

Data Harmonization

leDEA Global Consortium Annual Meeting

June 2016

Beverly Musick

Outline

In this presentation I will address the following broad topics:

- What data do we collect in leDEA?
- How are the data merged and harmonized to create regional and multi-regional research databases?
- What initiatives are being undertaken to enhance data harmonization and align the leDEA data with HICDEP and other standard terminologies to create a global data model for HIV observational data?

What data does leDEA have?

- leDEA gathers and curates data collected during routine clinical (outpatient) HIV care
- Currently, the 7 leDEA regional consortia follow over 1 million HIV patients enrolled in care and treatment programs

HIV care cascade*



leDEA has data in the latter four stages (i.e., data obtained while in care)

* [AIDS.gov](https://aids.gov)

Harmonization of leDEA data

leDEA data are harmonized in two stages:

- At the regional level
- At the global (multi-regional) level

Harmonization at the regional level

- At the regional level
 - Data from participating sites are transmitted to Regional Data Centers (RDCs)
 - Data are harmonized based on regional standards

Harmonization at the global level

- At the global (multi-regional) level
 - Within each region, data are harmonized to the global leDEA Data Exchange Standards (leDEA-DES)
 - Since leDEA does not have a global data coordinating center, harmonized data pertaining to a specific multi-regional research proposal are transmitted to the RDC leading the project

leDEA Data Exchange Standard

- Aimed at enhancing cross-region collaboration
- Developed through the leDEA Data Harmonization Working Group
- Based on **HIV Cohorts Data Exchange Protocol (HICDEP)**
- The current and extended versions have been developed in collaboration with EuroCoord

Current leDEA-DES

Table Name	Description
tblBAS	patient demographics
tblART	antiretroviral drugs
tblDIS	diseases (CDC & WHO stage events)
tblLAB_CD4	CD4 measurements
tblLAB_RNA	viral assay
tblTFU	death and drop-out
tblVIS	visit-related information
tblPROGRAM	linking sites to programs
tblCENTER	site-specific information

Extended leDEA DES

Table Name	Description
tblCANC	diagnosis of cancer
tblCEP	clinical events and procedures
tblDIS_TB	tuberculosis disease and treatment
tblPREG	general pregnancy-related information
tblPREG_OUT	pregnancy outcome
tblDELIVERY_CHILD	delivery information related to child
tblDELIVERY_MUM	delivery information related to mother
tblNEWBORN	information related to newborns
tblNEWBORN_ABNORM	information related to abnormalities of newborn
tblLAB	laboratory values
tblLAB_BP	laboratory values - blood pressure
tblLAB_RES	laboratory values - resistance testing
tblLAB_RES_LVL_2	mutations
tblLAB_RES_LVL_2	resistance test result
tblLAB_VIRO	laboratory values – viro/serology
tblMED	other medication
tblOVERLAP	participation in other cohorts

Conclusions

- leDEA harmonizes and curates data on more than 1,000,000 HIV-infected patients in care and treatment programs in 7 world regions
- leDEA data comprise information collected while patients are in care
- leDEA data exchange standards are aligned with other existing standard terminologies creating a global data model for HIV observational data