STANDARD PROCEDURE FOR DATA TRANSFER

Version 3 – September 2021

Contact: info@iedea-sa.org
www.iedea-sa.org / www.iedea.org/

Contents

1 Introduction

1.1 General remarks

- This document provides guidance on the preparation of data tables for the transfer of data for the IeDEA Southern Africa Collaboration.
- It is requested that each clinic prepares **separate data tables**, as described in the IeDEA Data Exchange Standard (DES) detail below. A minimum of 10 tables should be submitted by all sites, other tables could be completed by certain sites if data is available (see below).
- The tables can be sent in the format that is most convenient for the site, we can work with simple delimited files, common statistical data file format (Stata, SPSS, SAS, etc…) as well as database backup copy (.bak, MS SQL Server) or MS Access. Please contact the IeDEA data managers if you have any queries.
- **In case of necessity, the data manager from data center can guide and help with the preparation of the data into IeDEA DES tables. It is also requested to carefully document the variables you extract and how to transform them into IeDEA variables.**
- It is accepted that there will be missing data for some patients, and even entire missing tables from some sites who simply do not have that data in electronic format or do not collect that data routinely.
- It is requested that for security purposes, data tables be encrypted and compressed with a file compression utility (e.g WinZip, …) prior to sending. The encryption password (minimum of 10 characters long, including upper/lower case, numbers and special characters) should be communicated to the relevant data center contact person separately.
- Encrypted and compressed dataset should be sent via a secured cloud server, relevant to the data center (a data voucher for transfer, can be sent on request).
- Please ensure that ALL datasets have been stripped of personal identifying information prior to sending. (i.e clinic medical record number; folder number, etc.).
- Please include a unique anonymous identifier for each patient (PATIENT. This is a special identifier created for IeDEA Southern Africa. This anonymization key serves as cross-reference to your own database and must be maintained by the cohort/program under secure conditions. This unique identifier must remain consistent for the same patient over different data transfers to IeDEA Southern Africa through times.

Thank you very much for your contribution to this collaborative project!

1.2 Inclusion criteria for patients

Please include all patients with the following characteristics:

- Documented HIV infection (HIV positive patients)
- Patients in HIV care at the site for whom the date of first visit at the site is known exactly.

Notes:

- Where possible, it is intended that data be transferred on HIV-infected patients followed-up at the site irrespective of whether or not they received antiretroviral therapy (ART).
- When transferring data just on patients who received ART, it is preferable to include patients irrespective of whether or not they were exposed to anti-retroviral before the recorded ART start date. In other words, treatment-naïve and treatment-experienced patients are included.
- Sites should send all information on all patients (adults and/or children) in a single dataset. Pediatric specific fields must be entered as completely as possible for all patients whose **first visit at your facility is before their 16th birthday** even if their follow-up extends beyond the age of 16 years.
- Some patients will have been in care at another site/center prior to commencing care at your site/center. These patients should be included in the dataset, noting against the relevant field that they have been transferred in. All treatments and diseases (DIS) history prior to commencing care at your site/center should be reconstructed as far as possible and entered in the appropriate tables, with unknown codes for dates of start and end date of diseases/antiretroviral drugs where necessary.

1.3 Dates

- The term baseline will not be used as this creates confusion. We will rather make use of a set of key dates that will be entered into the first table, the **tblBAS** table. These are:

Variable name	Definition of key date
ENROL_D	Date of enrolment into the cohort
HIV_POS_D	Date of first positive HIV test
RECART_D	Date of first antiretroviral treatment initiation. *Leave blank if ART not yet initiated. This should be the first date at which antiretroviral therapy, regardless of regimen, was given as treatment irrespective of whether it was given at this center/program or not. It excludes antiretroviral regimens given only for PMTCT or other prophylaxis.*

- For all fields that require a date, the precise date should be entered in the format YYYY-MM-DD if it is known. However, it might be that some cohorts are limited to representing date data at the level of the month or year only.

> In case the date day is unknown, 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.

> In case both the month and day are unknown, the date should be coded from the mid-point of the year, so that 1999-??-?? becomes 1999-07-01.

> If the year is unknown but the presence of the date value is needed, a fictitious date should be used that couldn't be mistaken with an actual date. An unknown year should be coded as 1911-11-11.

> For issues regarding the precision of the dates, a character code is used to specify at which degree the day, month, or year date is precise. The annotation variable will have the same name as the date variable with the additional suffix _A. For example, the precision of BIRTH_D will be annotated using additional optional variables called BIRTH_D_A.

Character code	Precision of date
<	Before this date
D	Exact to the date
M	Exact to the month
Y	Exact to the year
>	After the date
U	Unknown

- For certain date fields a precise date is obligatory e.g. date of enrolment in your program/cohort (ENROL_D) and date of ART initiation (RECART_D). In patients who commenced ART at another site, if the precise date of ART initiation cannot be estimated reasonably accurately, the patient should be entered as treatment experienced and the date of first visit at your site will be regarded as the date of ART initiation.

1.4 Definitions

- ART is defined as HIV treatment regimen with either individual drugs or combination drugs of any class or classes.
- "Treatment experienced" is defined as previous exposure to any antiretroviral drug, **excluding** exposure for prevention of mother to child transmission (PMTCT) or post-exposure prophylaxis (PEP).

1.5 Standard codes

Certain codes will appear repeatedly in a number of lists for coded fields. In this instance, the same codes/coding format will be used in all fields where these codes appear as follows:

Codes	Description
0	No
1	Yes
95	Not ascertained/Not collected
99	Unknown despite attempting ascertainment
88	Not applicable

1.6 Data tables

For each clinic, the following data tables or files should be prepared, depending on data availability.

	Table Name	Description
1	tblART	Antiretroviral medication
2	tblBAS	Patient's basic information
3	tblCENTER	Site-specific information
4	tblDIS	CDC-C and WHO stage diseases
5	tblLAB	Laboratory values - general
6	tblLAB_CD4	Laboratory values - CD4+ cell count tests
7	tblLAB_RNA	Laboratory values - viral assay (HIV)
8	tblLTFU	Death and dropout
9	tblMED	Other medications
10	tblVIS	Visit-related information
11	tblTRANSFER	Transfer's information
12	tblTB	Tuberculosis (TB) information
13	tblART_MUM	Antiretroviral Medication of mother in cases where mother is not enrolled in cohort
14	tblCANC	Diagnosis of cancer
15	tblLAB_BP	Laboratory values - blood pressure
16	tblLAB_VIRO	Laboratory values - viro/serology
17	tblLAB_RES	Resistance testing
18	tblLAB_RES_LVL_2	Mutations
19	tblLAB_RES_LVL_3	Resistance test result
20	tblDELIVERY_CHILD	Delivery information related to child
21	tblDELIVERY_MUM	Delivery information related to mother
22	tblNEWBORN	Newborn information
23	tblNEWBORN_ABNORM	Newborn abnormalities
24	tblPREG	Pregnancy
25	tblPREG_OUT	Pregnancy outcome
26	tblOVERLAP	Cross-cohort identification
27	tblPROGRAM	Linkage of care programs/sites to regions

- Tables 1 to 10 are required by all sites.
- Table 13 (tblART_MUM) is required only from sites that record treatment of a patient's mother but the mother is not enrolled in the cohort.
- Table 14 (tblCANC) is required only from sites that record detailed information on cancer electronically.
- Table 15 (tblLAB_BP) is required only for sites that record blood pressures electronically
- Table 16 (tblLAB_VIRO) is required only for sites that record viral tests (e.g Hepatitis,…) electronically
- Table 17 to 19 (tblLAB_RES, tblLAB_RES_LVL_2, tblLAB_RES_LVL_3) are required only for sites that record information on drug resistance.

- Table 20 to 25 are required only for sites that have mother and child clinic, or do follow-up and record information electronically on pregnancy, as well as labor and delivery.
- In addition, information on the overall cohort as well as centers from which clinical data is extracted, and until when - or "meta-data" for the transfer - must be included with all transfers (tblCENTER).

2 Table structure and variables

All details about table structure and variables (definition, data types, values, etc…) can be consulted online using this link (except table tblTRANSFER and tblTB – see below). This is updated regularly so please ensure to check it when you prepare a data extract

https://redcap.vanderbilt.edu/plugins/iedea/des/

Additional table structures

2.1 Transfers (tblTRANSFER) table

tblTransfer details information on the patient transfer's in and transfer's out of facilities. There can be more than one records per patient.

Variables to be included in TRANSFER table

Name	Format	Description
MERGE	Numeric	Number of merge
PATIENT	Free (numerical or alphanumerical)	Unique, anonymous, patient identifier
TRANSFER_D	Date (YYYY-MM-DD)	Date of transferred into/out of cohort.
TRANSFER_TYPE	Numeric (see coding List 16)	Transferred out to/from?
TRANSFER_RS	Numeric (see coding List 17)	Reason for transfer

List 1 - Codes for transfer type (TRANSFER_TYPE)

Code source: IeDEA SA codes

Table name: LU: TRANSFER: TFI/ TFO

Codes	Type of Transfer
30	Transfer in from within the same service
31	Transfer in from another service
32	Transfer in unknown
33	Transfer out within the same service
34	Transfer out to another service
35	Transfer out unknown
90	Other (not listed above)
95	Not ascertained

List 2- Codes for transfer reason (TRANSFER_RS)

Code source: IeDEA SA codes

Table name: LU: TRANSFER: TRANSFER_RS

Codes	Transfer Reason
1	Moved away from the area
2	Moved into the area
3	Moved to another level of care
3.1	Moved to/from lower level of care
3.2	Moved to/from higher level of care
3.3	Moved to/from same level of care
4	Paediatric to adult facility (transition)
4.1	Child to adolescent/youth facility
4.2	Adolescent/Youth to adult facility
5	Antenatal/Pregnancy
5.1	Pregnancy started
5.2	Pregnancy ended
6.1	Transferred into club
6.2	Transferred from club
7	Partaking in a trial
7.1	Trial started
7.2	Trial ended
90	Other (not listed above)
95	Not ascertained

2.2 Tuberculosis information (TB table)

This table is for capturing details of the TB episodes during HIV follow-up. Tests related to TB can be included in the LAB table. Where possible this data can be derived from the electronic TB register.

Variables to be included in the table TB

Name	Format	Description
PATIENT	Free (numerical or alphanumerical)	Unique patient identifier
REG_D	Date (YYYY-MM-DD)	Date registered with TB
REGID	Text (eg. 2272007) -95 = Not ascertained -99 = Unknown despite attempting ascertainment	TB register number

Name	Format	Description
RAD	Numeric with codes 0 – Not done 1 – Normal 20 – Abnormal unspecified 21 – Abnormal - not consistent with current TB 22 – Abnormal - consistent with current TB unspecified 23 – Abnormal – consistent with current TB – Cavity on right 24 - Abnormal – consistent with current TB – Cavity on left 25 - Abnormal – consistent with current TB – Bilateral cavities 26 - Abnormal – consistent with current TB - No cavities 99 = Unknown despite attempting ascertainment	Radiography findings if done
RESISTANT	Numeric with codes 0 – No 1 – MDR 2 – XDR 95 = Not ascertained 99 = Unknown despite attempting ascertainment	Resistance data based on sensitivities Note: Exact results of sensitivities should be record in the LAB table. Code as MDR if … to more than one drug and XDR if… Categories MDR and XDR should be for the worst resistance status during the episode.
TB_START_D	Date (YYYY-MM-DD)	Date starting TB treatment
TB_END_D	Date (YYYY-MM-DD)	Date ending TB treatment or date of outcome
CAT	Numeric with codes 1 – Newly diagnosed for the first time 2 – After relapse 3 – After default 4 – After failure 95 = Not ascertained 99 = Unknown despite attempting ascertainment	TB Category
CLASS	Numeric with codes 1 - Pulmonary 2 – Extra-pulmonary 3 – Both pulmonary and extra-pulmonary 4 - Primary 95 = Not ascertained 99 = Unknown despite attempting ascertainment	Classification of episode

Name	Format	Description
SITE	Numeric with codes 1 – Bones/Joints (A18.0) 2 – Lymph nodes (A16.3) 3 – Meningitis (A17.0) 4 – Miliary (A19.9) 5 – Pleura (A16.5) 9 – Other sites (A18.8) 88 – Not applicable as pulmonary or primary only 95 = Not ascertained 99 = Unknown despite attempting ascertainment	Site of disease if extra-pulmonary component diagnosed
REGIMEN	Numeric with codes 1 – 2HRZE 4HR - Regimen 1 2 –2HRZES 1HRZE 5HRE - Regimen 2 3 – 2HRZ 4HR - Regimen 3 4 –Other Regimen 95 = Not ascertained 99 = Unknown despite attempting ascertainment	TB treatment regimen
REG_OTHER	Text	Text field for other regimen not included in codes for REGIMEN field above
TB_OUTCOME	Numeric with codes 1 – Completed 2 – Cured 3 – Failed 4 – Interrupted 5 – Defaulted 6 – Treatment ongoing 7 - Died 95 = Not ascertained 99 = Unknown despite attempting ascertainment	Outcome of TB episode