY](#) (What is the patient's current situation regarding labour?)
- Added optional field [CONTRACT](#) (If the patient is an employee, what is the type of the patient's employment contract?)
- Added optional field [SMOKING\\_Y](#) (Is the patient currently a smoker?)
- Added optional field [PREG\\_Y](#) (Is the patient currently pregnant?)
- Added optional field [CDC\\_STAGE](#)
- Added optional field [WHO\\_STAGE](#)

### [tblLTFU](#)

- Added optional fields for ICD9 cause of death coding (ICD9\_#)
- Added optional field DEATH\_SOURCE
- Added optional fields SUD\_DEATH\_Y and EXP\_DEATH\_Y. Replacing COR

#### [CoDe](#)

- Typos in the coding table (HBF -> HBV)
- New codes for cancer (04.XX)
- New codes for AMI (08.XX)
- New code: 92.1 - Unknown, competing risks

### [tblART](#)

[ART\\_ID](#): Coding extended/updated to match current ATC codes.

- Lopinavir / Ritonavir changed code from "J05AE06" to "J05AR10".
- Elvitegravir changed code from "J05AX-EVG" to "J05AX11"
- J05AR07 - Triomune
- J05AR08 - Eviplera / Complera
- J05AR09 - Stribild
- J05AR10 - Kaletra / Aluvia
- J05AX12 - Dolutegravir
- V03AX03 - Cobicistat
- [ART\\_FR](#): Added 0.33 and 0.5. Added special value -1 for "frequency unknown".

#### [ART\\_RS](#)

- Added new codes: 6.X, 92.31, 92.32, 92.33, 92.4, 92.9, 94.2, 96.1, 96.2, 97.1, 97.2, 97.6

### [tblLAB](#)

[LAB\\_ID](#)

- Removed redundant "APT" code. Use code "ALP"
- Removed "GLYCE" (synonymous to Glucose in blood)
- Added "AFP" (Alfa Fetoprotein) and "DIPP" (Dipstick result for protein in Urine).

[LAB\\_U](#)

- Removed WBC codes (WBC was already removed from LAB\_ID)
- Added generic "13 - µg/L" and "99 - No units" code

[LAB\\_ST](#)

- Added codes U and U24.

[tblLAB\\_RES](#)

- [SUBTYPE](#) Modified description to also include HCV

[tblLAB\\_RNA](#)[RNA\\_T](#)

- New code: 42 - Abbott RealTime HIV-1 m2000
- New code: 59 - Monitor unspecified

[tblMED](#)[MED\\_ID](#)

- Changed description from "Other HIV-related drugs" to "Other medication"
- Added codes for Boceprevir and Telaprevir

[MED\\_RS](#)

- Same changes as in tblART [ART\\_RS](#)

[tblLAB\\_VIRO](#)

- [VS\\_T](#) New codes 6,7,8.

**Version 1.60**[tblBAS:](#)

- The [ORIGIN](#) field is now coded using UN region and country codes.
- Added optional [SEROHOW](#) field indicating how the seroconversion date was determined.
- Added optional [CENS\\_D](#) field holding the last date the database was updated for a patient.
- [tblIREFILL](#) created to hold prescription refill data.
- [tblLTFU](#): The [DEATH\\_R#](#) fields now use [CoDe](#) codes for coding cause of death.

[tblLAB:](#)

- Added codes for Gamma-glutamyltransferase and Prothrombin rate.

[tblLAB\\_VIRO:](#)

- Changed code for "Other" of field [VS\\_T](#) to value 9.
- Added "HBVACIGM", "HBVACIGG", "HCVBD" and "HDVA" codes for [VS\\_ID](#) field.

[tblLAB\\_RES](#), [tblLAB\\_RES\\_LVL\\_1](#), [tblLAB\\_RES\\_LVL\\_2](#) and [tblLAB\\_RES\\_LVL\\_3:](#)

- Renamed SAMP\_ID to TEST\_ID, as it gives a better description of what the values should encode.

[tblLAB\\_RES:](#)

- Added a [VIRUSTYPE](#) variable allowing to distinguish HIV- and HCV-resistance tests.
- Added recommended [HIV-1 subtype codes](#) and [HCV geno- and subtype codes](#).

[tblLAB\\_RES\\_LVL\\_3:](#)

- ATC\_CODE replaces ART\_ID as HCV medication in tblMED may be encoded.

[tblDIS:](#)

- Added optional [DIS\\_ED](#) field for the end date of the disease.

[tblMED:](#)

- Added [MED\\_RS](#) describing the reason for stopping the treatment; needed for HCV.

[tblAE:](#)

- Added a number of codes for [AE\\_ID](#): ANG, ASCI, AVN, BYP, CERC, END, FIBS, FRA, HEP, HESY, LAC, LIVB, OESO, PAN and PERI.
- QA checks have been extended and updated where necessary.

**Version 1.50**



- Added the pediatric tables: [tbIPREG](#), [tbIPREG\\_OBS](#), [tbIPREG\\_OUT](#), [tbIDELIVERY\\_MUM](#), [tbIDELIVERY\\_CHILD](#), [tbNEWBORN](#) and [tbNEWBORN\\_ABNORM](#).
- Added new table [tbISAMPLES](#) which holds information regarding sample storage.
- [tbIART](#): Updated list of drugs

### Version 1.30

- [tbIART](#): Updated list of drugs
- [tbIAE](#):
  - added [EVENT\\_ID](#) as unique identifier and link to detailed tables for each event (see <sup>1</sup>) ? this replaces the optional [AE\\_NO](#) field.
  - added [AE\\_SPEC](#) to further specify an event by coding
  - a series of basic verification fields have been added to allow for tracking of event status for source documentation availability, verification of documentation (through monitoring) and final approval of the event.
  - in [AE\\_R\\_Y](#) - Relation to treatment: added more detailed codes.
- [tbILAB](#): added several codes for various biomarker tests.
- [tbILAB\\_CD4](#): added [CD4\\_U](#) as optional fields to discriminate between CD4% and CD4 cell count, so that the [tbILAB\\_CD4](#) table can hold both types of measurements.
- [tbILAB\\_VIRO](#): added several codes for various virology and serology tests.
- [CaseDefinitions](#) updated with end stage renal disease, chronic liver disease and non-AIDS defining malignancies

<sup>1</sup>:Detailed table definitions for the D:A:D events are available at [?http://www.cphiv.dk/HICDEP/Documents/tabid/159/Default.aspx](http://www.cphiv.dk/HICDEP/Documents/tabid/159/Default.aspx)

### Version 1.25

- [tbIART](#): Updated list of drugs
- [tbIMED](#): Updated list of drugs
- [tbIDIS](#):
  - Changed wording for CANO to ?Candidiasis, oesophageal, bronchi, trachea, or lungs?
  - Added COCC - Coccidioidomycosis, disseminated or extrapulmonary
- [tbILAB](#): Added [LAB\\_ST](#) as additional field to code for type of specimen used for the measurement
- [tbILAB\\_CD4](#): Added [CD4\\_U](#) as additional field so the table can hold both percentage and absolute CD4 measurements
- [tbILAB\\_RNA](#): Added [RNA\\_UL](#) (upper limit of detection) to the list of additional fields.
- Added more viral assays to the list of [RNA\\_T\\_codes](#)
- [tbILAB\\_VIRO](#): Added unit field to [tbILAB\\_VIRO](#) into the general format and [VS\\_LL](#) (lower limit of detection), [VS\\_UL](#) (upper limit of detection) and [VS\\_T](#) (type of test) and list of tests to the list of additional fields.
- [tbILTFU](#): Added [DEATH\\_RC#](#) to code for causal relation of the [DEATH\\_R#](#) code to the death in order to comply with [CoDe](#) and still maintain a format to be used for cohorts not using CoDe. [ICD10\\_#](#) fields have been moved to the list of additional fields.

### Version 1.21

- Added reasons for stopping treatment to table [tbIART\\_CODE\\_RS](#):

Code	Coding for Reason of Stopping Treatment
1.1	Virological failure
1.2	Partial virological failure
1.3	Immunological failure ? CD4 drop
1.4	Clinical progression
90	Side effects ? any of the above but unspecified
90.1	Comorbidity
92.1	Simplified treatment available
92.2	Treatment to complex
92.3	Drug interaction

93.1	Structured Treatment Interruption (STI) ? at high CD4
94.1	Non-compliance
96	Pregnancy
97	Study treatment

## Version 1.2

- added [CaseDefinitions](#)

## Version 1.1

### tbIBAS:

- The table was split into tbIBAS and tbILTFU. tbILTFU holds data on death and drop-out
- Renamed *LOS\_Y* to *LOSS\_Y*
- Renamed *GAL\_Y* to *GAIN\_Y*

### tbILAB\_BLP:

- Renamed table to *tbILAB\_BP*
- Renamed *BLP\_D* to *BP\_D*
- Renamed *BLP\_SYS* to *BP\_SYS*
- Renamed *BLP\_DIA* to *BP\_DIA*
- Renamed *BLP\_U* to *BP\_U*

## Version 1.00

### tbIBas:

- Renamed *BIRTHDAY* to *BIRTH\_D*
- Renamed *FIRSTVIS* to *FRSVIS\_D*
- Renamed *REC\_ART* to *RE CART\_Y*

### tbILAB:

- *LAB\_U*: has been dropped ? please use the ?unit codes/strings? as that is a safer way to code/represent the units ? prefixing all ?unit codes/strings? with a numeric value should however make analysis easier.

### tbILAB\_VIRO:

- New table added to capture mainly hepatitis measurements/tests

### tbILAB\_RES:

- *SEQ\_DT* was added to capture the time of sequencing in order to facilitate quality assurance of the data for contamination that might have happened during the sequencing.

### tbILAB\_RES\_LVL1:

- Renamed *SEQ\_ST* to *SEQ\_STAR*
- *SEQ\_STOP*: Added to the table to specify at which position in the sequence the sequencing was terminated

### tbILAB\_RES\_LVL2:

- The table has been optimised for ease of analysis so that the mutation codes have been split into their components of amino acid position, sub position for insertions and 4 our more fields for mixtures of amino acids found in the sample

## Version 0.50 and 0.90

First public versions that incorporated comments and corrections received from attendees at the 7th International Workshop on HIV Observational Databases, March 29th-30th 2003, Fiuggi, Italy and Stephen Hart.

## Version 0.38

Version presented at 7th International Workshop on HIV Observational Databases, March 29th-30th 2003, Fiuggi, Italy

## HICDEP 1.80

This article describes the HICDEP version 1.80 which was released on October 2nd 2014. For a more detailed version history, please refer to the [ChangeLog](#).

### General data format

The table pages referenced [in the overview](#) describe the specific tables' structure in detail and present a list of suggested codes, both standard and human readable.

All codes apart from trivial no, yes or unknown codes are presented as lookup tables, the usage of these are described in the the article [Considerations for using the format to create a database](#).

Along with the basic structure described in each ?Core fields? section, additional fields containing additional or more specific data are described in the ?Additional fields? sections. These fields were taken from several cohort collaborations but with the required changes that were needed for the specific data structures. This is presented to the reader to show that the core structure is not a fixed proposal but rather a basic structure, which can be altered by adding fields.

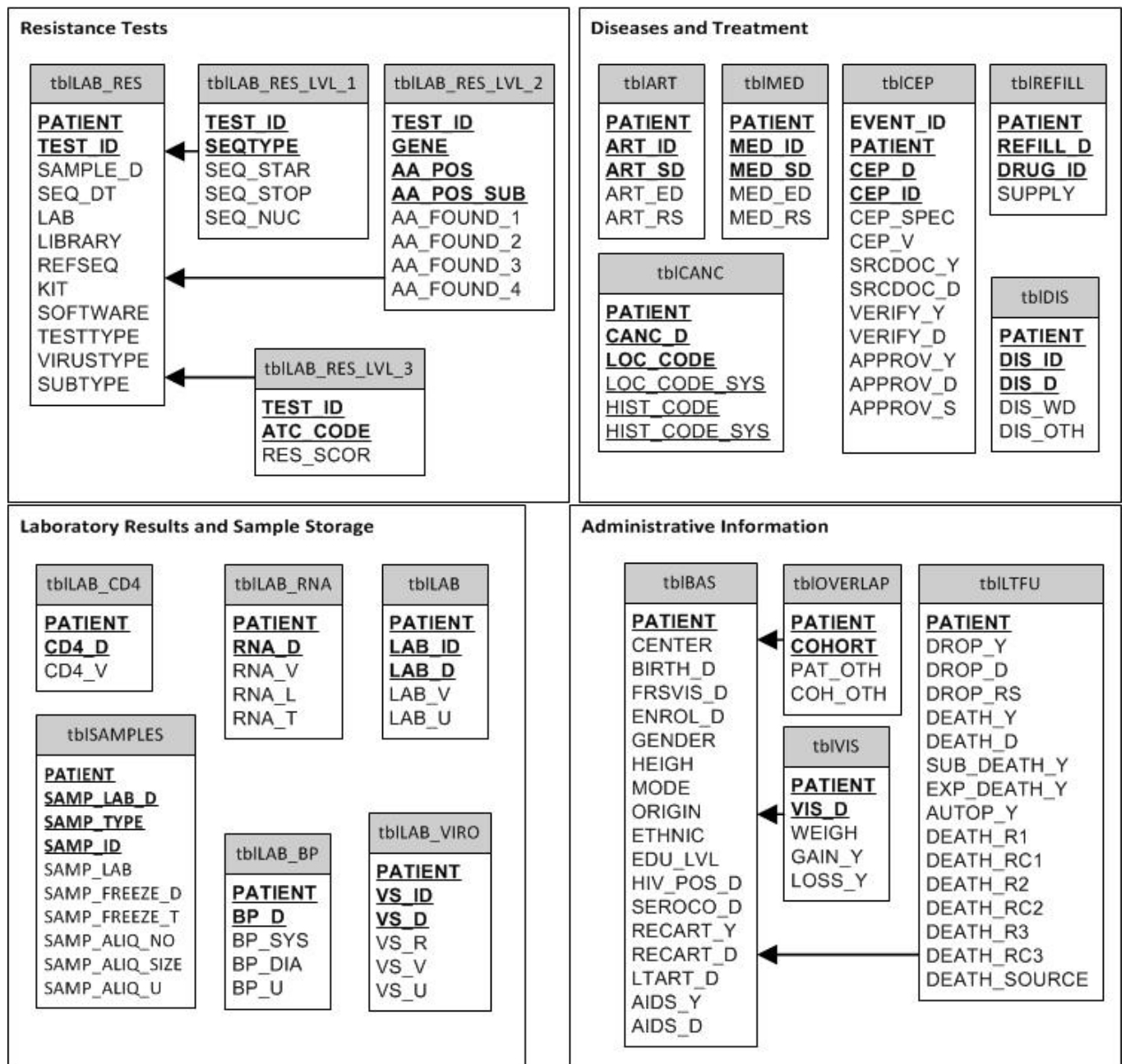
Issues regarding duplicates are discussed in [Considerations For Data Management](#).

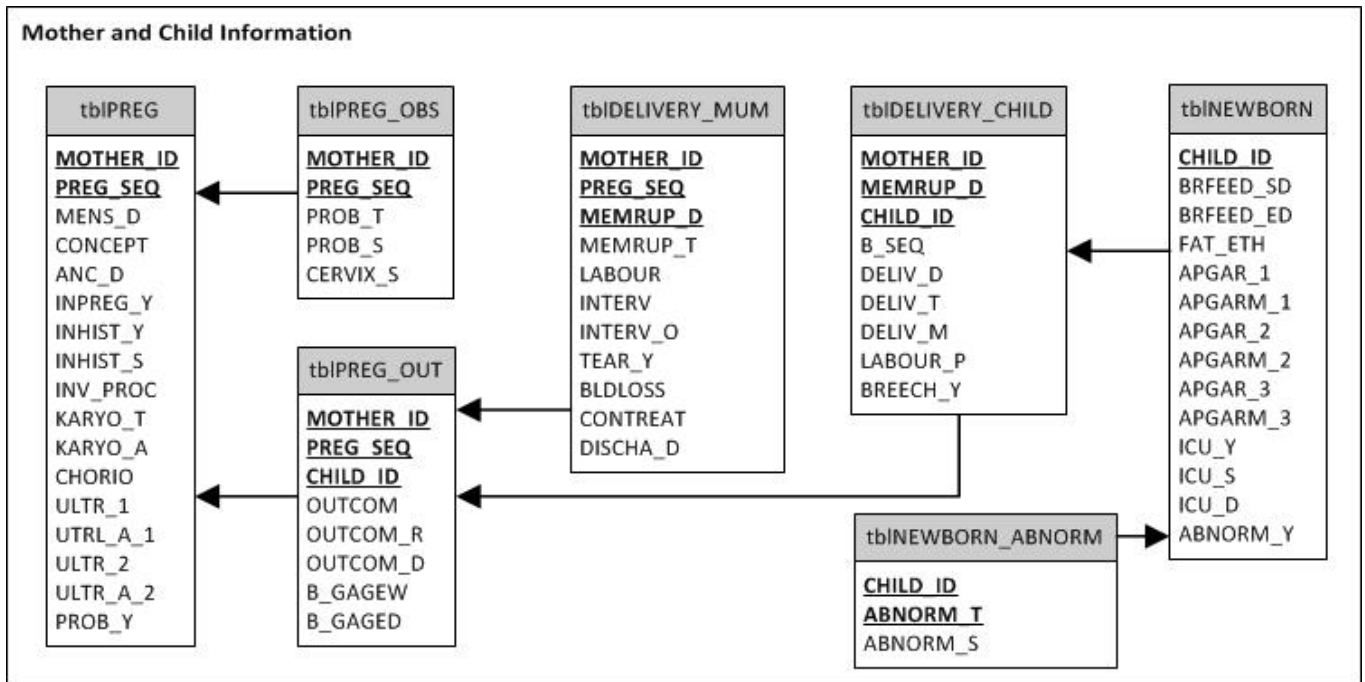
### Overview of data tables

Table	Content
<a href="#">tblART</a>	holds type of <b>antiretroviral drug</b> , start and stop dates and reason for stopping
<a href="#">tblBAS</a>	holds <b>basic</b> information such as demographics, basic clinical information and date of AIDS diagnosis
<a href="#">tblCANC</a>	holds type and date of <b>diagnosis of cancer</b>
<a href="#">tblCEP</a>	holds type and date of <b>clinical events and procedures</b> including serious non-AIDS conditions. Former known as tblAE (adverse event).
<a href="#">tblDELIVERY_CHILD</a>	holds <b>delivery</b> information related to the child
<a href="#">tblDELIVERY_MUM</a>	holds <b>delivery</b> information related to the mother
<a href="#">tblDIS</a>	holds type and date of CDC-C <b>diseases</b> .
<a href="#">tblLAB</a>	holds type, date, value and unit of <b>laboratory tests</b> .
<a href="#">tblLAB_BP</a>	holds date, diastolic and systolic values and unit of <b>blood pressure</b> measurements.
<a href="#">tblLAB_CD4</a>	holds date and value of <b>CD4 measurements</b> .
<a href="#">tblLAB_RES</a>	holds <b>background information</b> on the resistance test, <b>laboratory</b> , library, kit, software and type of test
<a href="#">tblLAB_RES_LVL_1</a>	holds nucleoside sequence for the PRO and RT sequences
<a href="#">tblLAB_RES_LVL_2</a>	holds mutations and positions of these.
<a href="#">tblLAB_RES_LVL_3</a>	holds resistance result in relation to antiretroviral drug.
<a href="#">tblLAB_RNA</a>	holds date, value, detection limit and type of <b>viral assay</b> .
<a href="#">tblLAB_VIRO</a>	holds test results for <b>viro-/serological tests</b> (hepatitis etc.)
<a href="#">tblLTFU</a>	holds data on <b>death and drop-out</b>
<a href="#">tblMED</a>	holds type, start and stop dates for <b>other medication/treatments</b> .
<a href="#">tblNEWBORN</a>	holds information related to <b>newborns</b>
<a href="#">tblNEWBORN_ABNORM</a>	holds information related to <b>abnormalities</b> of newborns
<a href="#">tblOVERLAP</a>	holds information on the patient's <b>participation in other cohorts</b>

<a href="#">tbIPREG</a>	holds general <b>pregnancy</b> -related information
<a href="#">tbIPREG_OBS</a>	holds information on <b>obstetrical problems during pregnancy</b>
<a href="#">tbIPREG_OUT</a>	describes the <b>pregnancy outcome</b>
<a href="#">tbIREFILL</a>	holds information on <b>prescription refills</b>
<a href="#">tbISAMPLES</a>	holds information on the storage of blood, urine and other <b>biological samples</b>
<a href="#">tbIVIS</a>	holds <b>visit related information</b> such as weight, wasting, smoking, occupational status etc.

**Diagram**





**Structure of data**

**From flat files towards a normalized structure**

The data collected in HIV collaborations is presented on the following pages in a set of data files/tables. Typically data would be put into one data file that would hold one line/record per patient where each field is represented as a separate column in that dataset. Often a dataset could contain more than 3000 columns of data.

The implication of going from thousands of fields to fewer fields means that data is in fact transposed from the flat format into the normalised format.

**Example of a flat file structure:**

PATIENT	ALAT_D	ALAT_V	ALAT_U	ASAT_D	ASAT_V	ASAT_U
999999	01-01-2000	15	U/l	01-01-2000	12	U/l

The normalised structure would then be like this:

PATIENT	TYPE_ID	LAB_DATE	LAB_VAL	LAB_UNIT
999999	1	01-01-2000	15	U/l
999999	2	01-01-2000	12	U/l

The type of measurement is identified through the TYPE\_ID field. Here 1 codes for ALAT and 2 codes for ASAT:

Code	Description
1	ALAT - Alanin-Aminotransferase
2	ASAT - Aspartat aminotransferase

**Technical considerations**

To enable a normalised structure that minimises the number of columns dramatically, the one file solution must be broken into several minor tables. These breakdowns are driven by the different data characteristics.

Each table has a basic structure that includes the patient identifier, a code that represents e.g. drug, adverse event or laboratory test performed. Along with this combination values like date, result, unit etc are present for each record.

A record for a laboratory measurement would include:

- Patient identifier
- Measurement type identifier
- Measured value
- Unit of value
- Date of determination

A record for usage of an antiretroviral drug would include:

- Patient identifier
- Drug identifier
- Start date for usage
- End date for usage
- Reason for discontinuation

These issues imply that a set of distinct tables must be generated based on the nature of the data. Since laboratory, medication and event data both cannot and should not be mixed at least 3 tables must be designed. Additionally there are other types of information that need their own domains: background information on the patient (height, birth date etc.), visit related data (weight, blood pressure, wasting etc.), and resistance testing (the latter requires more consideration due to the diversity of data present).

In this protocol further separation of data into different tables are presented. These separations are not only based on the rules for the relational model and normalisation, but they are culturally related.

For example: antiretroviral treatment medication is kept in one table and other medication in another table; CD4 cell measurements and HIV-RNA measurements are put into separate tables, that are also different from the general laboratory table. These separations are done simply because data in these tables are of distinct importance in analysis and often are gathered more frequently and with more attention than other variables.

## Coding Conventions

### Date codes

Although it is best to have precise dates in the format of YEAR-MONTH-DAY [ISO standard](#), it might be that some cohorts are limited to representing date data at the level of the month only, or information kept on the patient in the charts only defines dates to the month and in some cases only to the year. To solve this a set of date codes are presented here.

### Day unknown

In this case the date should be coded as the 15th of the month so that 1999-12-?? becomes 1999-12-15. This enables the date to be no more than 15 days away from the actual date.

### Month and day unknown

Best approach to this is to apply something similar, as with unknown dates, this would then mean that 1999-??-?? becomes 1999-07-01.

### Year unknown

If the year is unknown but the presence of the date value is needed as in case of opportunistic infections or adverse events (see later in this document) a fictive date should be used that couldn't be mistaken with an actual date. An unknown year should be coded as 1911-11-11.

### Specification of precision

An alternative to the above is to apply an additional field to each date field for which it is known that there might be issues regarding the precision of the dates. The field is then used to specify at which degree of the day, month or year the date is precise:

Code	Precision of date
<	Before this date
D	Exact to the date
M	Exact to the month
Y	Exact to the year

>	After this date
U	Unknown

### [?ICD-10 codes](#)

The coding system is the official standard for coding of diseases, however there is a wide set of ?homebrew? codes used within the HIV field in data coding in general, often it?s a 3 or 4 letter code which is an abbreviation for the AIDS defining disease. ICD-10 doesn?t have single codes that represent all single CDC-C events and as a consequence of this a list of 3 to 4 letter codes is the recommended way of coding for all CDC stage C events

ICD-10 codes are however the recommended for codes AE?s since it would become impossible for this protocol to maintain a complete list of all possible AE?s. ICD-10 is also recommended for causes of death.

### [?ATC codes](#)

ATC is a hierarchical structure for coding medication. The structure and hierarchy are best explained with an example of how a drug code is defined. Here it is on Indinavir:

J  
 ANTIINFECTIVES FOR SYSTEMIC USE (1st level, anatomical main group)  
 J05  
 ANTIVIRALS FOR SYSTEMIC USE (2nd level, therapeutic subgroup)  
 J05A  
 DIRECT ACTING ANTIVIRALS (3rd level, pharmacological subgroup)  
 J05AE  
 Protease inhibitors (4th level, chemical subgroup)  
 J05AE02  
 Indinavir (5th level, chemical substance)

This hierarchy has some benefits as will be explained later, but one of its limitations is that it?s impossible to ?read? the code compared to the widely used 3 letter mnemonic codes for antiretroviral drugs.

Example:

Drug	Code	ATC code
Indinavir	IDV	J05AE02

The difference is that the IDV code is easily readable, where the ATC code is not; going from a flat file structure to a normalised structure the human readable aspect becomes increasingly important. In the flat file format the column names and the possibility of labels makes data more or less readable; in the normalised format only the coding can help. Because of this the 3 letter codes are being presented in this document. However it must be stressed that usage of the ATC coding should be used to diminish the risk of several homebrew and non-compatible coding schemes.

**Currently however, the ATC scheme does not provide sufficient detail on the specific drugs, there is e.g. no official way to code Saquinavir as hard or soft gel. Thus a slight alteration to the set of codes will be presented in the sections of the ART and MED tables. The alterations are designed to extend the existing structure of ATC.**

One of the benefits is that the structure of ATC allows easier statistics on e.g. drug classes

- J05AE Protease inhibitors
  - J05AE01 Saquinavir
  - J05AE02 Indinavir
  - J05AE03 Ritonavir
  - J05AE04 Nelfinavir
  - J05AE05 Amprenavir
  - J05AE06 Lopinavir
- J05AF Nucleoside and nucleotide reverse transcriptase inhibitors
  - J05AF01 Zidovudine
  - J05AF02 Didanosine
  - J05AF03 Zalcitabine

- J05AF04 Stavudine
- J05AF05 Lamivudine
- J05AF06 Abacavir
- J05AF07 Tenofovir disoproxil
- J05AF30 Combinations1
- J05AG Non-nucleoside reverse transcriptase inhibitors
  - J05AG01 Nevirapine
  - J05AG02 Delavirdine
  - J05AG03 Efavirenz

Although the codes might be harder to read they provide grouping mechanisms in the way they are coded. Interested readers should go to the [?ATC Website](#) to learn about the structure of ATC. A fully updated database of ATC codes and DDD (Defined Daily Dosage) is available for querying.

### Other codes

It is often necessary to code for values like ?Yes?, ?No? and ?Unknown?, this document suggests that the following codes should be used:

Code	Description
0	No
1	Yes
9	Unknown

Unknown should be used to identify the difference between a value that has not yet been collected (Empty) and a value that cannot be collected (Unknown). Empty values should be required where Unknown values make little sense to keep querying for a value.

Example ? weight:

Depending on the unit in which weight is measured, a different value for Unknown should be applied. In the case of kg the ?Unknown? code should be 999 and not just 9 or 99, the last two could be actual values.

Blank values, for SAS users also known as " ." and for database programmers known as NULL, should be used wherever specified in this protocol. However, sometimes it might be more correct just to omit the record if no value has been recorded, test has not been performed etc.

### Quality Assurance

In order to verify the consistency and correctness of the data, QA checks are made before the data is used. The QA checks applying to a given table are listed at the bottom of its article. Additionally, a list of all QA checks, including checks which do not directly apply to the HICDEP tables themselves, is available [here](#).



## tbIART - Antiretroviral treatment

holds type of antiretroviral drug, start and stop dates and reason for stopping

### Core fields

Note: Fields marked **bold** form the unique identifier for a record of the table.

Field name	Format	Description
<a href="#">PATIENT</a>	character (or numeric if possible)	identifies patient
<a href="#">ART_ID</a>	character. see <a href="#">coding table</a> for valid codings.	represents the antiretroviral treatment
<a href="#">ART_SD</a>	yyyy-mm-dd	date of initiation of treatment
<a href="#">ART_ED</a>	yyyy-mm-dd	date of stopping treatment
<a href="#">ART_RS</a>	character. see <a href="#">coding table</a> for valid codings.	reason for stopping treatment

### Additional fields

Depending on the aim of the study it might be needed to gather both the dosage and the frequency of the dosage taken. However many cohorts do not collect this date and thus these fields are optional.

Field name	Format	Description
<a href="#">ART_DO</a>	numeric	Dosage (mg or mL) per intake unless ART_FR=-1
<a href="#">ART_FR</a>	numeric: <ul style="list-style-type: none"> <li>-1 = Frequency not known. ART_DO contains dosage per day</li> <li>0.33 = 1 dose every third day</li> <li>0.5 = 1 dose every second day</li> <li>1 = 1 daily dose/qd</li> <li>2 = 2 daily doses/bid</li> <li>3 = 3 daily doses/tid</li> <li>4... = code gives number of daily doses</li> </ul>	Frequency

It may also be necessary to record the start and end time:

Field name	Format	Description
<a href="#">ART_ST</a>	hh:mm	Start hour and minute of the day
<a href="#">ART_ET</a>	hh:mm	Stop hour and minute of the day

### QA Checks

Table	Crosstable	Error Code	Description	Study specific	HICDEP?
AllTables	CrossTable	ATC001	any date in database after DEATH_D in tbILTFU		YES
AllTables	CrossTable	ATC002	any date in database after DROP_D in tbILTFU		YES
AllTables	CrossTable	ATC003	any date in database before BIRTH_D in tbIBAS		YES

AllTables	CrossTable	ATC004	any date in database in the future		YES
AllTables	CrossTable	ATC005	patients submitted previously who have been missed out		YES
AllTables	CrossTable	ATC006	Any fields not coded as coding lists on table definition		YES
tblART	WithinTable	AW001	ART_RS not null, but end date NULL		YES
tblART	WithinTable	AW002	ART_RS null but end-date non NULL		YES
tblART	WithinTable	AW003	ART_RS=98 yet ART_OTH is null	EPPICC	NO
tblART	WithinTable	AW004	Duplicate records for same cohort, patient, art_id and art_sd		YES
tblART	WithinTable	AW005	Dose out of range for those dose units and frequency	PaediatricOnly	YES
tblART	WithinTable	AW006	Missing art_fr	PaediatricOnly	YES
tblART	WithinTable	AW007	Missing art_do	PaediatricOnly	YES
tblART	WithinTable	AW008	Missing patient		YES
tblART	WithinTable	AW009	Missing art_id		YES
tblART	WithinTable	AW010	Missing art_sd		YES
tblART	WithinTable	AW011	Overlapping periods of same drug		YES
tblART	WithinTable	AW012	Double reporting - records reported for both combination drugs and their components		YES
tblART	WithinTable	AW013	Periods of overlap of contra-indicated drugs		YES
tblART	WithinTable	AW014	Restart of same drug without a stop		YES
tblART	WithinTable	AW015	ART_SD greater than or equal to ART_ED		YES
tblART	CrossTable	AC001	Patient has no record in table BAS		YES
tblART	CrossTable	AC002	Records exist in tblART yet RECAR_T_Y=0 in tblBAS		YES

## tbIBAS - Basic clinical, background and demographic information

holds basic information such as demographics, basic clinical information, date of AIDS diagnosis, death and drop-out information

### Core fields

Note: Fields marked **bold** form the unique identifier for a record of the table.

Field name	Format	Description
<a href="#"><u>PATIENT</u></a>	character (or numeric if possible)	Code to identify patient (Cohort Patient ID)
<a href="#"><u>CENTER</u></a>	character	Code for Clinic/Centre/Hospital where patient is seen.
<a href="#"><u>BIRTH_D</u></a>	yyyy-mm-dd	Birth date
<a href="#"><u>FRSVIS_D</u></a>	yyyy-mm-dd	First seen at clinic
<a href="#"><u>ENROL_D</u></a>	yyyy-mm-dd	Date of enrolment into the cohort
<a href="#"><u>GENDER</u></a>	numeric: <ul style="list-style-type: none"> <li>1 = Male</li> <li>2 = Female</li> <li>9 = Unknown</li> </ul>	Gender/sex
<a href="#"><u>HEIGH</u></a>	numeric (metric): 999 = Unknown	Height of patient at visit/most current
<a href="#"><u>MODE</u></a>	numeric. see <a href="#">coding table</a> for valid codings.	Mode of infection
<a href="#"><u>ORIGIN</u></a>	character (1-3 letter/numeric codes). see <a href="#">coding table</a> for valid codings.	Country or region of birth
<a href="#"><u>ETHNIC</u></a>	numeric. see <a href="#">coding table</a> for valid codings.	Ethnicity of patient. Please take the <a href="#">additional notes</a> into consideration when using this field.
<a href="#"><u>EDU_LVL</u></a>	numeric. see <a href="#">coding table</a> for valid codings.	Last completed education Level. ISCED97 refers to the <a href="#">?1997 International Standard Classification of Education</a>
<a href="#"><u>HIV_POS_D</u></a>	yyyy-mm-dd	Date of first positive HIV test
<a href="#"><u>SEROCO_D</u></a>	yyyy-mm-dd	Date of seroconversion
<a href="#"><u>RECart_Y</u></a>	numeric: <ul style="list-style-type: none"> <li>1 = Yes</li> <li>0 = No</li> <li>9 = Unknown</li> </ul>	Has the patient received antiretroviral treatment?
<a href="#"><u>RECart_D</u></a>	yyyy-mm-dd	Date ART started
<a href="#"><u>LTART_D</u></a>	yyyy-mm-dd	Date last assessed for ART. If started ART, last date known to be on ART, or if not on ART, last date ART free.
<a href="#"><u>AIDS_Y</u></a>	numeric: <ul style="list-style-type: none"> <li>1 = Yes</li> <li>0 = No</li> <li>9 = Unknown</li> </ul>	Has patient been given an AIDS diagnosis?
<a href="#"><u>AIDS_D</u></a>	yyyy-mm-dd	IF YES, date of AIDS diagnosis

### Additional fields

For mode of infection and origin a set of other fields are often used to capture what cannot be coded. These fields are represented here as optional fields as it is the intention that the suggested codes applied to the MODE and ORIGIN should be able to cover all possible values.

Field name	Format	Description
<a href="#">MODE_OTH</a>	character	Mode of infection OTHER
<a href="#">ORI_OTH</a>	character	Origin of patient OTHER
<a href="#">CENS_D</a>	yyyy-mm-dd	The date of administrative censoring; the last date the database has been updated for this patient
<a href="#">SEROHOW</a>	numeric: <ul style="list-style-type: none"> <li>1 = Midpoint between last neg/first pos test</li> <li>2 = Lab evidence of seroconversion</li> <li>3 = Seroconversion illness</li> <li>4 = Other</li> <li>9 = Unknown</li> </ul>	For Seroconverters only: How was the seroconversion date determined?
<a href="#">NAIVE_Y</a>	numeric: <ul style="list-style-type: none"> <li>0 = No</li> <li>1 = Yes</li> <li>9 = Unknown</li> </ul>	Is the patient ART-naïve upon enrollment?

### [QA Checks](#)

Table	Crosstable	Error Code	Description	Study specific	HICDEP?
AllTables	CrossTable	ATC001	any date in database after DEATH_D in tbILTFU		YES
AllTables	CrossTable	ATC002	any date in database after DROP_D in tbILTFU		YES
AllTables	CrossTable	ATC003	any date in database before BIRTH_D in tbIBAS		YES
AllTables	CrossTable	ATC004	any date in database in the future		YES
AllTables	CrossTable	ATC005	patients submitted previously who have been missed out		YES
AllTables	CrossTable	ATC006	Any fields not coded as coding lists on table definition		YES
tbIBAS	WithinTable	BW001	AIDS date < SEROCO_D		YES
tbIBAS	WithinTable	BW002	Duplicate patients		YES
tbIBAS	WithinTable	BW003	First 3 chars of PATIENT don't form valid cohort code	CascadeOnly	YES
tbIBAS	WithinTable	BW004	Missing PATIENT		YES
tbIBAS	WithinTable	BW005	Missing CENTER		YES

tblBAS	WithinTable	BW006	Missing BIRTH_D		YES
tblBAS	WithinTable	BW007	Missing FRSVIS_D		YES
tblBAS	WithinTable	BW008	Missing ENROL_D		YES
tblBAS	WithinTable	BW009	Missing GENDER		YES
tblBAS	WithinTable	BW010	Missing HEIGH		YES
tblBAS	WithinTable	BW011	Missing MODE		YES
tblBAS	WithinTable	BW012	Missing MODE_OTH if MODE=90		YES
tblBAS	WithinTable	BW013	Missing ORIGIN		YES
tblBAS	WithinTable	BW014	Missing ETHNIC		YES
tblBAS	WithinTable	BW015	Missing SEROCO_D		YES
tblBAS	WithinTable	BW016	Missing RECARD_Y		YES
tblBAS	WithinTable	BW017	Missing AIDS_Y		YES
tblBAS	WithinTable	BW018	Missing AIDS_D if AIDS_Y=1		YES
tblBAS	WithinTable	BW019	BIRTH_D out of range (15-85 yrs)	CascadeOnly	YES
tblBAS	WithinTable	BW020	BIRTH_D out of range (<18)	PENTA	YES
tblBAS	CrossTable	BC001	RECARD_Y=1 but no records in tblART		YES
tblBAS	CrossTable	BC002	AIDS_Y=0, but AIDS-defining records in tblDIS		YES
tblBAS	CrossTable	BC003	AIDS_Y=1 but no AIDS-defining records in tblDIS table		YES

## tbICANC - Diagnosis of Cancer

There is a large heterogeneity across cohorts regarding the documentation of cancer diagnoses. Some use ICD9 or ICD10 for cancer location and have no information on histology. Others have cross-linked with cancer registries and are able to provide very granular ICD-O-3 histology codes. There are also proprietary systems such as the NA-ACCORD short list ([instructions](#), [list\\_1](#) & [list\\_2](#)). In tbICANC we suggest to not restrict to a single coding system but to allow cohorts to provide the data in different systems as long as definitions and mapping tables are submitted alongside the data.

For new cohorts we propose to use ICD9/10 coding schemes for cancer location and ICD-O-3 for histology if these results can be obtained from cancer registries.

Each row of tbICANC consists of a diagnosis of one cancer type.

If AIDS defining cancers are reported in tbICANC they should NOT be deleted from tbIDIS.

### Core fields

Note: Fields marked **bold** form the unique identifier for a record of the table.

### Additional fields

#### [QA Checks](#)

Table	Crosstable	Error Code	Description	Study specific	HICDEP?
AllTables	CrossTable	ATC001	any date in database after DEATH_D in tbILTFU		YES
AllTables	CrossTable	ATC002	any date in database after DROP_D in tbILTFU		YES
AllTables	CrossTable	ATC003	any date in database before BIRTH_D in tbIBAS		YES
AllTables	CrossTable	ATC004	any date in database in the future		YES
AllTables	CrossTable	ATC005	patients submitted previously who have been missed out		YES
AllTables	CrossTable	ATC006	Any fields not coded as coding lists on table definition		YES

## tbICEP - Clinical Events and Procedures

holds type and date of adverse events including serious non-AIDS conditions

### Core fields

Note: Fields marked **bold** form the unique identifier for a record of the table.

Field name	Format	Description
<a href="#">EVENT_ID</a>	numeric	Unique Event Identifier (foreign key to the different event tables)
<a href="#">PATIENT</a>	character (or numeric if possible)	Code to identify patient (Cohort Patient ID)
<a href="#">CEP_D</a>	yyyy-mm-dd	date of event
<a href="#">CEP_ID</a>	character. see <a href="#">coding table</a> for valid codings.	identifies type of event
<a href="#">CEP_SPEC</a>	character. see <a href="#">coding table</a> for valid codings.	further specification
<a href="#">CEP_V</a>	numeric. See <a href="#">coding table</a> for interpretation.	Depending on CEP_ID and CEP_SPEC: value of given event
<a href="#">SRCDOC_Y</a>	numeric: <ul style="list-style-type: none"> <li>1 = Yes</li> <li>0 = No</li> </ul>	whether the source documentation is available
<a href="#">SRCDOC_D</a>	yyyy-mm-dd	date for source documentation verification
<a href="#">VERIFY_Y</a>	numeric: <ul style="list-style-type: none"> <li>1 = Yes</li> <li>0 = No</li> </ul>	Has the monitor verified the source documentation?
<a href="#">VERIFY_D</a>	yyyy-mm-dd	date for monitor verification
<a href="#">APPROV_Y</a>	numeric: <ul style="list-style-type: none"> <li>1 = Yes</li> <li>0 = No</li> </ul>	final verification/approval
<a href="#">APPROV_D</a>	yyyy-mm-dd	final verification date
<a href="#">APPROV_S</a>	character	signature for final verification

### Additional fields

Field name	Format	Description
<a href="#">CEP_Y</a>	numeric: <ul style="list-style-type: none"> <li>1 = Yes</li> <li>0 = No</li> <li>9 = Unknown</li> </ul>	has the patient had an event?
<a href="#">CEP_NAME</a>	character	full name of the event
<a href="#">CEP_DESCRIP</a>	character	full description of the event

<a href="#">CEP_R_Y</a>	numeric: <ul style="list-style-type: none"> <li>• 0 = not related</li> <li>• 1 = definitive</li> <li>• 2 = remote/unlikely</li> <li>• 3 = possible</li> <li>• 4 = probable</li> </ul>	relation to treatment
-------------------------	---	-----------------------

### QA Checks

Table	Crosstable	Error Code	Description	Study specific	HICDEP?
AllTables	CrossTable	ATC001	any date in database after DEATH_D in tblLTFU		YES
AllTables	CrossTable	ATC002	any date in database after DROP_D in tblLTFU		YES
AllTables	CrossTable	ATC003	any date in database before BIRTH_D in tblBAS		YES
AllTables	CrossTable	ATC004	any date in database in the future		YES
AllTables	CrossTable	ATC005	patients submitted previously who have been missed out		YES
AllTables	CrossTable	ATC006	Any fields not coded as coding lists on table definition		YES



## tbDELIVERY\_CHILD - Delivery information related to the child(ren)

This table holds information about a delivery **specific to a child**. Mother-specific data is kept in [tbDELIVERY\\_MUM](#).

Please also read the [notes on pregnancy tables](#).

### Core fields

Note: Fields marked **bold** form the unique identifier for a record of the table.

Field name	Format	Description
<b>MOTHER_ID</b>	Character (or numeric if possible)	Patient ID of mother of the child
<b>MEMRUP_D</b>	yyyy-mm-dd	Date of rupture of membranes
<b>CHILD_ID</b>	Character (or numeric if possible)	Patient ID of the child
<b>B_SEQ</b>	numeric	If multiple births, indicate number (1=first born)
<b>DELIV_D</b>	yyyy-mm-dd	Date of delivery
<b>DELIV_T</b>	hh:mm	Time of delivery
<b>DELIV_M</b>	character: <ul style="list-style-type: none"> <li>1=Vaginally, spontaneous</li> <li>2=Vaginally, forceps</li> <li>3=Vaginally, vacuum</li> <li>10= Cesarean section, primary/elective (before onset of labour and rupture of membrane)</li> <li>11=Cesarean section, Secondary</li> </ul>	Mode of delivery
<b>LABOUR_P</b>	numeric	Duration of labour in hours (from cervical dilatation (>=3 cm) until delivery)
<b>BREECH_Y</b>	numeric: <ul style="list-style-type: none"> <li>1 = Yes</li> <li>0 = No</li> <li>9 = Unknown</li> </ul>	Was the child born from a breech presentation?

### QA Checks

Table	Crosstable	Error Code	Description	Study specific	HICDEP?
AllTables	CrossTable	ATC001	any date in database after DEATH_D in tbILTFU		YES
AllTables	CrossTable	ATC002	any date in database after DROP_D in tbILTFU		YES
AllTables	CrossTable	ATC003	any date in database before BIRTH_D in tbIBAS		YES
AllTables	CrossTable	ATC004	any date in database in the future		YES
AllTables	CrossTable	ATC005	patients submitted previously who have been missed out		YES

AllTables	CrossTable	ATC006	Any fields not coded as coding lists on table definition		YES
tbIDELIVERY_CHILD	CrossTable	DCC001	MOTHER_ID doesn't exist in tbIPREG		YES
tbIDELIVERY_CHILD	CrossTable	DCC002	CHILD_ID doesn't exist in tbIPREG_OUT		YES
tbIDELIVERY_CHILD	WithinTable	DCW001	Duplicate records for same Mother_ID, CHILD_ID		YES
tbIDELIVERY_CHILD	WithinTable	DCW002	B_SEQ >4		YES
tbIDELIVERY_CHILD	WithinTable	DCW003	B_SEQ>1, but no record exists for same Mother, Child with B_SEQ=1		YES
tbIDELIVERY_CHILD	WithinTable	DCW005	DELIV_D < MEMRUP_D		YES
tbIDELIVERY_CHILD	WithinTable	DCW004	LABOUR_P >72		YES
tbIDELIVERY_CHILD	WithinTable	DCW006	2 children, same mother, same MEMRUP_D, DELIV_D >DELIV_D , yet B_SEQ< B_SEQ		YES

## tbIDELIVERY\_MUM - Delivery information related to the mother

This table contains information about the delivery **specific to the mother**. Child-specific information is recorded in [tbIDELIVERY\\_CHILD](#) instead.

Please also read the [notes on pregnancy tables](#).

### Core fields

Note: Fields marked **bold** form the unique identifier for a record of the table.

Field name	Format	Description
<b>MOTHER_ID</b>	Character (or numeric if possible)	Patient ID of mother of the child
<b>PREG_SEQ</b>	numeric	Sequence number of the pregnancy for the specified mother
<b>MEMRUP_D</b>	yyyy-mm-dd	Date of rupture of membranes
<b>MEMRUP_T</b>	hh:mm	Time of rupture of membranes
<b>LABOUR</b>	character: <ul style="list-style-type: none"> <li>1=Spontaneous</li> <li>2=Induced</li> <li>3=No labour (elective C-section)</li> </ul>	Onset of labor
<b>INTERV</b>	character: <ul style="list-style-type: none"> <li>0=No</li> <li>1=Fetal blood sampling FBS</li> <li>2=Internal electrodes</li> <li>3=1+2</li> <li>90=Other</li> <li>91=1+Other</li> <li>92=2+Other</li> <li>93=1+2+Other</li> </ul>	Interventions during delivery
<b>INTERV_O</b>	character	Interventions during delivery - other
<b>TEAR_Y</b>	numeric: <ul style="list-style-type: none"> <li>1=Yes</li> <li>0=No</li> <li>9=Unknown</li> </ul>	Episiotomy/tear
<b>BLDLOSS</b>	numeric (mL)	Estimated blood loss during delivery
<b>CONTREAT</b>	character: <ul style="list-style-type: none"> <li>0=No, treatment has been interrupted</li> <li>1=Yes, at the foreseen intervals</li> <li>2=Yes, but not at foreseen time points</li> <li>9=Unknown</li> </ul>	Did the patient continue the usual antiretroviral therapy?
<b>DISCHA_D</b>	yyyy-mm-dd	Date of discharge from hospital

### QA Checks

Table	Crosstable	Error Code	Description	Study specific	HICDEP?
-------	------------	------------	-------------	----------------	---------

AllTables	CrossTable	ATC001	any date in database after DEATH_D in tblLTFU		YES
AllTables	CrossTable	ATC002	any date in database after DROP_D in tblLTFU		YES
AllTables	CrossTable	ATC003	any date in database before BIRTH_D in tblBAS		YES
AllTables	CrossTable	ATC004	any date in database in the future		YES
AllTables	CrossTable	ATC005	patients submitted previously who have been missed out		YES
AllTables	CrossTable	ATC006	Any fields not coded as coding lists on table definition		YES
tbDELIVERY_MUM	CrossTable	DMC001	MOTHER_ID doesn't exist in tblPREG		YES
tbDELIVERY_MUM	CrossTable	DMC002	MOTHER_ID doesn't exist in tblDELIVERY_CHILD with same MEMRUP_D		YES
tbDELIVERY_MUM	WithinTable	DMW001	PREG_SEQ invalid (PREG_SEQ>1 and PREG_SEQ=x, but PREG_SEQ=x-1 doesn't exist)		YES
tbDELIVERY_MUM	WithinTable	DMW002	DISCHA_D < MEMRUP_D		YES
tbDELIVERY_MUM	WithinTable	DMW004	INTERV=0-3 but INTERV_O non null		YES
tbDELIVERY_MUM	WithinTable	DMW003	INTERV=90,91, 92 or 93 but INTERV_O null		YES

## tbIDIS - Opportunistic infections

holds type and date of CDC-C diseases.

### Core fields

Note: Fields marked **bold** form the unique identifier for a record of the table.

Field name	Format	Description
<a href="#">PATIENT</a>	character (or numeric if possible)	Code to identify patient (Cohort Patient ID)
<a href="#">DIS_ID</a>	character. see <a href="#">coding table</a> for valid codings.	Code to identify event
<a href="#">DIS_D</a>	yyyy-mm-dd	Date of event
<a href="#">DIS_WD</a>	numeric. see <a href="#">coding table</a> for valid codings.	Means of diagnosis
<a href="#">DIS_OTH</a> <sup>1</sup>	character	Other location, only to be filled out if code alone is not sufficient

<sup>1</sup> DIS\_OTH might be part of the record's unique identification

### Additional fields

Please see [tbICEP](#) for specification on optional fields.

Field name	Format	Description
<a href="#">DIS_ED</a>	yyyy-mm-dd	End of opportunistic infection

### QA Checks

Table	Crosstable	Error Code	Description	Study specific	HICDEP?
AllTables	CrossTable	ATC001	any date in database after DEATH_D in tbILTFU		YES
AllTables	CrossTable	ATC002	any date in database after DROP_D in tbILTFU		YES
AllTables	CrossTable	ATC003	any date in database before BIRTH_D in tbIBAS		YES
AllTables	CrossTable	ATC004	any date in database in the future		YES
AllTables	CrossTable	ATC005	patients submitted previously who have been missed out		YES
AllTables	CrossTable	ATC006	Any fields not coded as coding lists on table definition		YES
tbIDIS	WithinTable	DW001	Duplicate records for same DIS_ID and same DIS_D		YES
tbIDIS	WithinTable	DW002	Miscoded DIS_WD as codes on table definition		YES

tblDIS	WithinTable	DW003	Miscoded DIS_ID - as in code list attached to table definition		YES
tblDIS	WithinTable	DW004	DIS_D missing		YES
tblDIS	WithinTable	DW005	DIS_ID missing		YES
tblDIS	WithinTable	DW006	Same event recorded twice - 2 records, same DIS_ID, DIS_D within 6 months		YES
tblDIS	WithinTable	DW007	DIS_ED present but before DIS_D		YES
tblDIS	CrossTable	DC001	Patient has no record in BAS		YES
tblDIS	CrossTable	DC002	AIDS-defining records, yet AIDS=0 in tblBAS		YES
tblDIS	CrossTable	DC003	First AIDS-defining DIS_D not equal to AIDS_D in tblBAS		YES

## tbILAB - Laboratory values

holds type, date, value and unit of laboratory tests.

### Core fields

Note: Fields marked **bold** form the unique identifier for a record of the table.

Field name	Format	Description
<a href="#">PATIENT</a>	character (or numeric if possible)	Code to identify patient (Cohort Patient ID)
<a href="#">LAB_ID</a>	character. see <a href="#">coding table</a> for valid codings.	Code representing the measurement
<a href="#">LAB_D</a>	yyyy-mm-dd	Date of measurement/sample
<a href="#">LAB_V</a>	numeric: -1 = undetectable or detection limit as negative value  for dipstick results: <ul style="list-style-type: none"> <li>• 0 = Negative</li> <li>• 9 = Trace</li> <li>• 1 = 1+</li> <li>• 2 = 2+</li> <li>• 3 = 3+</li> <li>• 4 = 4+</li> </ul>	Value of measurement
<a href="#">LAB_U</a>	numeric. see <a href="#">coding table</a> for valid codings.	Unit of measurement

### Additional fields

Other detailed information regarding the patient and the measurement would be relevant, like the proposed fasting information shown below.

Field name	Format	Description
<a href="#">LAB_FA</a>	numeric: <ul style="list-style-type: none"> <li>• 1 = Yes</li> <li>• 0 = No</li> <li>• 9 = Unknown</li> </ul>	Was the blood sample taken while fasting?
<a href="#">LAB_ST</a>	<ul style="list-style-type: none"> <li>• WB = Whole Blood</li> <li>• P = Plasma</li> <li>• S = Serum</li> <li>• U24 = 24h Urine</li> <li>• U = Urine</li> </ul>	Specimen type

Depending on the set of measurements collected and the mandatory fields applicable to these individual measurements, it might be useful to separate the LAB table into several sub tables. This is already shown for the CD4 and RNA measurements: the level of detail needed for CD4 is less than for the LAB variables in general (no unit since it's always the same), while for RNA the data required is more detailed (assay and detection limit).

### QA Checks

Table	Crosstable	Error Code	Description	Study specific	HICDEP?
AllTables	CrossTable	ATC001	any date in database after DEATH_D in tbILTFU		YES

AllTables	CrossTable	ATC002	any date in database after DROP_D in tblLTFU		YES
AllTables	CrossTable	ATC003	any date in database before BIRTH_D in tblBAS		YES
AllTables	CrossTable	ATC004	any date in database in the future		YES
AllTables	CrossTable	ATC005	patients submitted previously who have been missed out		YES
AllTables	CrossTable	ATC006	Any fields not coded as coding lists on table definition		YES
tblLAB	WithinTable	LW001	Duplicate records		YES
tblLAB	WithinTable	LW002	Missing LAB_D		YES
tblLAB	WithinTable	LW003	Missing LAB_V		YES
tblLAB	WithinTable	LW004	Missing LAB_U		YES
tblLAB	WithinTable	LW005	Missing LAB_ID		YES
tblLAB	WithinTable	LW008	Missing LAB_FA		YES
tblLAB	WithinTable	LW009	Missing LAB_ST		YES
tblLAB	WithinTable	LW010	LAB_V out of range for unit LAB_U		YES
tblLAB	CrossTable	LC001	Patient has no record in BAS table		YES



## tbILAB\_BP - Laboratory values - Blood pressure

holds date, diastolic and systolic values and unit of blood pressure measurements.

### Core fields

Note: Fields marked **bold** form the unique identifier for a record of the table.

Field name	Format	Description
<b><u>PATIENT</u></b>	character (or numeric if possible)	Code to identify patient (Cohort Patient ID)
<b><u>BP_D</u></b>	yyyy-mm-dd	Date of Measurement/Sample
<b><u>BP_SYS</u></b>	numeric	Systolic Blood Pressure
<b><u>BP_DIA</u></b>	numeric	Diastolic Blood Pressure
<b><u>BP_U</u></b>	numeric. see <a href="#">coding table</a> for valid codings.	Unit of measurement

### QA Checks

Table	Crosstable	Error Code	Description	Study specific	HICDEP?
AllTables	CrossTable	ATC001	any date in database after DEATH_D in tbILTFU		YES
AllTables	CrossTable	ATC002	any date in database after DROP_D in tbILTFU		YES
AllTables	CrossTable	ATC003	any date in database before BIRTH_D in tbIBAS		YES
AllTables	CrossTable	ATC004	any date in database in the future		YES
AllTables	CrossTable	ATC005	patients submitted previously who have been missed out		YES
AllTables	CrossTable	ATC006	Any fields not coded as coding lists on table definition		YES

## tbILAB\_CD4 - Laboratory values

holds date and value of CD4 measurements.

*Note:* If needed, a CD8 table (tbILAB\_CD8) could be formed from the same structure.

### Core fields

*Note:* Fields marked **bold** form the unique identifier for a record of the table.

Field name	Format	Description
<a href="#">PATIENT</a>	character (or numeric if possible)	Code to identify patient (Cohort Patient ID)
<a href="#">CD4_D</a>	yyyy-mm-dd	Date of measurement
<a href="#">CD4_V</a>	numeric (per microliter): -1 = undetectable or detection limit as negative value	Value of CD4 measurement

### Additional fields

[CD4\\_V](#) is assumed to contain absolute CD4 cell counts per mL as standard. In case CD4 % (with respect to CD45+ lymphocytes as denominator) should be collected as well, please append the following field to the table:

Field name	Format	Description
<a href="#">CD4_U</a>	numeric with codes (or full string): <ul style="list-style-type: none"> <li>1 = cells/<math>\mu</math>l</li> <li>2 = %</li> </ul>	Unit of measurement

### QA Checks

Table	Crosstable	Error Code	Description	Study specific	HICDEP?
AllTables	CrossTable	ATC001	any date in database after DEATH_D in tbILTFU		YES
AllTables	CrossTable	ATC002	any date in database after DROP_D in tbILTFU		YES
AllTables	CrossTable	ATC003	any date in database before BIRTH_D in tbIBAS		YES
AllTables	CrossTable	ATC004	any date in database in the future		YES
AllTables	CrossTable	ATC005	patients submitted previously who have been missed out		YES
AllTables	CrossTable	ATC006	Any fields not coded as coding lists on table definition		YES
tbILAB_CD4	WithinTable	CW001	CD4 value out of range		YES
tbILAB_CD4	WithinTable	CW002	Duplicate records for same date		YES
tbILAB_CD4	WithinTable	CW003	Missing CD4_D		YES
tbILAB_CD4	WithinTable	CW004	Missing CD4_V		YES

tblLAB_CD4	WithinTable	CW005	Missing CD4_U		YES
tblLAB_CD4	WithinTable	CW006	Miscoded CD4_U as coding list on table definition		YES
tblLAB_CD4	WithinTable	CW007	CD4_U=2 (percentage) and CD4_V>100		YES
tblLAB_CD4	WithinTable	CW008	CD4_U=(1 or 3) and CD4_V>3000		YES
tblLAB_CD4	WithinTable	CW009	CD4 counts spike up or down suddenly (large change in less than a year)		YES
tblLAB_CD4	CrossTable	CC001	Patient has no record in BAS		YES

## tbILAB\_RES - Resistance testing

holds background information on the resistance test, laboratory, library, kit, software and type of test

Note: This table is tightly linked to [tbILAB\\_RES\\_LVL\\_1](#), [tbILAB\\_RES\\_LVL\\_2](#) and [tbILAB\\_RES\\_LVL\\_3](#).

**Resistance should be reported at lowest level of interpretation possible ? so if the nucleotide sequence is available this should be reported rather than the list of mutations or resistance scores. However, the resistance test results should be captured if they have been part of the physician's treatment decisions for the patient.**

**Non-amplifiable resistance tests should not be reported.**

These four tables are designed to capture several possible formats the clinics and cohorts might have recorded resistance test data in. Once this data is gathered it should like all other tables be quality assessed.

### Core fields

Note: Fields marked **bold** form the unique identifier for a record of the table.

Field name	Format	Description
<a href="#">PATIENT</a>	character (or numeric if possible)	Code to identify patient (Cohort Patient ID)
<a href="#">TEST_ID</a>	character (or numeric if possible)	An arbitrary value identifying a resistance test result
<a href="#">SAMPLE_D</a>	yyyy-mm-dd	Date of the actual sample taken (NOT the test date)
<a href="#">SEQ_DT</a>	yyyy-mm-dd hh:mm	Date and time when the sequencing was performed
<a href="#">LAB</a>	character	Name of laboratory where the test was performed
<a href="#">LIBRARY</a>	character	Library/algorithm used to identify resistance mutations
<a href="#">REFSEQ</a>	character	Name/identifier of reference strain used to find mutations
<a href="#">KIT</a>	character	Vendor and version/name of the kit used for the test
<a href="#">SOFTWARE</a>	character	Software and version used to determine resistance
<a href="#">TESTTYPE</a>	numeric: <ul style="list-style-type: none"> <li>• 1 = Genotype</li> <li>• 2 = Phenotype</li> <li>• 9 = Other</li> </ul>	Type of test
<a href="#">VIRUSTYPE</a>	numeric: <ul style="list-style-type: none"> <li>• 1 = HIV</li> <li>• 2 = HCV</li> </ul>	Type of Virus
<a href="#">SUBTYPE</a>	character	Subtype of HIV- or HCV-RNA

### Additional fields

As shown with the core fields, the *TEST\_ID* is the link between the 3 levels of data and the test background information table.

Some prior assessment of the assigned test identifiers has to be done in order to avoid duplicates.

In a running database the duplicate issues are easily resolved by adding a unique auto-generated key as the identifier between 3 levels of data and the test background information table.

Along with the *TEST\_ID* it might be necessary to store the ID assigned to the sample at both the testing laboratory but also the centres laboratory in order to track the sample.

Field name	Format	Description
<a href="#">SAMP_LAB</a>	character (or numeric if possible)	The assigned sample ID at the lab where the resistance test is preformed.
<a href="#">SAMP_INT</a>	character (or numeric if possible)	The assigned sample ID from the centre.

### QA Checks

Table	Crosstable	Error Code	Description	Study specific	HICDEP?
AllTables	CrossTable	ATC001	any date in database after DEATH_D in tblLTFU		YES
AllTables	CrossTable	ATC002	any date in database after DROP_D in tblLTFU		YES
AllTables	CrossTable	ATC003	any date in database before BIRTH_D in tblBAS		YES
AllTables	CrossTable	ATC004	any date in database in the future		YES
AllTables	CrossTable	ATC005	patients submitted previously who have been missed out		YES
AllTables	CrossTable	ATC006	Any fields not coded as coding lists on table definition		YES
tblLAB_RES	WithinTable	LRW001	Duplicate records for same patient on same date		YES
tblLAB_RES	WithinTable	LRW002	Missing PATIENT		YES
tblLAB_RES	WithinTable	LRW003	Missing TEST_ID		YES
tblLAB_RES	WithinTable	LRW004	Missing SAMPLE_D		YES
tblLAB_RES	WithinTable	LRW005	Missing SEQ_DT		YES
tblLAB_RES	WithinTable	LRW006	Missing LAB		YES
tblLAB_RES	WithinTable	LRW007	Missing LIBRARY		YES
tblLAB_RES	WithinTable	LRW008	Missing REFSEQ		YES
tblLAB_RES	WithinTable	LRW009	Missing KIT		YES
tblLAB_RES	WithinTable	LRW010	Missing SOFTWARE		YES
tblLAB_RES	WithinTable	LRW011	Missing TESTTYPE		YES
tblLAB_RES	WithinTable	LRW012	Missing SUBTYPE		YES
tblLAB_RES	WithinTable	LRW013	SEQ_DT has no time part		YES
tblLAB_RES	WithinTable	LRW014	Missing VIRUSTYPE		YES

tblLAB_RES	CrossTable	LRC001	PATIENT has no record in tblBAS		YES
tblLAB_RES	CrossTable	LRC002	This TEST_ID has both LVL_1 and LVL_2 records		YES

## tbILAB\_RES\_LVL\_1 - Nucleotide sequences (PRO, RT, GP41, GP120)

holds nucleoside sequence for the PRO and RT sequences. No entry is made if the test was a phenotype test.

*Note:* This table is tightly linked to [tbILAB\\_RES](#).

### Core fields

*Note:* Fields marked **bold** form the unique identifier for a record of the table.

Field name	Format	Description
<a href="#">TEST_ID</a>	character (or numeric if possible)	Identifier linking this record to <a href="#">tbILAB_RES</a>
<a href="#">SEQTYPE</a>	character: <ul style="list-style-type: none"> <li>PRO = PRO sequence</li> <li>RT = RT sequence</li> <li>GP41 = GP41 sequence</li> <li>GP120 = GP120 sequence</li> </ul>	Type of nucleotide sequence if available
<a href="#">SEQ_START</a>	numeric	Start position for the sequence
<a href="#">SEQ_STOP</a>	numeric	Stop position for the sequence
<a href="#">SEQ_NUC</a>	character	Nucleotide sequence if available

### Additional fields

In cases where the amino acid sequence is collected rather than the nucleotide sequence, the field *SEQ\_NUC* might be replaced with *SEQ\_AA*, which is the nucleotide sequence, expressed in an amino acid sequence:

Field name	Format	Description
<a href="#">SEQ_AA</a>	character	Amino acid sequence if available (empty if test was phenotype)

However using the amino acid sequence does not give the same detail of data as the nucleoside sequence: wobbles in the nucleoside sequence can either complicate the reading and alignment of the amino acid sequence or the wobbles can be lost and silent mutations are lost.

### QA Checks

Table	Crosstable	Error Code	Description	Study specific	HICDEP?
AllTables	CrossTable	ATC001	any date in database after DEATH_D in tbILTFU		YES
AllTables	CrossTable	ATC002	any date in database after DROP_D in tbILTFU		YES
AllTables	CrossTable	ATC003	any date in database before BIRTH_D in tbIBAS		YES
AllTables	CrossTable	ATC004	any date in database in the future		YES
AllTables	CrossTable	ATC005	patients submitted previously who have been missed out		YES

AllTables	CrossTable	ATC006	Any fields not coded as coding lists on table definition		YES
tblLAB_RES_LVL_1	WithinTable	L1W001	Duplicate records per TEST_ID and SEQTYPE		YES
tblLAB_RES_LVL_1	WithinTable	L1W002	SEQ_START > SEQ_STOP		YES
tblLAB_RES_LVL_1	WithinTable	L1W003	SEQ_NUC contains invalid IUPAC character		YES
tblLAB_RES_LVL_1	WithinTable	L1W004	Missing TEST_ID		YES
tblLAB_RES_LVL_1	WithinTable	L1W005	Missing SEQ_START		YES
tblLAB_RES_LVL_1	WithinTable	L1W006	Missing SEQ_STOP		YES
tblLAB_RES_LVL_1	WithinTable	L1W007	Missing SEQ_NUC		YES
tblLAB_RES_LVL_1	CrossTable	L1C001	TEST_ID not in tblLAB_RES.TEST_ID		YES



## tbILAB\_RES\_LVL\_2 - Mutations

holds mutations and positions of PRO and RT sequences.

*Note:* This table is tightly linked to [tbILAB\\_RES](#).

### Core fields

*Note:* Fields marked **bold** form the unique identifier for a record of the table.

Field name	Format	Description
<a href="#">TEST_ID</a>	character (or numeric if possible)	Identifier linking this record to <a href="#">tbILAB_RES</a>
<a href="#">GENE</a>	character: <ul style="list-style-type: none"> <li>PRO = PRO sequence</li> <li>RT = RT sequence</li> <li>GP41 = GP41 sequence</li> <li>GP120 = GP120 sequence</li> </ul>	Type of sequence/gene (PRO, RT, GP41, GP120)
<a href="#">AA_POS</a>	numeric	Position of the mutation in the sequence
<a href="#">AA_POS_SUB</a>	character: <ul style="list-style-type: none"> <li>a = first</li> <li>b = second</li> <li>etc.</li> </ul>	Subposition used to code insertions
<a href="#">AA_FOUNDED_1</a>	character. empty = Amino acid has been deleted.	Mutation (Amino acid) found in the sequence
<a href="#">AA_FOUNDED_2</a>	character. empty = Amino acid has been deleted.	Mutation (Amino acid) found in the sequence (if more than 1)
<a href="#">AA_FOUNDED_3</a>	character. empty = Amino acid has been deleted.	Mutation (Amino acid) found in the sequence (if more than 2)
<a href="#">AA_FOUNDED_4</a>	character. empty = Amino acid has been deleted.	Mutation (Amino acid) found in the sequence (if more than 3)

[AA\\_FOUNDED\\_#](#) could be extended if mixtures with more than 4 amino acids are found.

### QA Checks

Table	Crosstable	Error Code	Description	Study specific	HICDEP?
AllTables	CrossTable	ATC001	any date in database after DEATH_D in tbILTFU		YES
AllTables	CrossTable	ATC002	any date in database after DROP_D in tbILTFU		YES
AllTables	CrossTable	ATC003	any date in database before BIRTH_D in tbIBAS		YES
AllTables	CrossTable	ATC004	any date in database in the future		YES
AllTables	CrossTable	ATC005	patients submitted previously who have been missed out		YES

AllTables	CrossTable	ATC006	Any fields not coded as coding lists on table definition		YES
tblLAB_RES_LVL_2	WithinTable	L2W001	AA_FOUND_x but nothing in AA_FOUND(x-1)		YES
tblLAB_RES_LVL_2	WithinTable	L2W002	Duplicate records per TEST_ID, GENE, AA_POS and AA_POS_SUB		YES
tblLAB_RES_LVL_2	WithinTable	L2W003	Missing TEST_ID		YES
tblLAB_RES_LVL_2	WithinTable	L2W004	Missing GENE		YES
tblLAB_RES_LVL_2	WithinTable	L2W005	Missing AA_POS		YES
tblLAB_RES_LVL_2	WithinTable	L2W006	Missing AA_POS_SUB		YES
tblLAB_RES_LVL_2	WithinTable	L2W007	Missing AA_FOUND_1		YES
tblLAB_RES_LVL_2	CrossTable	L2C001	TEST_ID not in tblLAB_RES.TEST_ID		YES

## tbILAB\_RES\_LVL\_3 - Resistance test result

holds resistance result in relation to antiretroviral drug.

*Note:* This table is tightly linked to [tbILAB\\_RES](#).

### Core fields

*Note:* Fields marked **bold** form the unique identifier for a record of the table.

Field name	Format	Description
<b>TEST_ID</b>	character (or numeric if possible)	Identifier linking this record to <a href="#">tbILAB_RES</a>
<b>ATC_CODE</b>	character	? <a href="#">ATC code</a> of the medication
<b>RES_SCOR</b>	character	Score of resistance or recommendation given from the test.

### Additional fields

For phenotype test results it will be necessary to extend the table with a field to store the cut-off value:

Field name	Format	Description
<b>RES_CUT</b>	character	Cut-off value for phenotype test result

However using the amino acid sequence does not give the same detail of data as the nucleoside sequence: wobbles in the nucleoside sequence can either complicate the reading and alignment of the amino acid sequence or the wobbles can be lost and silent mutations are lost.

### QA Checks

Table	Crosstable	Error Code	Description	Study specific	HICDEP?
AllTables	CrossTable	ATC001	any date in database after DEATH_D in tbILTFU		YES
AllTables	CrossTable	ATC002	any date in database after DROP_D in tbILTFU		YES
AllTables	CrossTable	ATC003	any date in database before BIRTH_D in tbIBAS		YES
AllTables	CrossTable	ATC004	any date in database in the future		YES
AllTables	CrossTable	ATC005	patients submitted previously who have been missed out		YES
AllTables	CrossTable	ATC006	Any fields not coded as coding lists on table definition		YES
tbILAB_RES_LVL_3	WithinTable	L3W001	Duplicate records for same TEST_ID and ATC_CODE		YES
tbILAB_RES_LVL_3	WithinTable	L3W002	Missing TEST_ID		YES
tbILAB_RES_LVL_3	WithinTable	L3W003	Missing ATC_CODE		YES
tbILAB_RES_LVL_3	WithinTable	L3W004	Missing RES_SCOR		YES

tblLAB_RES_LVL_3	CrossTable	L3C001	TEST_ID not in tblLAB_RES.TEST_ID		YES
------------------	------------	--------	--------------------------------------	--	-----

## tbILAB\_RNA - Laboratory values

holds date, value, detection limit and type of viral assay.

### Core fields

Note: Fields marked **bold** form the unique identifier for a record of the table.

Field name	Format	Description
<a href="#">PATIENT</a>	character (or numeric if possible)	Code to identify patient (Cohort Patient ID)
<a href="#">RNA_D</a>	yyyy-mm-dd	Date of Measurement/Sample
<a href="#">RNA_V</a>	numeric: -1 = undetectable or detection limit as negative value	HIV-RNA measurement value
<a href="#">RNA_L</a>	numeric	Lower Limit of HIV-RNA Assay
<a href="#">RNA_T</a>	numeric. see <a href="#">coding table</a> for valid codings.	IF AVAILABLE, What type of VIRAL ASSAY was used for this measurement?

### Additional fields

Field name	Format	Description
<a href="#">RNA_UL</a>	numeric	IF AVAILABLE, Upper Limit of assay

### QA Checks

Table	Crosstable	Error Code	Description	Study specific	HICDEP?
AllTables	CrossTable	ATC001	any date in database after DEATH_D in tbILTFU		YES
AllTables	CrossTable	ATC002	any date in database after DROP_D in tbILTFU		YES
AllTables	CrossTable	ATC003	any date in database before BIRTH_D in tbIBAS		YES
AllTables	CrossTable	ATC004	any date in database in the future		YES
AllTables	CrossTable	ATC005	patients submitted previously who have been missed out		YES
AllTables	CrossTable	ATC006	Any fields not coded as coding lists on table definition		YES
tbILAB_RNA	WithinTable	RW001	Duplicate records for same RNA_D		YES
tbILAB_RNA	WithinTable	RW002	RNA_V < 0 and RNA_L missing		YES
tbILAB_RNA	WithinTable	RW003	RNA_V > 10 000 and (RNA_V modulo 1000) = 1 and RNA_UL missing		YES
tbILAB_RNA	WithinTable	RW004	Missing RNA_D		YES

tblLAB_RNA	WithinTable	RW005	Missing RNA_V		YES
tblLAB_RNA	WithinTable	RW006	RNA_V > 10 million		YES
tblLAB_RNA	WithinTable	RW007	RNA_V < RNA_L and RNA_V >= 0		YES
tblLAB_RNA	WithinTable	RW008	RNA_V > RNA_UL		YES
tblLAB_RNA	CrossTable	RC001	Patient has no record in BAS		YES

## tbILAB\_VIRO - Laboratory values - viro-/serology

holds test results for viro-/serological tests (hepatitis etc.)

### Core fields

Note: Fields marked **bold** form the unique identifier for a record of the table.

Field name	Format	Description
<b><u>PATIENT</u></b>	character (or numeric if possible)	Code to identify patient (Cohort Patient ID)
<b><u>VS_ID</u></b>	character. see <a href="#">coding table</a>	Viral test
<b><u>VS_D</u></b>	yyyy-mm-dd	Measurement date
<b><u>VS_R</u></b>	numeric: <ul style="list-style-type: none"> <li>1 = Positive</li> <li>0 = Negative</li> <li>9 = Unknown/borderline</li> </ul>	Measurement result
<b><u>VS_V</u></b>	numeric	Measurement value (HCV-RNA & HBV-DNA only) (copies/ml)
<b><u>VS_U</u></b>	character. see <a href="#">coding table</a> for valid codings.	Measurement unit

### Additional fields

Field name	Format	Description
<b><u>VS_LL</u></b>	numeric	IF AVAILABLE, Lower limit of assay
<b><u>VS_UL</u></b>	numeric	IF AVAILABLE, Upper limit of assay
<b><u>VS_T</u></b>	character. see <a href="#">coding table</a> for valid codings.	IF AVAILABLE, type of ASSAY used for this measurement

### QA Checks

Table	Crosstable	Error Code	Description	Study specific	HICDEP?
AllTables	CrossTable	ATC001	any date in database after DEATH_D in tbILTFU		YES
AllTables	CrossTable	ATC002	any date in database after DROP_D in tbILTFU		YES
AllTables	CrossTable	ATC003	any date in database before BIRTH_D in tbIBAS		YES
AllTables	CrossTable	ATC004	any date in database in the future		YES
AllTables	CrossTable	ATC005	patients submitted previously who have been missed out		YES
AllTables	CrossTable	ATC006	Any fields not coded as coding lists on table definition		YES

tblLAB_VIRO	CrossTable	LVC001	Patient doesn't have a record in BAS		YES
tblLAB_VIRO	WithinTable	LVW002	Missing patient		YES
tblLAB_VIRO	WithinTable	LVW003	Missing VS_ID		YES
tblLAB_VIRO	WithinTable	LVW004	Missing VS_D		YES
tblLAB_VIRO	WithinTable	LVW005	Missing VS_R		YES
tblLAB_VIRO	WithinTable	LVW006	Missing VS_V		YES
tblLAB_VIRO	WithinTable	LVW007	Missing VS_U		YES
tblLAB_VIRO	WithinTable	LVW008	More or less than exactly 1 positive HIV test per patient		YES
tblLAB_VIRO	WithinTable	LVW009	More than 1 negative HIV test for a patient		YES
tblLAB_VIRO	WithinTable	LVW010	Date of negative test after date of positive test		YES
tblLAB_VIRO	WithinTable	LVW011	Duplicate records		YES



## tbILTFU - Death and drop-out

holds data in death and drop-out

### Core fields

Note: Fields marked **bold** form the unique identifier for a record of the table.

Field name	Format	Description
<u><a href="#">PATIENT</a></u>	character (or numeric if possible)	Code to identify patient (Cohort Patient ID)
<u><a href="#">DROP_Y</a></u>	numeric: <ul style="list-style-type: none"> <li>1 = Yes</li> <li>0 = No</li> </ul>	Has the patient DROPPED OUT?
<u><a href="#">DROP_D</a></u>	yyyy-mm-dd	IF YES, Date of Last Visit
<u><a href="#">DROP_RS</a></u>	numeric. see <a href="#">coding table</a> for valid codings.	IF YES, Reason for DROP
<u><a href="#">DEATH_Y</a></u>	numeric: <ul style="list-style-type: none"> <li>1 = Yes</li> <li>0 = No</li> </ul>	Has the patient died?
<u><a href="#">DEATH_D</a></u>	yyyy-mm-dd	Date of Death
<u><a href="#">SUD_DEATH_Y</a></u>	numeric: <ul style="list-style-type: none"> <li>1 = Yes</li> <li>0 = No</li> <li>9 = Unknown</li> </ul>	Sudden Death?
<u><a href="#">EXP_DEATH_Y</a></u>	numeric: <ul style="list-style-type: none"> <li>1 = Yes</li> <li>0 = No</li> <li>9 = Unknown</li> </ul>	Expected Death?
<u><a href="#">AUTOP_Y</a></u>	numeric: <ul style="list-style-type: none"> <li>1 = Yes</li> <li>0 = No</li> <li>9 = Unknown</li> </ul>	Was an autopsy Performed?
<u><a href="#">DEATH_R1</a></u>	character. see <a href="#">coding table</a> for valid codings.	Cause of death
<u><a href="#">DEATH_RC1</a></u>	character with codes: <ul style="list-style-type: none"> <li>I = Immediate cause</li> <li>U = Underlying cause/condition</li> <li>C = Contributing cause</li> <li>N = Not available</li> </ul>	Coding of causal relation of the code given in DEATH_R1 to the death
<u><a href="#">DEATH_R2</a></u>	character. see <a href="#">coding table</a> for valid codings.	Cause of death
<u><a href="#">DEATH_RC2</a></u>	character with codes: <ul style="list-style-type: none"> <li>I = Immediate cause</li> <li>U = Underlying cause/condition</li> <li>C = Contributing cause</li> <li>N = Not available</li> </ul>	Coding of causal relation of the code given in DEATH_R2 to the death

<i>DEATH_R3</i>	character. see <a href="#">coding table</a> for valid codings.	Cause of death
<i>DEATH_RC3</i>	character with codes: <ul style="list-style-type: none"> <li>• I = Immediate cause</li> <li>• U = Underlying cause/condition</li> <li>• C = Contributing cause</li> <li>• N = Not available</li> </ul>	Coding of causal relation of the code given in DEATH_R3 to the death
<i>DEATH_SOURCE</i>	character	Source of information for coding of death (e.g. CoDe within own cohort, CoDe from D:A:D, CoDe from ART-CC, etc.)

List of *DEATH\_R#* and *DEATH\_RC#* should be continued for as many reasons that are recorded.

The *DEATH\_RC#* fields should enable cohorts to transfer data in accordance with the [?Coding of Death project \(CoDe\)](#). You are welcome to contact the CoDe group for electronic sample forms for detailed collection of data used for the CoDe review process.

CoDe defines 1 immediate, 2 contributing and 1 underlying cause of death.

### Additional fields

Field name	Format	Description
<a href="#">ICD10_1</a>	character	Cause of death as ICD-10 if available
<i>ICD10_2</i>	character	Cause of death as ICD-10 if available
<i>ICD10_3<sup>1</sup></i>	character	Cause of death as ICD-10 if available
<a href="#">ICD9_1?</a>	character	Cause of death as ICD-9 if available
<i>ICD9_2</i>	character	Cause of death as ICD-9 if available
<i>ICD9_3<sup>1</sup></i>	character	Cause of death as ICD-9 if available
<a href="#">DEATH_OT</a>	character	Reason for death ? other - description
<a href="#">L_ALIVE</a>	yyyy-mm-dd	Last date known to be alive

<sup>1</sup>: List of *ICD10\_#* and *ICD9\_#* in place of or together with *DEATH\_R#* and together *DEATH\_RC#* and should be continued for as many reasons that are recorded.

CoDe defines 1 immediate, 2 contributing and 1 underlying cause of death.

### QA Checks

Table	Crosstable	Error Code	Description	Study specific	HICDEP?
AllTables	CrossTable	ATC001	any date in database after DEATH_D in tblLTFU		YES
AllTables	CrossTable	ATC002	any date in database after DROP_D in tblLTFU		YES
AllTables	CrossTable	ATC003	any date in database before BIRTH_D in tblBAS		YES
AllTables	CrossTable	ATC004	any date in database in the future		YES
AllTables	CrossTable	ATC005	patients submitted previously who have been missed out		YES

AllTables	CrossTable	ATC006	Any fields not coded as coding lists on table definition		YES
tbLTFU	WithinTable	LFW001	DROP_Y and DEATH_Y both non-null		YES
tbLTFU	WithinTable	LFW002	DEATH_Y and DROP_RS both non-null		YES
tbLTFU	WithinTable	LFW003	Any of DEATH_Rx or DEATH_RCx non-null but DEATH_Y=0		YES
tbLTFU	WithinTable	LFW004	R2/RC2 non-null but R1/RC1 null		YES
tbLTFU	WithinTable	LFW005	R3/RC3 non-null but R2/RC2 null		YES
tbLTFU	WithinTable	LFW006	Duplicate patients		YES
tbLTFU	WithinTable	LFW007	Missing DROP_Y		YES
tbLTFU	WithinTable	LFW008	Missing DROP_D if DROP_Y=1		YES
tbLTFU	WithinTable	LFW009	Missing DROP_RS if DROP_Y=1		YES
tbLTFU	WithinTable	LFW010	Missing DEATH_Y		YES
tbLTFU	WithinTable	LFW011	Missing DEATH_D if DEATH_Y=1		YES
tbLTFU	WithinTable	LFW012	DEATH_D non null but DEATH_Y=0		YES
tbLTFU	CrossTable	LFC001	Patient not found in tbIBAS		YES
tbLTFU	CrossTable	LFC002	patient in tbIBAS hasn't got a record in tbLTFU		YES
tbLTFU	CrossTable	LFC003	tbIBAS has AIDS=0 but DEATHRx =8.1		YES

## tbIMED - Other medication

holds type, start and stop dates for other medication/treatments.

### Core fields

Note: Fields marked **bold** form the unique identifier for a record of the table.

Field name	Format	Description
<b><u>PATIENT</u></b>	character (or numeric if possible)	Code to identify patient (Cohort Patient ID)
<b><u>MED_ID</u></b>	character. see <a href="#">coding table</a> for valid codings.	Code representing the treatment
<b><u>MED_SD</u></b>	yyyy-mm-dd	Date of Initiation of Treatment
<b><u>MED_ED</u></b>	yyyy-mm-dd	Date of stopping treatment
<b><u>MED_RS</u></b>	character. see <a href="#">coding table</a> for valid codings.	reason for stopping treatment

### Additional fields

Please see [tblART - Antiretroviral treatment](#).

### QA Checks

Table	Crosstable	Error Code	Description	Study specific	HICDEP?
AllTables	CrossTable	ATC001	any date in database after DEATH_D in tblLTFU		YES
AllTables	CrossTable	ATC002	any date in database after DROP_D in tblLTFU		YES
AllTables	CrossTable	ATC003	any date in database before BIRTH_D in tblBAS		YES
AllTables	CrossTable	ATC004	any date in database in the future		YES
AllTables	CrossTable	ATC005	patients submitted previously who have been missed out		YES
AllTables	CrossTable	ATC006	Any fields not coded as coding lists on table definition		YES
tbIMED	WithinTable	MW001	Duplicate records		YES
tbIMED	WithinTable	MW002	MED_ONG=0 and MED_ED null	EPPICC	NO
tbIMED	WithinTable	MW003	MED_ONG=1 and MED_ED non-null	EPPICC	NO
tbIMED	WithinTable	MW004	Missing MED_ID		YES
tbIMED	WithinTable	MW005	Missing MED_SD		YES
tbIMED	WithinTable	MW006	Mising MED_ED		YES
tbIMED	WithinTable	MW007	Overlapping periods of same drug		YES
tbIMED	WithinTable	MW008	MED_ED < MED_SD		YES

tbIMED	WithinTable	MW009	MED_RS not null, but end date NULL		YES
tbIMED	CrossTable	MC001	Patient has no records in tbIBAS		YES

## tbINWBORN - Newborn

Please also read the [notes on pregnancy tables](#).

### Core fields

Note: Fields marked **bold** form the unique identifier for a record of the table.

Field name	Format	Description
<b>CHILD_ID</b>	Character (or numeric if possible)	Patient ID of the child
BRFEED_SD	yyyy-mm-dd	Breastfeeding, start date.
BRFEED_ED	yyyy-mm-dd	Breastfeeding, end date.
FAT_ETH	See <a href="#">Coding Table</a> .	Ethnicity of father
APGAR_1	numeric	1 <sup>st</sup> APGAR score
APGARM_1	numeric	minute at which the 1 <sup>st</sup> APGAR test was performed
APGAR_2	numeric	2 <sup>nd</sup> APGAR score
APGARM_2	numeric	minute at which the 2 <sup>nd</sup> APGAR test was performed
APGAR_3	numeric	3 <sup>rd</sup> APGAR score
APGARM_3	numeric	minute at which the 3 <sup>rd</sup> APGAR test was performed
ICU_Y	numeric: <ul style="list-style-type: none"> <li>• 1 = Yes</li> <li>• 0 = No</li> <li>• 9 = Unknown</li> </ul>	Referral to intensive/intermediate care unit?
ICU_S	character	if yes, specify reason
ICU_D	yyyy-mm-dd	Date of entry to intensive/intermediate care unit
ABNORM_Y	numeric: <ul style="list-style-type: none"> <li>• 1 = Yes</li> <li>• 0 = No</li> <li>• 9 = Unknown</li> </ul>	did any abnormalities occur? (if yes, recorded in <a href="#">tbINWBORN_ABNORM</a> )

### QA Checks

Table	Crosstable	Error Code	Description	Study specific	HICDEP?
AllTables	CrossTable	ATC001	any date in database after DEATH_D in tbILTFU		YES
AllTables	CrossTable	ATC002	any date in database after DROP_D in tbILTFU		YES
AllTables	CrossTable	ATC003	any date in database before BIRTH_D in tbIBAS		YES
AllTables	CrossTable	ATC004	any date in database in the future		YES

AllTables	CrossTable	ATC005	patients submitted previously who have been missed out		YES
AllTables	CrossTable	ATC006	Any fields not coded as coding lists on table definition		YES
tblNEWBORN	CrossTable	NC001	CHILD_ID doesn't exist in tblDELIVERY_CHILD		YES
tblNEWBORN	CrossTable	NC002	ABNORM_Y = 1, yet no records in tblNEWBORN_ABNORM		YES
tblNEWBORN	CrossTable	NC003	ABNORM_Y = 0 or 9, yet records in tblNEWBORN_ABNORM		YES
tblNEWBORN	WithinTable	NW001	BRFEED_SD>BRFEED_ED		YES
tblNEWBORN	WithinTable	NW002	APGARM_x's out of order (e.g APGARM_3 < APGARM_2)		YES
tblNEWBORN	WithinTable	NW003	ICU_Y=1, but ICU_S or ICS_D null		YES
tblNEWBORN	WithinTable	NW004	ICU_Y=0 or 9, but ICU_S or ICU_D non null		YES

## tbNEWBORN\_ABNORM - Abnormalities

Abnormalities in newborns are recorded here, one abnormality per row. The absence of a record is to be interpreted as "unknown whether the abnormality existed" since most cohorts only record positive events.

Please also read the [notes on pregnancy tables](#).

### Core fields

*Note:* Fields marked **bold** form the unique identifier for a record of the table.

Field name	Format	Description
<b>CHILD_ID</b>	Character (or numeric if possible)	Patient ID of the child
<b>ABNORM_T</b>	character: <ul style="list-style-type: none"> <li>• 1 = Birth defect(s) (detectable in physical examination including skin abnormalities)</li> <li>• 2 = Congenital infection(s)</li> <li>• 3 = Drug withdrawal syndrome</li> <li>• 4 . 1 = Neurological disorder(s): abnormal reflexes</li> <li>• 4 . 2 = Neurological disorder(s): abnormal motility</li> <li>• 4 . 3 = Neurological disorder(s): abnormal tonus</li> <li>• 90 = Other health problems</li> </ul>	type of abnormality
<b>ABNORM_S</b>	character	further specification of the abnormality

### QA Checks

Table	Crosstable	Error Code	Description	Study specific	HICDEP?
AllTables	CrossTable	ATC001	any date in database after DEATH_D in tblLTFU		YES
AllTables	CrossTable	ATC002	any date in database after DROP_D in tblLTFU		YES
AllTables	CrossTable	ATC003	any date in database before BIRTH_D in tblBAS		YES
AllTables	CrossTable	ATC004	any date in database in the future		YES
AllTables	CrossTable	ATC005	patients submitted previously who have been missed out		YES
AllTables	CrossTable	ATC006	Any fields not coded as coding lists on table definition		YES
tbNEWBORN_ABNORM	CrossTable	NAC001	CHILD_ID doesn't exist in NEWBORN		YES
tbNEWBORN_ABNORM	WithinTable	NAW001	ABNORM_T=90 but ABNORM_S null		YES



## tbOVERLAP - Cross-cohort identification

holds information on the patient's participation in other cohorts

### Core fields

Note: Fields marked **bold** form the unique identifier for a record of the table.

Field name	Format	Description
<u>PATIENT</u>	character (or numeric if possible)	Code to identify patient (Cohort Patient ID)
<u>COHORT</u>	character	Code/name of the cohort
<u>PAT_OTH</u>	character	Unique patient identifier in other cohorts
<u>COH_OTH</u>	character	Name of the cohort

Patients of an "original"-cohort who also participate in a "super"-cohort should be analysed within the "original"-cohort only. To suppress these patients from the datasets of the "super"-cohorts the identifier used in the "super"-cohort is needed. It is suggested that "original"-cohorts report id's from the "super"-cohorts, since the "super"-cohorts might not even know the other ID's. Often this information is only available at centre level.

A record should be present for each cohort that the patient is participating in (apart from it's own "original"-cohort).

### QA Checks

Table	Crosstable	Error Code	Description	Study specific	HICDEP?
AllTables	CrossTable	ATC001	any date in database after DEATH_D in tblLTFU		YES
AllTables	CrossTable	ATC002	any date in database after DROP_D in tblLTFU		YES
AllTables	CrossTable	ATC003	any date in database before BIRTH_D in tblBAS		YES
AllTables	CrossTable	ATC004	any date in database in the future		YES
AllTables	CrossTable	ATC005	patients submitted previously who have been missed out		YES
AllTables	CrossTable	ATC006	Any fields not coded as coding lists on table definition		YES
tbOVERLAP	WithinTable	OW001	Invalid other cohort		YES
tbOVERLAP	CrossTable	OC001	PATIENT not found in tblBAS for that cohort		YES
tbOVERLAP	CrossTable	OC002	PAT_OTH not found in tblBAS for that overlapping cohort		YES

## tbIPREG - Pregnancy

Please also read the [notes on pregnancy tables](#).

### Core fields

Note: Fields marked **bold** form the unique identifier for a record of the table.

Field name	Format	Description
<b>MOTHER_ID</b>	Character (or numeric if possible)	Patient ID of mother of the child
<b>PREG_SEQ</b>	numeric	Sequence number of the pregnancy for the specified mother
<b>MENS_D</b>	yyyy-mm-dd	Date of last menstrual period (If date not known exactly please give approximated date)
<b>CONCEPT</b>	character: <ul style="list-style-type: none"> <li>1=Natural</li> <li>2=Infertility treatment unspecified</li> <li>2 . 1=IVF (In Vitro Fertilisation)</li> <li>2 . 2=ICSI (IntraCyttoplasmic Sperm Injection)</li> <li>2 . 3=Ovulation induction</li> <li>3=Artificial insemination</li> <li>4=Self insemination</li> <li>9=Unknown</li> </ul>	Conception
<b>ANC_D</b>	yyyy-mm-dd	Date of first antenatal care contact
<b>INPREG_Y</b>	character: <ul style="list-style-type: none"> <li>1=Yes</li> <li>2=No: ectopic</li> <li>3=No: missed abortion</li> <li>4=No: death in utero (IUFT)</li> </ul>	At first gynaecological visit: intact intrauterine pregnancy?
<b>INHIST_Y</b>	numeric: <ul style="list-style-type: none"> <li>0=No</li> <li>1=Yes</li> <li>9=Unknown</li> </ul>	If no in INPREG_Y was a histological investigation of reason made.
<b>INHIST_S</b>	character	If yes, please specify reason
<b>INV_PROC</b>	<ul style="list-style-type: none"> <li>0=No</li> <li>1=Yes, chorionic villus sampling (CVS)</li> <li>2=Cordocentesis</li> <li>3=Amniocentesis</li> <li>9=Unknown</li> </ul>	Invasive procedure
<b>KARYO_T</b>	<ul style="list-style-type: none"> <li>0=Not done</li> <li>1=Normal</li> <li>2=Abnormal</li> <li>9=Unknown</li> </ul>	Karyotype
<b>KARYO_A</b>	character	If abnormal, please specify

<i>CHORIO</i>	character: <ul style="list-style-type: none"> <li>• 1=Monochorionic</li> <li>• 2=Dichorionic</li> <li>• 3=Trichorionic</li> <li>• etc.</li> <li>• 9=Unknown</li> </ul>	For multiple pregnancies
<i>ULTR_1</i>	character: <ul style="list-style-type: none"> <li>• 0=No</li> <li>• 1=Yes, normal</li> <li>• 2=Yes, abnormal</li> <li>• 9=Unknown</li> </ul>	Ultrasound 1. trimester
<i>ULTR_A_1</i>	character	If abnormal, please specify
<i>ULTR_2</i>	character: <ul style="list-style-type: none"> <li>• 0=No</li> <li>• 1=Yes, normal</li> <li>• 2=Yes, abnormal</li> <li>• 9=Unknown</li> </ul>	Ultrasound 2. trimester
<i>ULTR_A_2</i>	character	If abnormal, please specify
<i>PROB_Y</i>	numeric: <ul style="list-style-type: none"> <li>• 1 = Yes</li> <li>• 0 = No</li> <li>• 9 = Unknown</li> </ul>	did any obstetrical problems occur? (if yes, recorded in <a href="#">tblPREG_OBS</a> )

### QA Checks

Table	Crosstable	Error Code	Description	Study specific	HICDEP?
AllTables	CrossTable	ATC001	any date in database after DEATH_D in tblLTFU		YES
AllTables	CrossTable	ATC002	any date in database after DROP_D in tblLTFU		YES
AllTables	CrossTable	ATC003	any date in database before BIRTH_D in tblBAS		YES
AllTables	CrossTable	ATC004	any date in database in the future		YES
AllTables	CrossTable	ATC005	patients submitted previously who have been missed out		YES
AllTables	CrossTable	ATC006	Any fields not coded as coding lists on table definition		YES
tblPREG	WithinTable	PW001	ANC_D		YES
tblPREG	WithinTable	PW002	INPREG_Y=2,3,4 but INHIST_Y null		YES

tblPREG	WithinTable	PW003	INPREG_Y=1 but INHIST_Y non null		YES
tblPREG	WithinTable	PW004	INHIST_Y=1 but INHIST_S null		YES
tblPREG	WithinTable	PW004	INHIST_Y=0 or 9 but INHIST_S non null		YES
tblPREG	WithinTable	PW005	KARYO_T=2 but KARYO_A null		YES
tblPREG	WithinTable	PW006	KARYO_T<>2 but KARYO_A non null		YES
tblPREG	WithinTable	PW007	ULTRA_x =2, but ULTR_A_x null		YES
tblPREG	WithinTable	PW008	ULTRA_x <>2, but ULTR_A_x non null		YES
tblPREG	CrossTable	PC001	PROB_Y = 1 but no records in tblPREG_OBS		YES
tblPREG	CrossTable	PC002	PROB_Y = 0 or 9 but records in tblPREG_OBS		YES

## tbIPREG\_OBS - Obstetrical problems during pregnancy

This table describes problems **during a pregnancy**. Abnormalities in newborns are recorded in [tbINEWBORN\\_ABNORM](#) instead.

Please also read the [notes on pregnancy tables](#).

### Core fields

*Note:* Fields marked **bold** form the unique identifier for a record of the table.

Field name	Format	Description
<b>MOTHER_ID</b>	Character (or numeric if possible)	patient id of mother
<b>PREG_SEQ</b>	numeric	Sequence number of the pregnancy for the specified mother
<b>PROB_T</b>	character: <ul style="list-style-type: none"> <li>• 1 = Preterm contractions</li> <li>• 2 = Shortened cervix</li> <li>• 3 = Preterm rupture of membranes</li> <li>• 4 = Antepartum bleeding</li> <li>• 5 = Intrauterine growth retardation (IUGR)</li> <li>• 6 = Preeclampsia/HELLP</li> <li>• 7 = Hypertension</li> <li>• 8 = Gestational diabetes (unspecified type)</li> <li>• 8.1 = Gestational diabetes (Diet)</li> <li>• 8.2 = Gestational diabetes (Insulin)</li> <li>• 9 = Placental abruption</li> <li>• 10 = Placenta praevia</li> <li>• 99 = Other, specify in <b>PROB_S</b></li> </ul>	type of obstetrical problem
<b>PROB_S</b>	character.	description of other (99) obstetrical problem
<b>CERVIX_S</b>	numeric (mm)	In case of shortened cervix, the length of the cervix in millimeters.

### QA Checks

Table	Crosstable	Error Code	Description	Study specific	HICDEP?
AllTables	CrossTable	ATC001	any date in database after DEATH_D in tbILTFU		YES
AllTables	CrossTable	ATC002	any date in database after DROP_D in tbILTFU		YES
AllTables	CrossTable	ATC003	any date in database before BIRTH_D in tbIBAS		YES
AllTables	CrossTable	ATC004	any date in database in the future		YES
AllTables	CrossTable	ATC005	patients submitted previously who have been missed out		YES

AllTables	CrossTable	ATC006	Any fields not coded as coding lists on table definition		YES
tblPREG_OBS	CrossTable	POC001	MOTHER_ID+PREG_SEQ doesn't exist in tblPREG		YES
tblPREG_OBS	WithinTable	POW001	PROB_T=99 but PROB_S null		YES
tblPREG_OBS	WithinTable	POW002	PROB_T<>99 but PROB_S non null		YES
tblPREG_OBS	WithinTable	POW003	PROB_T=2 but CERVIX_S null		YES
tblPREG_OBS	WithinTable	POW003	PROB_T<>2 but CERVIX_S non null		YES

## tbIPREG\_OUT - Pregnancy outcome

This table describes outcomes of pregnancies described in [tbIPREG](#).

Please also read the [notes on pregnancy tables](#).

### Core fields

Note: Fields marked **bold** form the unique identifier for a record of the table.

Field name	Format	Description
<b>MOTHER_ID</b>	Character (or numeric if possible)	Patient ID of mother of the child
<b>PREG_SEQ</b>	numeric	Sequence number of the pregnancy for the specified mother
<b>CHILD_ID</b>	Character (or numeric if possible)	Patient ID of the child
<b>OUTCOM</b>	character: <ul style="list-style-type: none"> <li>1=Born alive, HIV negative</li> <li>2=Born alive, HIV positive</li> <li>3=Born alive, unknown HIV status</li> <li>10=Stillborn</li> <li>11=Spontaneous miscarriage</li> <li>20=Termination: surgical</li> <li>21=Termination: medication</li> </ul>	Pregnancy outcome
<b>OUTCOM_R</b>	<ul style="list-style-type: none"> <li>1=Fetus with malformation</li> <li>2=Dead fetus</li> <li>3=Unwanted pregnancy</li> <li>8=Other</li> <li>9=Unknown</li> </ul>	Reason for termination
<b>OUTCOM_D</b>	yyyy-mm-dd	Date of birth or termination of pregnancy
<b>B_GAGEW</b>	numeric	Gestational age in complete weeks at birth or termination
<b>B_GAGED</b>	numeric	Gestational age in days in addition to weeks at birth or termination

### QA Checks

Table	Crosstable	Error Code	Description	Study specific	HICDEP?
AllTables	CrossTable	ATC001	any date in database after DEATH_D in tbILTFU		YES
AllTables	CrossTable	ATC002	any date in database after DROP_D in tbILTFU		YES
AllTables	CrossTable	ATC003	any date in database before BIRTH_D in tbIBAS		YES
AllTables	CrossTable	ATC004	any date in database in the future		YES

AllTables	CrossTable	ATC005	patients submitted previously who have been missed out		YES
AllTables	CrossTable	ATC006	Any fields not coded as coding lists on table definition		YES
tblPREG_OUT	CrossTable	PTC001	MOTHER_ID+PREG_SEQ doesn't exist in tblPREG		YES
tblPREG_OUT	CrossTable	PTC002	OUTCOM=1,2 or 3 and CHILD_ID doesn't exist in tblNEWBORN or tblDELIVERYCHILD		YES
tblPREG_OUT	CrossTable	PTC003	OUTCOM=10,11,20 or 21 and CHILD_ID exists in tblNEWBORN or tblDELIVERYCHILD		YES
tblPREG_OUT	WithinTable	PTW001	OUTCOME=20 or 21 and OUTCOM_R null		YES
tblPREG_OUT	WithinTable	PTW002	OUTCOME not 20 or 21 and OUTCOM_R non null		YES



## tbIREFILL - Prescription refill data

holds drug, refill date and supply of prescriptions

### Core fields

Note: Fields marked **bold** form the unique identifier for a record of the table.

Field name	Format	Description
<u><a href="#">PATIENT</a></u>	character (or numeric if possible)	Code to identify patient (Cohort Patient ID)
<u><a href="#">REFILL_D</a></u>	yyyy-mm-dd	Date of the prescription refill
<u><a href="#">DRUG_ID</a></u>	character. see <a href="#">coding table</a> for valid codings.	Drug which was prescribed
<u><a href="#">SUPPLY</a></u>	numeric	How many days supply of the drug was supplied

### QA Checks

Table	Crosstable	Error Code	Description	Study specific	HICDEP?
AllTables	CrossTable	ATC001	any date in database after DEATH_D in tbILTFU		YES
AllTables	CrossTable	ATC002	any date in database after DROP_D in tbILTFU		YES
AllTables	CrossTable	ATC003	any date in database before BIRTH_D in tbIBAS		YES
AllTables	CrossTable	ATC004	any date in database in the future		YES
AllTables	CrossTable	ATC005	patients submitted previously who have been missed out		YES
AllTables	CrossTable	ATC006	Any fields not coded as coding lists on table definition		YES
tbIREFILL	CrossTable	PRC001	Patient doesn't have a record in BAS		YES
tbIREFILL	WithinTable	PRW002	Missing PATIENT		YES
tbIREFILL	WithinTable	PRW003	Missing REFILL_D		YES
tbIREFILL	WithinTable	PRW004	Missing DRUG_ID		YES
tbIREFILL	WithinTable	PRW005	Missing SUPPLY		YES
tbIREFILL	WithinTable	PRW006	SUPPLY < 1		YES

## tbISAMPLES - Blood Samples

This table contains information on the storage of blood, urine and other samples stored in a laboratory.

### Core fields

Note: Fields marked **bold** form the unique identifier for a record of the table.

Field name	Format	Description
<b>PATIENT</b>	character (or numeric if possible)	patient cohort identifier
<b>SAMP_LAB_D</b>	yyyy-mm-dd	date when the sample was taken
<b>SAMP_TYPE</b>	character: <ul style="list-style-type: none"> <li>• BS = blood serum</li> <li>• BP = blood plasma</li> <li>• C = viable cells</li> <li>• D = cell pellet (DNA)</li> <li>• S = semen</li> <li>• OTH:x = other sample type x (none of the above)</li> </ul>	type of the sample
<b>SAMP_ID</b>	character	identification symbol allowing the localization of the sample in the laboratory
<b>SAMP_LAB</b>	character	laboratory where the samples are stored
<b>SAMP_FREEZE_D</b>	yyyy-mm-dd	date when the sample was frozen
<b>SAMP_FREEZE_T</b>	hh:mm	time when the sample was frozen
<b>SAMP_ALIQ_NO</b>	numeric	number of aliquots available
<b>SAMP_ALIQ_SIZE</b>	numeric	size of the aliquot: <ul style="list-style-type: none"> <li>• in ml for serum, plasma and cell pellet aliquots</li> <li>• in millions of cells for viable cell aliquots</li> </ul>
<b>SAMP_ALIQ_U</b>	character: <ul style="list-style-type: none"> <li>• 0 = millions of cells</li> <li>• 1 = ml</li> </ul>	unit of measurement for the SAMP_ALIQ_SIZE value

### Additional fields

Field name	Format	Description
<b>SAMP_LAB_T</b>	hh:mm	time when the sample was taken
<b>SAMP_TEMP</b>	numeric	temperature of the storage unit containing the samples (in °C)
<b>SAMP_DEFROST</b>	numeric: <ul style="list-style-type: none"> <li>• 1 = Yes</li> <li>• 0 = No</li> <li>• 9 = Unknown</li> </ul>	have the samples already been defrosted?

### [QA Checks](#)

Table	Crosstable	Error Code	Description	Study specific	HICDEP?
AllTables	CrossTable	ATC001	any date in database after DEATH_D in tblLTFU		YES
AllTables	CrossTable	ATC002	any date in database after DROP_D in tblLTFU		YES
AllTables	CrossTable	ATC003	any date in database before BIRTH_D in tblBAS		YES
AllTables	CrossTable	ATC004	any date in database in the future		YES
AllTables	CrossTable	ATC005	patients submitted previously who have been missed out		YES
AllTables	CrossTable	ATC006	Any fields not coded as coding lists on table definition		YES

## tbIVIS - Basic follow-up/visit related data

holds visit related information such as weight, wasting, smoking, occupational status etc.

### Core fields

Note: Fields marked **bold** form the unique identifier for a record of the table.

Field name	Format	Description
<a href="#">PATIENT</a>	character (or numeric if possible)	Code to identify patient (Cohort Patient ID)
<a href="#">VIS_D</a>	yyyy-mm-dd	Date of patient visit
<a href="#">WEIGH</a>	numeric (metric: kg): 999 = Unknown	Weight of patient at visit
<a href="#">GAIN_Y</a>	numeric: <ul style="list-style-type: none"> <li>• 1 = Yes</li> <li>• 0 = No</li> <li>• 9 = Unknown</li> </ul>	Is the patient gaining fat in the abdomen, neck, breast or other defined locations?
<a href="#">LOSS_Y</a>	numeric: <ul style="list-style-type: none"> <li>• 1 = Yes</li> <li>• 0 = No</li> <li>• 9 = Unknown</li> </ul>	Is the patient experiencing loss of fat from extremities, buttocks or face?

Depending on the collaboration this data might be collected in intervals of a year, e.g. from July last to July this year. In that case all visit dates or a fixed number of visit dates for that period should be gathered, if the patient did not have a visit in the defined period, a record with the PATIENT id and empty fields for VIS\_D etc. should be included.

### Additional fields

Field name	Format	Description
<a href="#">CDC_STAGE</a>	character. see <a href="#">coding table</a> for valid codings.	Clinical CDC stage at time of visit?
<a href="#">WHO_STAGE</a>	numeric. <ul style="list-style-type: none"> <li>• 1</li> <li>• 2</li> <li>• 3</li> <li>• 4</li> <li>• 9 = Unknown</li> </ul>	Clinical WHO stage at time of visit?

The following fields are meant to be used for **adolescents and adults**.

Field name	Format	Description
<a href="#">EMPLOY</a>	numeric. see <a href="#">coding table</a> for valid codings.	What is the patient's current situation regarding labour?
<a href="#">CONTRACT</a>	numeric. see <a href="#">coding table</a> for valid codings.	If the patient is an employee, what is the type of the patient's employment contract?
<a href="#">SMOKING_Y</a>	numeric: <ul style="list-style-type: none"> <li>• 1 = Yes</li> <li>• 0 = No</li> <li>• 9 = Unknown</li> </ul>	Is the patient currently a smoker?

<a href="#">PREG_Y</a>	numeric: <ul style="list-style-type: none"> <li>• 1 = Yes</li> <li>• 0 = No</li> <li>• 9 = Unknown</li> </ul>	Is the patient currently pregnant? If possible, provide additional details in <a href="#">tblPREG</a>
------------------------	---	---

The following fields are meant to be used **for children and infants**.

The following fields are meant to be used **for infants**:

Field name	Format	Description
<a href="#">HEIGH_P</a>	numeric	Height/length of patient at visit in percentiles
<a href="#">WEIGH_P</a>	numeric	Weight of patient at visit in percentiles
<a href="#">HEADC</a>	numeric	Head circumference measured in millimeters (mm)
<a href="#">HEADC_P</a>	numeric	Head circumference in percentiles
<a href="#">BREASTF_Y</a>	numeric: <ul style="list-style-type: none"> <li>• 1 = Yes</li> <li>• 0 = No</li> <li>• 9 = Unknown</li> </ul>	Currently Breastfeeding?

### QA Checks

Table	Crosstable	Error Code	Description	Study specific	HICDEP?
AllTables	CrossTable	ATC001	any date in database after DEATH_D in tblLTFU		YES
AllTables	CrossTable	ATC002	any date in database after DROP_D in tblLTFU		YES
AllTables	CrossTable	ATC003	any date in database before BIRTH_D in tblBAS		YES
AllTables	CrossTable	ATC004	any date in database in the future		YES
AllTables	CrossTable	ATC005	patients submitted previously who have been missed out		YES
AllTables	CrossTable	ATC006	Any fields not coded as coding lists on table definition		YES
tblVIS	WithinTable	VW001	Duplicate records, same VIS_D		YES
tblVIS	WithinTable	VW002	Height decreasing over time		YES
tblVIS	WithinTable	VW003	Height out of acceptable range		YES
tblVIS	WithinTable	VW004	Weight out of acceptable range		YES

tblVIS	CrossTable	VC001	patient has no record in BAS table		YES
tblVIS	CrossTable	VC002	No weights within 3 mths of starting FPV/DRV	EPPICC	YES

## Considerations for data management

### Uniqueness of *PATIENT ID*

Each table in this document is shown with the *PATIENT* field as part of the unique identifier for each record, in many cases it might be necessary to specify both *PATIENT* and *COHORT* as identifier since the *PATIENT* id might not be unique across cohorts, this should however be implemented only if needed.

Another solution, which might be worth considering when building the final dataset for the analysis, is to concatenate the patient id and the cohort id into a single unique new patient id.

### Duplicate records

Transposition of the flat format data, where there is one record per patient, into a normalised structure that has multiple rows per patient requires that the combination of *PATIENT*, *TYPE\_ID* and *LAB\_DATE* is unique for each row in the table.

The normalised structure and the relational model does not allow for duplicate records like in the following table to exist:

PATIENT	TYPE_ID	LAB_DATE	LAB_VAL	LAB_UNIT
999999	1	2000-01-01	15	U/l
999999	1	2000-01-01	15	U/l

The typical solution to this is to generate an auto-incremented value ? *RECORD\_ID* - for each record in the table:

RECORD_ID	PATIENT	TYPE_ID	LAB_DATE	LAB_VAL	LAB_UNIT
1	999999	1	2000-01-01	15	U/l
2	999999	1	2000-01-01	15	U/l

But it can, and should be dealt with in the design of the study. The presence of duplicate records like in the example shown above might not make any sense; if they did occur it then would be on account of a recording error. If, however, it was important to keep records that show measurements taken the same day but at different times, the format of the *LAB\_DATE* shouldn't be date (YYYY-MM-DD) but instead should be date-time (YYYY-MM-DD hh:mm:ss):

PATIENT	TYPE_ID	LAB_DATE	LAB_VAL	LAB_UNIT
999999	1	2000-01-01 10:00:00	15	U/l
999999	1	2000-01-01 14:00:00	15	U/l

Different and more sophisticated methods to make records unique will be presented later in this document under the definition of the LAB table.

## Considerations for using the format to create a database

### Administrative fields

Sometimes it might be needed to have a fixed value that shows from which visit or merger a value originates, this does not only apply to the VIS table but could be applied to all tables. This however does depend on the nature of the database and needs for data management, the field below should be considered an administrative support field for data management.

VISIT
Visit number
Numeric: 0 = Baseline Visit 1 = First follow up visit 2 = Second follow up visit etc.

Often the above field is used for clinical trials databases where there is a need to associate the data directly with a given week's follow-up. Codes could then be the week number e.g. 4, 12, 24 etc or ?1 for screening/randomisation and 0 for baseline visits.

In some cases it might be useful to have a separate field that defines the correct order of the periods. This becomes important where several dates are incomplete (unknown days, unknown months and possibly unknown years). The ordering by date would then not be correct.

One solution to this is use a *PERI\_ID* field that numbers the periods from the 1<sup>st</sup> until N<sup>th</sup> usage:

PERI_ID
Period of usage (1 <sup>st</sup> , 2 <sup>nd</sup> , 3 <sup>rd</sup> etc.)
Numeric

However this is an optional field that for most cohorts may not be needed. It also requires additional maintenance to keep it updated.

For databases that work with double data entry, such as most clinical databases, it becomes necessary to make each data entry unique and backwards traceable. For this to work a field like the above would have to be part of the primary key of each table that requires double data entry.

ENTRY_ID
Number of data entry
Numeric: 1 = first data entry 2 = second data entry 3 = comparison of 1 <sup>st</sup> and 2 <sup>nd</sup> data entry 4 = final approved record including corrections

With respect to performance, it might also be a good design to have 3 copies of each table, one to hold the data while being entered and compared, one for the two data entries to be archived into once a final record has been approved and a table holding the final and approved values. This way it is avoided that queries will have to work on checking for *ENTRY\_ID* = 4 and to select amongst a table holding 3 times the almost same data.

As part of an audit trail in a database a time stamp field could be added for each record to fix the exact time when the record last was inserted or updated. Along with the time stamp name of the user who entered or altered data can be recorded.

T_STAMP	USER_LOG
Date and time of data entry	Username of user that last inserted or updated data
yyyy-mm-dd hh:mm:ss	character

Often it's necessary to keep a log of user action in a separate table. The above suggestion will only be valid for inserts and updates, and only be valid for the most recent action performed.

To record a complete audit trail a logging facility must be implemented. In most database management systems this is done using triggers on the tables. For any insert, update or delete actions performed on the data, the user, time stamp, old value and new value are recorded in the logging table.



The *T\_STAMP* field could also include information about which time zone is relevant for data entry. Depending on database requirements this might in fact be mandatory if the [FDA's 21 CFR part 11](#) on electronic records and signatures applies.

## Further normalisation

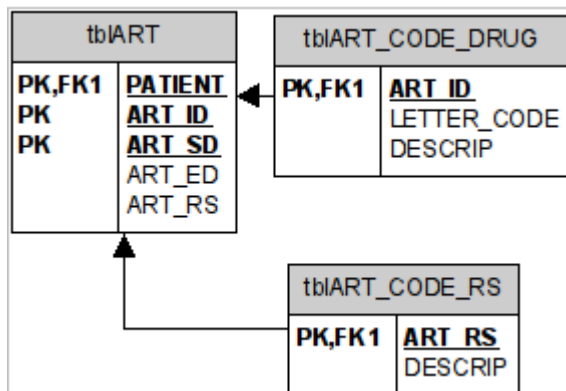
Depending on performance considerations it might be worth looking at how data are queried for data entry and data analysis. A smaller *tbIBAS* table might increase performance: Since processing a smaller table is always faster than processing a larger table, one could put drop-out, death, birthday, date of aids diagnosis, etc. into separate tables and keep the core patient list in a separate master table

But if the database is used e.g. for BMI calculations directly on the running database, performance might be enhanced by keeping the patient list and the height together in the same table so that a query involves 2 tables (*tbIBAS* and *tbIVIS*) rather than perhaps 3 or more.

Another consideration is space. Although it may not be much of an issue, it will be possible to minimise the actual size of the database by putting fields that may be empty for most patients, like death information, into a separate table in a 0-1 to 1 relation to the master table.

## Lookup tables

In a running database the *#\_ID* fields could be implemented as a foreign key to a linked lookup table containing all possible codes and their corresponding definitions in a text string.



This setup not only enables integrity of the data, but also defines the domain<sup>1</sup> for the *#\_ID* values and enables data to both become human readable and easily recoded<sup>2</sup>.

An important note on lookup tables is that the performance on a large database can be slowed significantly by using character based keys to link lookup tables with the primary table as it is presented in this document. A work around is to use numeric value for the codes.

<sup>1</sup>: Domain is a term in the definition of the relational database model that defines a set of allowed values for a given set of fields (attributes), the mixing of different domains is not allowed in order to preserve the integrity of a relational and normalised model.

<sup>2</sup>: Easily recoded permanently if the relation is specified as cascade on update or recoded dynamic by selecting a different column from the lookup table when querying the data through SQL

## Performance

As already outlined in the above section, there are also performance issues that may have to be considered.

When using the suggested data types presented in this document for a database implementation, it may be worth looking at the actual data at hand when defining the final data types. The aim of this document is to present a format that will work between cohorts with rather different setups.

If it is at all possible in many cases there may be a large performance gain by using numeric instead of character fields. Character fields have been suggested here for, amongst others, the *PATIENT* field. If the *PATIENT* id is purely numeric it's worth using a numeric data type since it always faster for querying than a character field.

Whenever the field has to be character, make sure that only the needed amount of space is assigned for the field length; there is no need to assign 50 characters of memory if the field in fact only stores a 3-letter code.

## QA checks

This page lists all QA checks currently defined for [HICDEP 1.80](#). The attached `QA.csv` file also serves as data source for the QA sections in the articles describing the individual tables.

Checks added in so far have been collated from 3 sources:

- CASCADE
- EPPICC FPV/DRV mergers
- PENTA Eurocoord merger

Checks on a table's additional fields have not been included at this stage

Only a few checks for non-HICDEP tables, non-HICDEP fields in HICDEP tables and Study-specific checks have been included at this stage for demo/testing purposes.

## Error Codes

The error codes allow to unambiguously reference a certain problem with a given data set. When merging data from different cohorts, the data manager may report problems for instance as follows:

```
[ATC006] in tblART/ART_ID: there is no drug named 'XYZ' in the coding list
```

**Note:** the use of the error code alone is not sufficient to locate a problem since some of them are fairly generic.

Used like above, error codes can help to clearly communicate issues. This holds especially if the QA requirements are being checked by an automatic system.

## QA checks listed nowhere else

These are QA checks which are either not directly applicable to a specific table or to tables which are not (yet) part of HICDEP.

Table	Crosstable	Error Code	Description	Study specific	HICDEP?
Appendix	CrossTable	GC001	New patients who were not submitted last time we did this merger	EPPICC	NO
Appendix	CrossTable	GC002	Patients left out who were submitted last time we did this merger	EPPICC	NO
tblAE_NEW	CrossTable	ANC001	tblIBAS says AE_FPV=1, yet records in tblAE_NEW whilst on FPV	EPPICC	NO
tblAE_NEW	CrossTable	ANC002	AE_DRUG not reported in tblART or tblMED	EPPICC	NO
tblART	WithinTable	AW003	ART_RS=98 yet ART_OTH is null	EPPICC	NO
tblMED	WithinTable	MW002	MED_ONG=0 and MED_ED null	EPPICC	NO
tblMED	WithinTable	MW003	MED_ONG=1 and MED_ED non-null	EPPICC	NO
Appendix	CrossTable	GC003	Spot-check dates to ensure no Excel date-zero errors	EPPICC	NO
tblAE_NADM	WithinTable	ANDW001	Duplicate records for same EVENT_ID	EPPICC	NO

tblAE_NADM	CrossTable	ANDC001	EVENT_ID doesn't exist in tblAE_NEW	EPPICC	NO
tblAE_NADM	WithinTable	ANDW002	Missing Patient	EPPICC	NO
tblAE_NADM	WithinTable	ANDW003	Missing EVENT_ID	EPPICC	NO
tblAE_NADM	WithinTable	ANDW004	Missing CANTYP_T	EPPICC	NO
tblAE_NADM	WithinTable	ANDW005	Missing PATREP_S	EPPICC	NO
tblAE_NADM	WithinTable	ANDW006	DIAOTH_Y=1 and DIA_S is missing	EPPICC	NO
tblAE_NADM	CrossTable	ANDC002	PATIENT doesn't exist in tblBAS	EPPICC	NO
tblAE_NEW	CrossTable	ANC003	PATIENT has no record in BAS	EPPICC	NO
tblAE_NEW	CrossTable	ANC004	AE happened before that drug was started	EPPICC	NO
tblAE_NEW	WithinTable	ANW001	More than one record for same patient on same date with same AE_ID	EPPICC	NO
tblAE_NEW	WithinTable	ANW002	More than one record for same patient with same EVENT_ID	EPPICC	NO
tblAE_NEW	WithinTable	ANW003	Missing PATIENT	EPPICC	NO
tblAE_NEW	WithinTable	ANW004	Missing AE_D	EPPICC	NO
tblAE_NEW	WithinTable	ANW005	Missing AE_ID	EPPICC	NO
tblAE_NEW	WithinTable	ANW006	Missing EVENT_ID	EPPICC	NO
tblAE_NEW	WithinTable	ANW007	Missing AE_RES	EPPICC	NO
tblAE_NEW	WithinTable	ANW008	Missing AE_TEXT	EPPICC	NO
tblAE_NEW	WithinTable	ANW009	Missing AE_DRUG	EPPICC	NO
tblAE_NEW	CrossTable	ANC005	AE_ART=1 but doesn't agree with records in ART table	EPPICC	NO

### All QA checks

A complete list of all QA checks listed in the QA.csv file:

Table	Crosstable	Error Code	Description	Study specific	HICDEP?
AllTables	CrossTable	ATC001	any date in database after DEATH_D in tblLTFU		YES
AllTables	CrossTable	ATC002	any date in database after DROP_D in tblLTFU		YES
AllTables	CrossTable	ATC003	any date in database before BIRTH_D in tblBAS		YES
AllTables	CrossTable	ATC004	any date in database in the future		YES

AllTables	CrossTable	ATC005	patients submitted previously who have been missed out		YES
AllTables	CrossTable	ATC006	Any fields not coded as coding lists on table definition		YES
Appendix	CrossTable	GC001	New patients who were not submitted last time we did this merger	EPPICC	NO
Appendix	CrossTable	GC002	Patients left out who were submitted last time we did this merger	EPPICC	NO
tblAE_NEW	CrossTable	ANC001	tblBAS says AE_FPV=1, yet records in tblAE_NEW whilst on FPV	EPPICC	NO
tblAE_NEW	CrossTable	ANC002	AE_DRUG not reported in tblART or tblMED	EPPICC	NO
tblART	WithinTable	AW001	ART_RS not null, but end date NULL		YES
tblART	WithinTable	AW002	ART_RS null but end-date non NULL		YES
tblART	WithinTable	AW003	ART_RS=98 yet ART_OTH is null	EPPICC	NO
tblART	WithinTable	AW004	Duplicate records for same cohort, patient, art_id and art_sd		YES
tblART	WithinTable	AW005	Dose out of range for those dose units and frequency	PaediatricOnly	YES
tblART	WithinTable	AW006	Missing art_fr	PaediatricOnly	YES
tblART	WithinTable	AW007	Missing art_do	PaediatricOnly	YES
tblART	WithinTable	AW008	Missing patient		YES
tblART	WithinTable	AW009	Missing art_id		YES
tblART	WithinTable	AW010	Missing art_sd		YES
tblART	WithinTable	AW011	Overlapping periods of same drug		YES
tblART	WithinTable	AW012	Double reporting - records reported for both combination drugs and their components		YES
tblART	WithinTable	AW013	Periods of overlap of contra-indicated drugs		YES
tblART	WithinTable	AW014	Restart of same drug without a stop		YES
tblART	WithinTable	AW015	ART_SD greater than or equal to ART_ED		YES
tblART	CrossTable	AC001	Patient has no record in table BAS		YES

tblART	CrossTable	AC002	Records exist in tblART yet RECART_Y=0 in tblBAS		YES
tblBAS	WithinTable	BW001	AIDS date < SEROCO_D		YES
tblBAS	WithinTable	BW002	Duplicate patients		YES
tblBAS	WithinTable	BW003	First 3 chars of PATIENT don't form valid cohort code	CascadeOnly	YES
tblBAS	WithinTable	BW004	Missing PATIENT		YES
tblBAS	WithinTable	BW005	Missing CENTER		YES
tblBAS	WithinTable	BW006	Missing BIRTH_D		YES
tblBAS	WithinTable	BW007	Missing FRSVIS_D		YES
tblBAS	WithinTable	BW008	Missing ENROL_D		YES
tblBAS	WithinTable	BW009	Missing GENDER		YES
tblBAS	WithinTable	BW010	Missing HEIGH		YES
tblBAS	WithinTable	BW011	Missing MODE		YES
tblBAS	WithinTable	BW012	Missing MODE_OTH if MODE=90		YES
tblBAS	WithinTable	BW013	Missing ORIGIN		YES
tblBAS	WithinTable	BW014	Missing ETHNIC		YES
tblBAS	WithinTable	BW015	Missing SEROCO_D		YES
tblBAS	WithinTable	BW016	Missing RECART_Y		YES
tblBAS	WithinTable	BW017	Missing AIDS_Y		YES
tblBAS	WithinTable	BW018	Missing AIDS_D if AIDS_Y=1		YES
tblBAS	WithinTable	BW019	BIRTH_D out of range (15-85 yrs)	CascadeOnly	YES
tblBAS	WithinTable	BW020	BIRTH_D out of range (<18)	PENTA	YES
tblBAS	CrossTable	BC001	RECART_Y=1 but no records in tblART		YES
tblBAS	CrossTable	BC002	AIDS_Y=0, but AIDS-defining records in tblDIS		YES
tblBAS	CrossTable	BC003	AIDS_Y=1 but no AIDS-defining records in tblDIS table		YES
tblDIS	WithinTable	DW001	Duplicate records for same DIS_ID and same DIS_D		YES
tblDIS	WithinTable	DW002	Miscoded DIS_WD as codes on table definition		YES
tblDIS	WithinTable	DW003	Miscoded DIS_ID - as in code list attached to table definition		YES

tbIDIS	WithinTable	DW004	DIS_D missing		YES
tbIDIS	WithinTable	DW005	DIS_ID missing		YES
tbIDIS	WithinTable	DW006	Same event recorded twice - 2 records, same DIS_ID, DIS_D within 6 months		YES
tbIDIS	WithinTable	DW007	DIS_ED present but before DIS_D		YES
tbIDIS	CrossTable	DC001	Patient has no record in BAS		YES
tbIDIS	CrossTable	DC002	AIDS-defining records, yet AIDS=0 in tbIBAS		YES
tbIDIS	CrossTable	DC003	First AIDS-defining DIS_D not equal to AIDS_D in tbIBAS		YES
tbILAB	WithinTable	LW001	Duplicate records		YES
tbILAB	WithinTable	LW002	Missing LAB_D		YES
tbILAB	WithinTable	LW003	Missing LAB_V		YES
tbILAB	WithinTable	LW004	Missing LAB_U		YES
tbILAB	WithinTable	LW005	Missing LAB_ID		YES
tbILAB	WithinTable	LW008	Missing LAB_FA		YES
tbILAB	WithinTable	LW009	Missing LAB_ST		YES
tbILAB	WithinTable	LW010	LAB_V out of range for unit LAB_U		YES
tbILAB	CrossTable	LC001	Patient has no record in BAS table		YES
tbILAB_CD4	WithinTable	CW001	CD4 value out of range		YES
tbILAB_CD4	WithinTable	CW002	Duplicate records for same date		YES
tbILAB_CD4	WithinTable	CW003	Missing CD4_D		YES
tbILAB_CD4	WithinTable	CW004	Missing CD4_V		YES
tbILAB_CD4	WithinTable	CW005	Missing CD4_U		YES
tbILAB_CD4	WithinTable	CW006	Miscoded CD4_U as coding list on table definition		YES
tbILAB_CD4	WithinTable	CW007	CD4_U=2 (percentage) and CD4_V>100		YES
tbILAB_CD4	WithinTable	CW008	CD4_U=(1 or 3) and CD4_V>3000		YES
tbILAB_CD4	WithinTable	CW009	CD4 counts spike up or down suddenly (large change in less than a year)		YES
tbILAB_CD4	CrossTable	CC001	Patient has no record in BAS		YES

tblLAB_RES	WithinTable	LRW001	Duplicate records for same patient on same date		YES
tblLAB_RES	WithinTable	LRW002	Missing PATIENT		YES
tblLAB_RES	WithinTable	LRW003	Missing TEST_ID		YES
tblLAB_RES	WithinTable	LRW004	Missing SAMPLE_D		YES
tblLAB_RES	WithinTable	LRW005	Missing SEQ_DT		YES
tblLAB_RES	WithinTable	LRW006	Missing LAB		YES
tblLAB_RES	WithinTable	LRW007	Missing LIBRARY		YES
tblLAB_RES	WithinTable	LRW008	Missing REFSEQ		YES
tblLAB_RES	WithinTable	LRW009	Missing KIT		YES
tblLAB_RES	WithinTable	LRW010	Missing SOFTWARE		YES
tblLAB_RES	WithinTable	LRW011	Missing TESTTYPE		YES
tblLAB_RES	WithinTable	LRW012	Missing SUBTYPE		YES
tblLAB_RES	WithinTable	LRW013	SEQ_DT has no time part		YES
tblLAB_RES	WithinTable	LRW014	Missing VIRUSTYPE		YES
tblLAB_RES	CrossTable	LRC001	PATIENT has no record in tblBAS		YES
tblLAB_RES	CrossTable	LRC002	This TEST_ID has both LVL_1 and LVL_2 records		YES
tblLAB_RES	CrossTable	LRC003	TESTTYPE in (1,9) (i.e. not phenotype) yet no records in LVL_1 or LVL_2		YES
tblLAB_RES_LVL_1	WithinTable	L1W001	Duplicate records per TEST_ID and SEQTYPE		YES
tblLAB_RES_LVL_1	WithinTable	L1W002	SEQ_START > SEQ_STOP		YES
tblLAB_RES_LVL_1	WithinTable	L1W003	SEQ_NUC contains invalid IUPAC character		YES
tblLAB_RES_LVL_1	WithinTable	L1W004	Missing TEST_ID		YES
tblLAB_RES_LVL_1	WithinTable	L1W005	Missing SEQ_START		YES
tblLAB_RES_LVL_1	WithinTable	L1W006	Missing SEQ_STOP		YES
tblLAB_RES_LVL_1	WithinTable	L1W007	Missing SEQ_NUC		YES
tblLAB_RES_LVL_1	CrossTable	L1C001	TEST_ID not in tblLAB_RES.TEST_ID		YES
tblLAB_RES_LVL_2	WithinTable	L2W001	AA_FOUND_x but nothing in AA_FOUND(x-1)		YES
tblLAB_RES_LVL_2	WithinTable	L2W002	Duplicate records per TEST_ID, GENE, AA_POS and AA_POS_SUB		YES

tbILAB_RES_LVL_2	WithinTable	L2W003	Missing TEST_ID		YES
tbILAB_RES_LVL_2	WithinTable	L2W004	Missing GENE		YES
tbILAB_RES_LVL_2	WithinTable	L2W005	Missing AA_POS		YES
tbILAB_RES_LVL_2	WithinTable	L2W006	Missing AA_POS_SUB		YES
tbILAB_RES_LVL_2	WithinTable	L2W007	Missing AA_FOUNDED_1		YES
tbILAB_RES_LVL_2	CrossTable	L2C001	TEST_ID not in tbILAB_RES.TEST_ID		YES
tbILAB_RES_LVL_3	WithinTable	L3W001	Duplicate records for same TEST_ID and ATC_CODE		YES
tbILAB_RES_LVL_3	WithinTable	L3W002	Missing TEST_ID		YES
tbILAB_RES_LVL_3	WithinTable	L3W003	Missing ATC_CODE		YES
tbILAB_RES_LVL_3	WithinTable	L3W004	Missing RES_SCOR		YES
tbILAB_RES_LVL_3	CrossTable	L3C001	TEST_ID not in tbILAB_RES.TEST_ID		YES
tbILAB_RNA	WithinTable	RW001	Duplicate records for same RNA_D		YES
tbILAB_RNA	WithinTable	RW002	RNA_V < 0 and RNA_L missing		YES
tbILAB_RNA	WithinTable	RW003	RNA_V > 10 000 and (RNA_V modulo 1000) = 1 and RNA_UL missing		YES
tbILAB_RNA	WithinTable	RW004	Missing RNA_D		YES
tbILAB_RNA	WithinTable	RW005	Missing RNA_V		YES
tbILAB_RNA	WithinTable	RW006	RNA_V > 10 million		YES
tbILAB_RNA	WithinTable	RW007	RNA_V < RNA_L and RNA_V >= 0		YES
tbILAB_RNA	WithinTable	RW008	RNA_V > RNA_UL		YES
tbILAB_RNA	CrossTable	RC001	Patient has no record in BAS		YES
tbILAB_VIRO	CrossTable	LVC001	Patient doesn't have a record in BAS		YES
tbILAB_VIRO	WithinTable	LVW002	Missing patient		YES
tbILAB_VIRO	WithinTable	LVW003	Missing VS_ID		YES
tbILAB_VIRO	WithinTable	LVW004	Missing VS_D		YES
tbILAB_VIRO	WithinTable	LVW005	Missing VS_R		YES
tbILAB_VIRO	WithinTable	LVW006	Missing VS_V		YES
tbILAB_VIRO	WithinTable	LVW007	Missing VS_U		YES
tbILAB_VIRO	WithinTable	LVW008	More or less than exactly 1 positive HIV test per patient		YES
tbILAB_VIRO	WithinTable	LVW009	More than 1 negative HIV test for a patient		YES



tblLAB_VIRO	WithinTable	LVW010	Date of negative test after date of positive test		YES
tblLAB_VIRO	WithinTable	LVW011	Duplicate records		YES
tblLTFU	WithinTable	LFW001	DROP_Y and DEATH_Y both non-null		YES
tblLTFU	WithinTable	LFW002	DEATH_Y and DROP_RS both non-null		YES
tblLTFU	WithinTable	LFW003	Any of DEATH_Rx or DEATH_RCx non-null but DEATH_Y=0		YES
tblLTFU	WithinTable	LFW004	R2/RC2 non-null but R1/RC1 null		YES
tblLTFU	WithinTable	LFW005	R3/RC3 non-null but R2/RC2 null		YES
tblLTFU	WithinTable	LFW006	Duplicate patients		YES
tblLTFU	WithinTable	LFW007	Missing DROP_Y		YES
tblLTFU	WithinTable	LFW008	Missing DROP_D if DROP_Y=1		YES
tblLTFU	WithinTable	LFW009	Missing DROP_RS if DROP_Y=1		YES
tblLTFU	WithinTable	LFW010	Missing DEATH_Y		YES
tblLTFU	WithinTable	LFW011	Missing DEATH_D if DEATH_Y=1		YES
tblLTFU	WithinTable	LFW012	DEATH_D non null but DEATH_Y=0		YES
tblLTFU	CrossTable	LFC001	Patient not found in tblBAS		YES
tblLTFU	CrossTable	LFC002	patient in tblBAS hasn't got a record in tblLTFU		YES
tblLTFU	CrossTable	LFC003	tblBAS has AIDS=0 but DEATHRx =8.1		YES
tblMED	WithinTable	MW001	Duplicate records		YES
tblMED	WithinTable	MW002	MED_ONG=0 and MED_ED null	EPPICC	NO
tblMED	WithinTable	MW003	MED_ONG=1 and MED_ED non-null	EPPICC	NO
tblMED	WithinTable	MW004	Missing MED_ID		YES
tblMED	WithinTable	MW005	Missing MED_SD		YES
tblMED	WithinTable	MW006	Mising MED_ED		YES
tblMED	WithinTable	MW007	Overlapping periods of same drug		YES
tblMED	WithinTable	MW008	MED_ED < MED_SD		YES
tblMED	WithinTable	MW009	MED_RS not null, but end date NULL		YES

tblMED	CrossTable	MC001	Patient has no records in tblBAS		YES
tblOVERLAP	WithinTable	OW001	Invalid other cohort		YES
tblOVERLAP	CrossTable	OC001	PATIENT not found in tblBAS for that cohort		YES
tblOVERLAP	CrossTable	OC002	PAT_OTH not found in tblBAS for that overlapping cohort		YES
tblVIS	WithinTable	VW001	Duplicate records, same VIS_D		YES
tblVIS	WithinTable	VW002	Height decreasing over time		YES
tblVIS	WithinTable	VW003	Height out of acceptable range		YES
tblVIS	WithinTable	VW004	Weight out of acceptable range		YES
tblVIS	CrossTable	VC001	patient has no record in BAS table		YES
tblVIS	CrossTable	VC002	No weights within 3 mths of starting FPV/DRV	EPPICC	YES
Appendix	CrossTable	GC003	Spot-check dates to ensure no Excel date-zero errors	EPPICC	NO
tblAE_NADM	WithinTable	ANDW001	Duplicate records for same EVENT_ID	EPPICC	NO
tblAE_NADM	CrossTable	ANDC001	EVENT_ID doesn't exist in tblAE_NEW	EPPICC	NO
tblAE_NADM	WithinTable	ANDW002	Missing Patient	EPPICC	NO
tblAE_NADM	WithinTable	ANDW003	Missing EVENT_ID	EPPICC	NO
tblAE_NADM	WithinTable	ANDW004	Missing CANTYP_T	EPPICC	NO
tblAE_NADM	WithinTable	ANDW005	Missing PATREP_S	EPPICC	NO
tblAE_NADM	WithinTable	ANDW006	DIAOTH_Y=1 and DIA_S is missing	EPPICC	NO
tblAE_NADM	CrossTable	ANDC002	PATIENT doesn't exist in tblBAS	EPPICC	NO
tblAE_NEW	CrossTable	ANC003	PATIENT has no record in BAS	EPPICC	NO
tblAE_NEW	CrossTable	ANC004	AE happened before that drug was started	EPPICC	NO
tblAE_NEW	WithinTable	ANW001	More than one record for same patient on same date with same AE_ID	EPPICC	NO
tblAE_NEW	WithinTable	ANW002	More than one record for same patient with same EVENT_ID	EPPICC	NO
tblAE_NEW	WithinTable	ANW003	Missing PATIENT	EPPICC	NO

tblAE_NEW	WithinTable	ANW004	Missing AE_D	EPPICC	NO
tblAE_NEW	WithinTable	ANW005	Missing AE_ID	EPPICC	NO
tblAE_NEW	WithinTable	ANW006	Missing EVENT_ID	EPPICC	NO
tblAE_NEW	WithinTable	ANW007	Missing AE_RES	EPPICC	NO
tblAE_NEW	WithinTable	ANW008	Missing AE_TEXT	EPPICC	NO
tblAE_NEW	WithinTable	ANW009	Missing AE_DRUG	EPPICC	NO
tblAE_NEW	CrossTable	ANC005	AE_ART=1 but doesn't agree with records in ART table	EPPICC	NO
tblDELIVERY_CHILD	CrossTable	DCC001	MOTHER_ID doesn't exist in tblPREG		YES
tblDELIVERY_CHILD	CrossTable	DCC002	CHILD_ID doesn't exist in tblPREG_OUT		YES
tblDELIVERY_CHILD	WithinTable	DCW001	Duplicate records for same Mother_ID, CHILD_ID		YES
tblDELIVERY_CHILD	WithinTable	DCW002	B_SEQ >4		YES
tblDELIVERY_CHILD	WithinTable	DCW003	B_SEQ>1, but no record exists for same Mother, Child with B_SEQ=1		YES
tblDELIVERY_CHILD	WithinTable	DCW005	DELIV_D < MEMRUP_D		YES
tblDELIVERY_CHILD	WithinTable	DCW004	LABOUR_P >72		YES
tblDELIVERY_CHILD	WithinTable	DCW006	2 children, same mother, same MEMRUP_D, DELIV_D >DELIV_D , yet B_SEQ< B_SEQ		YES
tblDELIVERY_MUM	CrossTable	DMC001	MOTHER_ID doesn't exist in tblPREG		YES
tblDELIVERY_MUM	CrossTable	DMC002	MOTHER_ID doesn't exist in tblDELIVERY_CHILD with same MEMRUP_D		YES
tblDELIVERY_MUM	WithinTable	DMW001	PREG_SEQ invalid (PREG_SEQ>1 and PREG_SEQ=x, but PREG_SEQ=x-1 doesn't exist)		YES
tblDELIVERY_MUM	WithinTable	DMW002	DISCHA_D < MEMRUP_D		YES
tblDELIVERY_MUM	WithinTable	DMW004	INTERV=0-3 but INTERV_O non null		YES
tblDELIVERY_MUM	WithinTable	DMW003	INTERV=90,91, 92 or 93 but INTERV_O null		YES
tblNEWBORN	CrossTable	NC001	CHILD_ID doesn't exist in tblDELIVERY_CHILD		YES

tblNEWBORN	CrossTable	NC002	ABNORM_Y = 1, yet no records in tblNEWBORN_ABNORM		YES
tblNEWBORN	CrossTable	NC003	ABNORM_Y = 0 or 9, yet records in tblNEWBORN_ABNORM		YES
tblNEWBORN	WithinTable	NW001	BRFEED_SD>BRFEED_ED		YES
tblNEWBORN	WithinTable	NW002	APGARM_x's out of order (e.g APGARM_3 < APGARM_2)		YES
tblNEWBORN	WithinTable	NW003	ICU_Y=1, but ICU_S or ICS_D null		YES
tblNEWBORN	WithinTable	NW004	ICU_Y=0 or 9, but ICU_S or ICU_D non null		YES
tblNEWBORN_ABNORM	CrossTable	NAC001	CHILD_ID doesn't exist in NEWBORN		YES
tblNEWBORN_ABNORM	WithinTable	NAW001	ABNORM_T=90 but ABNORM_S null		YES
tblPREG_OBS	CrossTable	POC001	MOTHER_ID+PREG_SEQ doesn't exist in tblPREG		YES
tblPREG_OBS	WithinTable	POW001	PROB_T=99 but PROB_S null		YES
tblPREG_OBS	WithinTable	POW002	PROB_T<>99 but PROB_S non null		YES
tblPREG_OBS	WithinTable	POW003	PROB_T=2 but CERVIX_S null		YES
tblPREG_OBS	WithinTable	POW003	PROB_T<>2 but CERVIX_S non null		YES
tblPREG_OUT	CrossTable	PTC001	MOTHER_ID+PREG_SEQ doesn't exist in tblPREG		YES
tblPREG_OUT	CrossTable	PTC002	OUTCOM=1,2 or 3 and CHILD_ID doesn't exist in tblNEWBORN or tblDELIVERYCHILD		YES
tblPREG_OUT	CrossTable	PTC003	OUTCOM=10,11,20 or 21 and CHILD_ID exists in tblNEWBORN or tblDELIVERYCHILD		YES
tblPREG_OUT	WithinTable	PTW001	OUTCOME=20 or 21 and OUTCOM_R null		YES
tblPREG_OUT	WithinTable	PTW002	OUTCOME not 20 or 21 and OUTCOM_R non null		YES
tblPREG	WithinTable	PW001	ANC_D		YES
tblPREG	WithinTable	PW002	INPREG_Y=2,3,4 but INHIST_Y null		YES
tblPREG	WithinTable	PW003	INPREG_Y=1 but INHIST_Y non null		YES
tblPREG	WithinTable	PW004	INHIST_Y=1 but INHIST_S null		YES

tblPREG	WithinTable	PW004	INHIST_Y=0 or 9 but INHIST_S non null		YES
tblPREG	WithinTable	PW005	KARYO_T=2 but KARYO_A null		YES
tblPREG	WithinTable	PW006	KARYO_T<>2 but KARYO_A non null		YES
tblPREG	WithinTable	PW007	ULTRA_x =2, but ULTR_A_x null		YES
tblPREG	WithinTable	PW008	ULTRA_x <>2, but ULTR_A_x non null		YES
tblPREG	CrossTable	PC001	PROB_Y = 1 but no records in tblPREG_OBS		YES
tblPREG	CrossTable	PC002	PROB_Y = 0 or 9 but records in tblPREG_OBS		YES
tblREFILL	CrossTable	PRC001	Patient doesn't have a record in BAS		YES
tblREFILL	WithinTable	PRW002	Missing PATIENT		YES
tblREFILL	WithinTable	PRW003	Missing REFILL_D		YES
tblREFILL	WithinTable	PRW004	Missing DRUG_ID		YES
tblREFILL	WithinTable	PRW005	Missing SUPPLY		YES
tblREFILL	WithinTable	PRW006	SUPPLY < 1		YES