

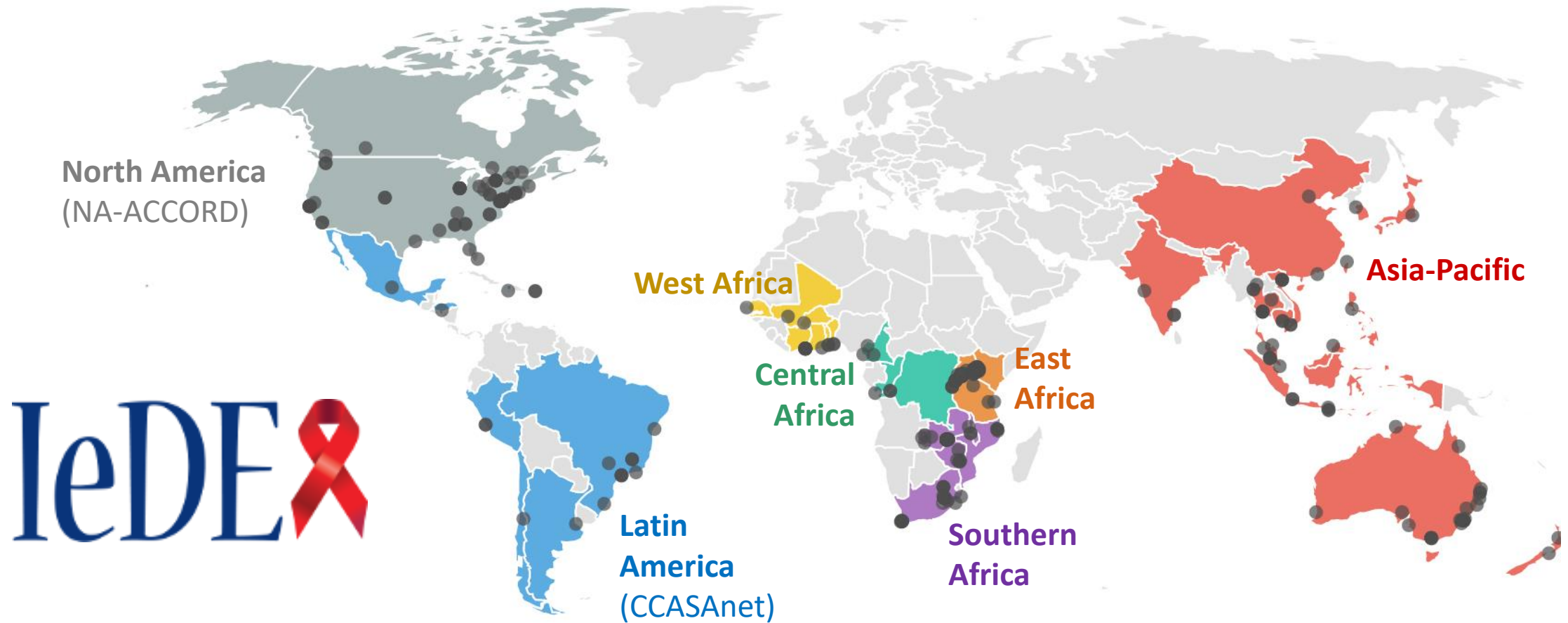
Streamlining Data Sharing in a Global HIV Research Consortium

Stephany Duda PhD and Judy Lewis PhD

Vanderbilt University Medical Center

October 1, 2021

leDEA – A global HIV research consortium

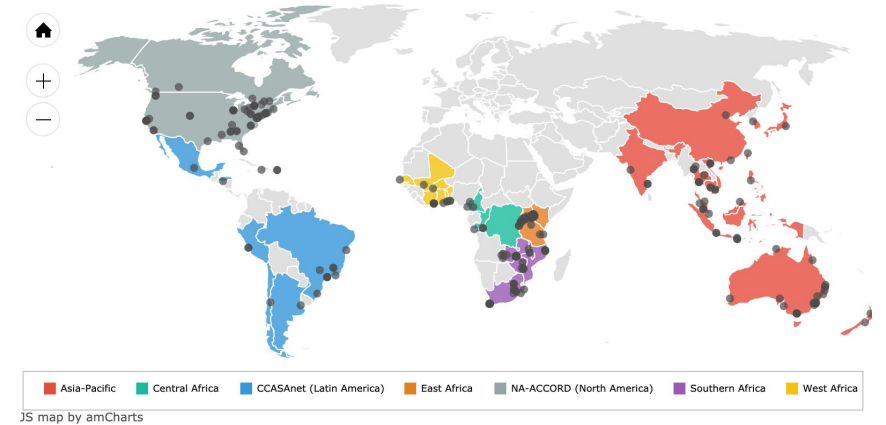


International epidemiology Databases to Evaluate AIDS

A global, NIAID-funded collaboration of seven regional HIV observational research networks with combined data on nearly two million persons living with HIV (PLWH)

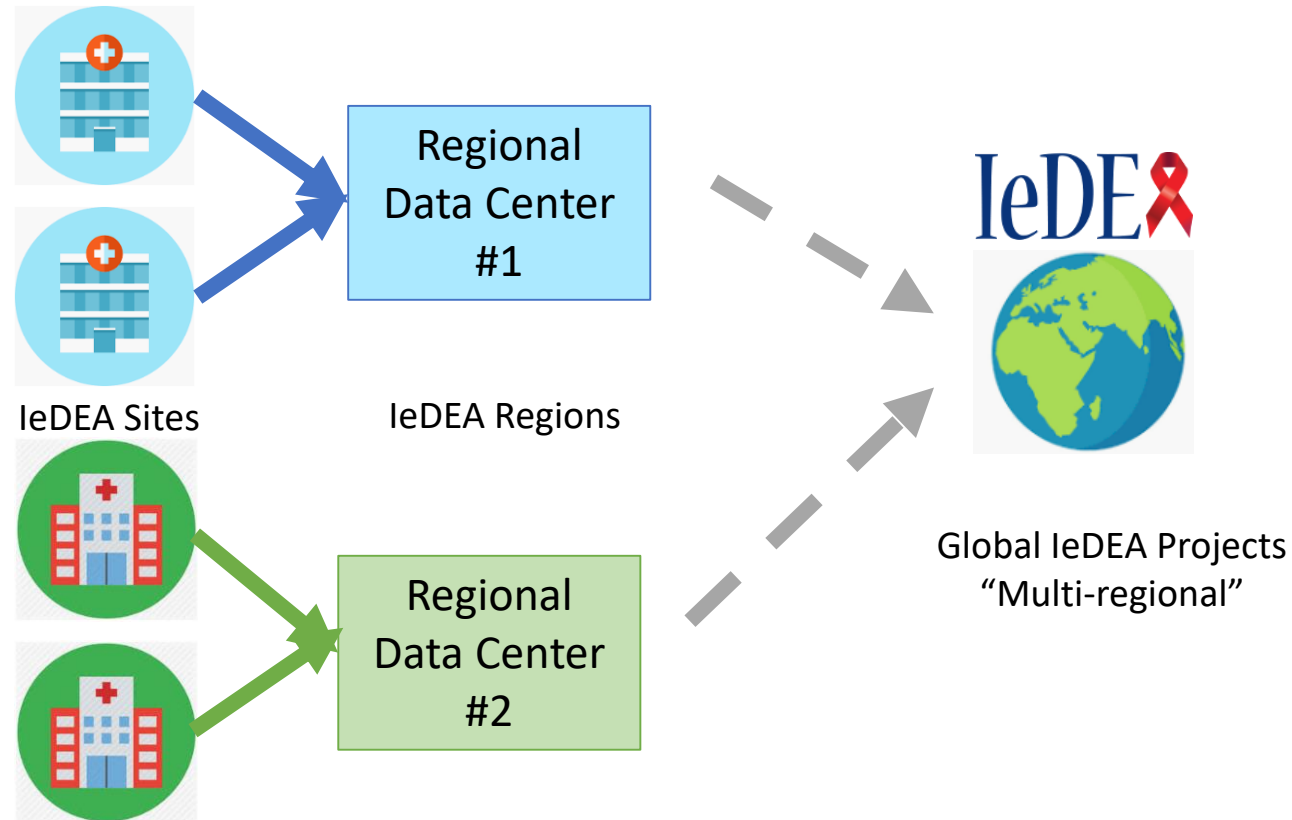
leDEA is a collaboration, not a single network

- Each region operates independently, has its own coordinating center, and conducts regional research.
- Regions come together for multiregional research projects.
- There is no overall coordinating center for leDEA (as of July 2016).
- There has never been an overall data coordinating center for leDEA



(This turned out to be great)

Flow of leDEA data

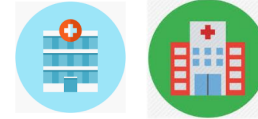


In leDEA

- Sites generate the data.
- Regional Data Centers combine all the data from one region.
- Researchers can get data from multiple regions for a global leDEA project.

Data Considerations

- Data from every clinic can be different.



- Data at every Regional Data Center can be different.

Regional Data
Center #1

Regional Data
Center #2

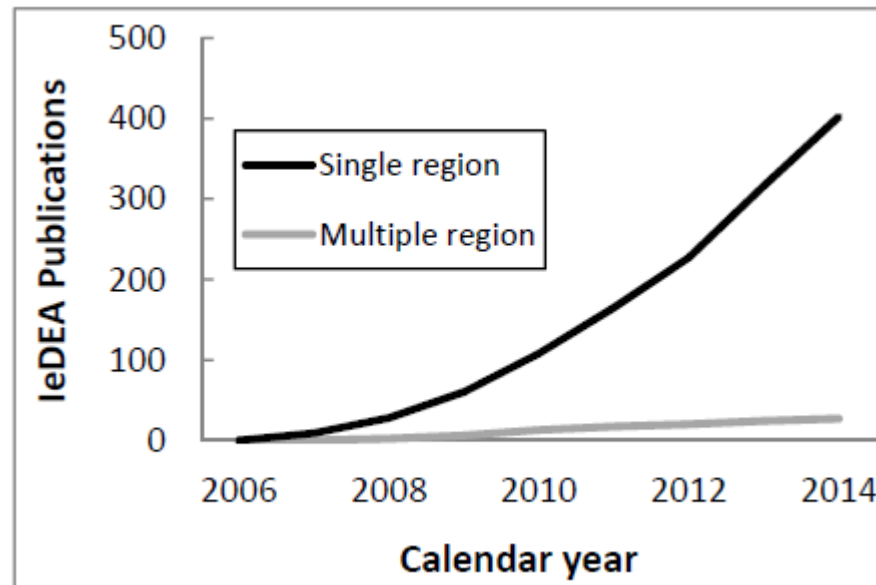
- Global IeDEA data are not stored centrally – subsets of the data are merged for specific projects.
- Sites and Regions have the ultimate say in whether their data is included for a specific project.



In the Early Days of leDEA...

- We had **no standardized way to share data** for global projects.
- Multi-regional projects (projects with 3+ leDEA regions) were **very slow**, in part because it was difficult to merge the data.

Cumulative number of leDEA publications by publication year
(figure from Constantin Yiannoutsos)





The Harmonist Project

Researchers wanted to collaborate efficiently, but they **needed the right tools to support them**. We focused on three areas:

- (1) Support for Data Model
leDEA Data Exchange Standard → Clear data standards
- (2) Consortium Portal
Harmonist Hub → Managed research workflow
- (3) Data Quality and Sharing Tool
Harmonist Data Toolkit → High quality data

Goal: address leDEA's needs but always look for generalizable solutions that can benefit others.

The Harmonist Project



(1) Support for Data Model
leDEA Data Exchange Standard

→ Clear data standards

(2) Consortium Portal
Harmonist Hub


→ Managed research workflow

(3) Data Quality and Sharing Tool
Harmonist Data Toolkit

→ High quality data

(1) IeDEA Data Exchange Standard (DES)

The **IeDEA Data Exchange Standard (DES)** is a **common data model** for HIV data sharing within IeDEA. The document defines variable names, variable definitions, and code lists.



Data Transfer Protocol for IeDEA Multi-regional Collaboration

Approved by the IeDEA Executive Committee on Thursday August 13, 2015

Appendix A (IeDEA-DES Tables) revised by IeDEA Data Harmonization Working Group on April 14, 2015

Appendix A: IeDEA-DES Tables
 (note: the Data Harmonization Working Group will be responsible for recent revisions of this section of the data transfer protocol both in part and in whole. Tables that are approved and designated as either HICDEP+ or Non-HICDEP+ provide additional documentation of the modified or additional data elements.)

Date last revised: Tuesday April 14, 2015 (Text in green represent recent changes.)
 Designations based on HICDEP version 1.8

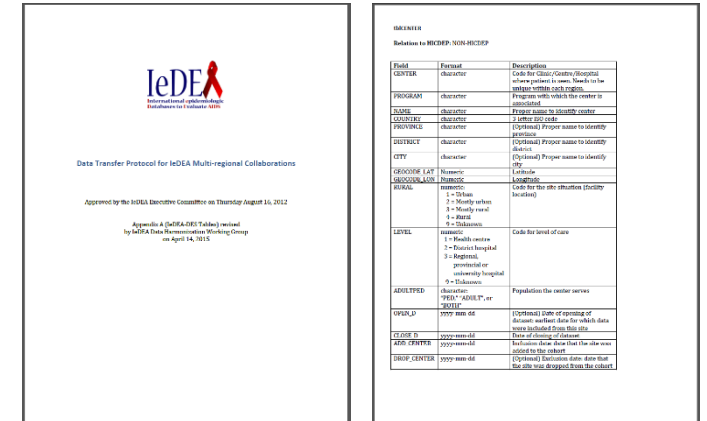
Table Name	Description	Not Yet Designated
tbART	antiretroviral drugs	
tbBAS	basic information	
tbCANC	cancer diagnoses	
tbCENTER	site-specific information	
tbCEP	clinical events including serious non-AIDS conditions	X
tbDELIVERY_CHILD	delivery information related to child	X
tbDELIVERY_MUM	delivery information related to mother	X
tbDIS	diseases (CDC-C)	
tbLAB	laboratory tests	X
tbLAB_BP	blood pressure	X
tbLAB_CD4	CD4 measurements	
tbLAB_RES	resistance testing information	X
tbLAB_RES_LVL_1	nucleoside sequence for PRO and RT	X
tbLAB_RES_LVL_2	mutations and positions of PRO and RT sequences	X
tbLAB_RES_LVL_3	resistance result	X
tbLAB_RNA	viral assay	
tbLAB_VIRO	viro-/serological Tests	X
tbLTFU	death and drop-out	
tbMED	medicines	X
tbNEWBORN	information related to newborns	X
tbNEWBORN_ABNORM	information related to abnormalities of newborn	X
tbOVERLAP	participation in other cohorts	X
tbPREG	general pregnancy-related information	X
tbPREG_OBS	obstetrical problems	X
tbPREG_OUT	pregnancy outcome	X
tbPROGRAM	linking sites to programs	
tbREFILL	prescription refills	X
tbSAMPLES	biological sample storage	X
tbVIS	visit-related information	

tbCENTER
 Relation to HICDEP: NON-HICDEP

Field	Format	Description
CENTER	character	Code for Clinic/Centre/Hospital where patient is seen. Needs to be unique within each region.
PROGRAM	character	Program with which the center is associated
NAME	character	Proper name to identify center
COUNTRY	character	3-letter ISO code
PROVINCE	character	(Optional) Proper name to identify province
DISTRICT	character	(Optional) Proper name to identify district
CITY	character	(Optional) Proper name to identify city
GEOCODE_LAT	Numeric	Latitude
GEOCODE_LON	Numeric	Longitude
RURAL	numeric:	Code for the site situation (facility location)
	1 = Urban	
	2 = Mostly urban	
	3 = Mostly rural	
	4 = Rural	
	9 = Unknown	
LEVEL	numeric:	Code for level of care
	1 = Health centre	
	2 = District hospital	
	3 = Regional, provincial or university hospital	
	9 = Unknown	
ADULTPED	character: "PED," "ADULT", or "BOTH"	Population the center serves
OPEN_D	yyyy-mm-dd	(Optional) Date of opening of dataset: earliest date for which data were included from this site
CLOSE_D	yyyy-mm-dd	Date of closing of dataset
ADD_CENTER	yyyy-mm-dd	Inclusion date: date that the site was added to the cohort
DROP_CENTER	yyyy-mm-dd	(Optional) Exclusion date: date that the site was dropped from the cohort

Maintaining the leDEA DES


- Challenges with MS Word documents
 - Multiple versions, potentially conflicting edits
 - Hard to find latest version in files, email
 - Single copy is not group editable
 - **Not machine-readable**
- Needed a machine-readable solution that was easy to edit and didn't require technical training.
- **Solution:** Use REDCap to create human-readable forms that produce machine-readable structures




Representing the leDEA DES in REDCap

To represent the DES in REDCap, we designed **three data entry forms**:

1. Information about Tables (e.g, demographics, visits, labs, meds)
2. Information about Variables
3. Information about Code Lists

Actions:  Modify instrument  Download PDF of instrument(s) **Table Metadata**

 Editing existing Record ID **3** tblBAS

Record ID 3
To rename the record, see the record action drop-down at top of the [Record Home Page](#).

Table Definition

Table name

Table type
 One row per patient
 Multiple rows per patient
 N/A (e.g., tblCENTER) [reset](#)

Table definition
(brief text)

Display Settings

Display this table in human-readable documents and forms? Yes No [reset](#)

Table display name
(optional, if different title is needed for human-readable documents)

Display order for this table
(e.g., use 1.5 to place a table between tables 1 and 2)

Text (HTML-formatted) to display before the table definiton

Every Patient ID must have one and only one entry in tblBAS.

Text (HTML-formatted) to display before the table definiton

Every Patient ID must have one and only one entry in tblBAS.

[Expand](#)

Initial REDCap Mapping

leDEA
International epi
Databases to the
World

Data Transfer Protocol for leDEA

Approved by the leDEA Executive Comm

Appendix A (leDEA-DES Tables)
by leDEA Data Harmonization
on April 14

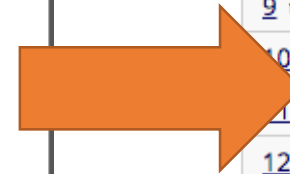
Appendix A: leDEA-DES Tables
(note: the Data Harmonization Working Group will be responsible for recent revisions of this section of the data transfer protocol both tables are approved and designated as either HICDEP* or Non-HICDEP* provide additional documentation of the modified or additional tables.)

Date last revised: Tuesday April 14, 2015 (Text in green represents Designations based on HICDEP version 1.0)

Table Name	Description	Not Yet Designated
tbIART	antiretroviral drugs	
tbIBAS	basic information	
tbICANC	cancer diagnoses	
tbICENTER	site-specific information	
tbICEP	clinical events including serious non-AIDS conditions	X
tbIDELIVERY_CHILD	delivery information related to child	X
tbIDELIVERY_MUM	delivery information related to mother	X
tbIDIS	diseases (CDC-C)	
tbILAB	laboratory tests	X
tbILAB_BP	blood pressure	X
tbILAB_CD4	CD4 measurements	
tbILAB_RES	resistance testing information	X
tbILAB_RES_LVL_1	nucleoside sequence for PRO and RT	X
tbILAB_RES_LVL_2	mutations and positions of PRO and RT sequences	X
tbILAB_RES_LVL_3	resistance result	X
tbILAB_RNA	viral assay	
tbILAB_VIRO	viro /serological Tests	X
tbILTFTU	death and drop-out	
tbIMED	medications	X
tbINNEWBORN	information related to newborns	X
tbINNEWBORN_ABNORM	information related to abnormalities of newborn	X
tbINOVERLAP	participation in other cohorts	X
tbIPREG	general pregnancy-related information	X
tbIPREG_OBS	obstetrical problems	X
tbIPREG_OUT	pregnancy outcome	X
tbIPROGRAM	linking sites to programs	
tbIPREFFILL	prescription refills	X
tbISAMPLES	biological sample storage	X
tbIVIS	viro-related information	

tbICENTER
Relation to HICDEP: NON- HICDEP

Field	Format	Description
CENTER	character	Code for Clinic/Centre/Hospital where patient is seen. Needs to be unique within each region.
PROGRAM	character	Program with which the center is associated
NAME	character	Proper name to identify center
COUNTRY	character	3-letter ISO code
PROVINCE	character	(Optional) Proper name to identify province
DISTRICT	character	(Optional) Proper name to identify district
CITY	character	(Optional) Proper name to identify city
GEOCODE_LAT	Numeric	Latitude
GEOCODE_LON	Numeric	Longitude
RURAL	numeric: 1 = Urban 2 = Mostly urban 3 = Mostly rural 4 = Rural 9 = Unknown	Code for the site situation (facility location)
LEVEL	numeric: 1 = Health centre 2 = District hospital 3 = Regional, provincial or university hospital 9 = Unknown	Code for level of care
ADULTPED	character: "PED," "ADULT," or "BOTH"	Population the center serves
OPEN_D	yyyy-mm-dd	(Optional) Date of opening of dataset: earliest date for which data were included from this site
CLOSE_D	yyyy-mm-dd	Date of closing of dataset
ADD_CENTER	yyyy-mm-dd	Inclusion date: date that the site was added to the cohort
DROP_CENTER	yyyy-mm-dd	(Optional) Exclusion date: date that the site was dropped from the cohort



Record ID	Table Metadata	Variable Metadata
1	tbIART	+
2	tbIART_MUM	+
3	tbIBAS	+
4	tbICANC	+
5	tbICENTER	+
6	tbIDELIVERY_CHILD	+
7	tbIDELIVERY_MUM	+
8	tbIDIS	+
9	tbILAB	+
10	tbILAB_BP	+
11	tbILAB_CD4	+
12	tbILAB_RES	+
13	tbILAB_RES_LVL_2	+
14	tbILAB_RES_LVL_3	+
15	tbILAB_RNA	+
16	tbILAB_VIRO	+
17	tbILTFTU	+
18	tbIMED	+
19	tbINNEWBORN	+
20	tbINNEWBORN_ABNOR	+
21	tbIPREG	+

Word version of the leDEA DES

Part of Machine-Readable DES

(JSON formatted export via REDCap API)

```
[{"record_id": "1", "redcap_repeat_instrument": "", "redcap_repeat_instance": "", "table_name": "tblART", "table_format": "2", "table_definition": "antiretroviral medication", "table_display_y": "1", "table_display_name": "", "table_order": "2", "text_top": "", "text_bottom": "", "table_deprecated__1": "0", "table_deprecated_d": "", "table_metadata_complete": "2", "table_link": "", "variable_name": "", "data_format": "", "description": "", "variable_order": "", "code_text": "", "variable_required__1": "", "variable_deprecated__1": "", "variable_deprecated_d": "", "has_codes": "", "code_format": "", "code_list": "", "code_file": "", "codes_print": "", "variable_metadata_complete": ""}, {"record_id": "1", "redcap_repeat_instrument": "variable metadata", "redcap_repeat_instance": "1", "table_name": "", "table_format": "", "table_definition": "", "table_display_y": "", "table_display_name": "", "table_order": "", "text_top": "", "text_bottom": "", "table_deprecated__1": "", "table_deprecated_d": "", "table_metadata_complete": "", "table_link": "1", "variable_name": "PATIENT", "data_format": "1", "description": "Code to identify patient (Cohort Patient ID)", "variable_order": "1", "code_text": "", "variable_required__1": "1", "variable_deprecated__1": "0", "variable_deprecated_d": "", "has_codes": "0", "code_format": "", "code_list": "", "code_file": "", "codes_print": "", "variable_metadata_complete": "2"}, {"record_id": "1", "redcap_repeat_instrument": "variable metadata", "redcap_repeat_instance": "2", "table_name": "", "table_format": "", "table_definition": "", "table_display_y": "", "table_display_name": "", "table_order": "", "text_top": "", "text_bottom": "", "table_deprecated__1": "", "table_deprecated_d": "", "table_metadata_complete": "", "table_link": "1", "variable_name": "ART ID", "data_format": "1", "description": "ATC code for the patient's ARV medication", "variable_order": "2", "code_text": "Codes are drawn from the Anatomical Therapeutic Chemical (ATC) classification system. Use ART code list maintained by HICDEP: http://www.hicdep.org/wiki/Hicdep\_1.100/TableArt/FieldArtId#CodingTable A searchable index is available at http://www.whocc.no/atc\_ddd\_index/", "variable_required__1": "1", "variable_deprecated__1": "0", "variable_deprecated_d": "", "has_codes": "1", "code_format": "3", "code_list": "", "code_file": "", "codes_print": "0", "variable_metadata_complete": "0"}, {"record_id": "1", "redcap_repeat_instrument": "variable metadata", "redcap_repeat_instance": "3", "table_name": "", "table_format": "", "table_definition": "", "table_display_y": "", "table_display_name": "", "table_order": "", "text_top": "", "text_bottom": "", "table_deprecated__1": "", "table_deprecated_d": "", "table_metadata_complete": "", "table_link": "1", "variable_name": "ART SD", "data_format": "4", "description": "Date of initiation of treatment", "variable_order": "3", "code_text": "", "variable_required__1": "1", "variable_deprecated__1": "0", "variable_deprecated_d": "", "has_codes": "0", "code_format": "", "code_list": "", "code_file": "", "codes_print": "", "variable_metadata_complete": "2"}, {"record_id": "1", "redcap_repeat_instrument": "variable metadata", "redcap_repeat_instance": "4", "table_name": "", "table_format": "", "table_definition": "", "table_display_y": "", "table_display_name": "", "table_order": "", "text_top": "", "text_bottom": "", "table_deprecated__1": "", "table_deprecated_d": "", "table_metadata_complete": "", "table_link": "1", "variable_name": "ART SD A", "data_format": "1", "description": "Date approximation for ART SD", "variable_order": "4", "code_text": "", "variable_required__1": "0", "variable_deprecated__1": "0", "variable_deprecated_d": "", "has_codes": "1", "code_format": "1", "code_list": "< = Before this date | D = Exact to the date | M = Exact to the month | Y = Exact to the year | > = After the date | U = Unknown", "code_file": "", "codes_print": "0", "variable_metadata_complete": "2"}, {"record_id": "1", "redcap_repeat_instrument": "variable metadata", "redcap_repeat_instance": "5", "table_name": "", "table_format": "", "table_definition": "", "table_display_y": "", "table_display_name": "", "table_order": "", "text_top": "", "text_bottom": "", "table_deprecated__1": "", "table_deprecated_d": "", "table_metadata_complete": "", "table_link": "1", "variable_name": "ART ED", "data_format": "4", "description": "Date of stopping of treatment", "variable_order": "5", "code_text": "", "variable_required__1": "0", "variable_deprecated__1": "0", "variab
```

Data Model Browser



IeDEA Data Exchange Standard

This site provides an auto-generated, web-browsable version of the IeDEA Data Exchange Standard (IeDEA DES), a **common data model for sharing observational HIV data** developed by the [International epidemiology Databases to Evaluate AIDS \(IeDEA\)](#). More information on the data model is available on our [GitHub page](#).

IeDEA DES Quick Reference Diagram ([↓ Download PDF](#), last updated 2020-01-03)

IeDEA Multiregional Data Transfer Protocol (Word Document) ([↓ Download](#), last updated 2017-02-17)

Show Draft

Show Deprecated

Data Tables	
Table	Content
tbIART	Antiretroviral medication
tbIART_MUM	Antiretroviral Medication of mother in cases where mother is not enrolled in cohort
tbIBAS	*Required Basic information
tbICANC	Diagnosis of cancer
tbICENTER	Site-specific information
tbIDELIVERY_CHILD	Delivery information related to child
tbIDELIVERY_MUM	Delivery information related to mother

iedeades.org

Website is auto-generated based on the content of the three REDCap projects.



Why this matters

- One “source of truth” for the data model
- One place for users to visit
- Easy to add new variables and codes
- Machine-readable data model can inform additional software tools

- REDCap templates and module code are shared on GitHub
 - <https://github.com/vanderbilt-redcap/data-model-browser>
 - Currently developing an “External Module” version (one-click installation) to share via the REDCap Repository of External Modules

The Harmonist Project



(1) Support for Data Model
leDEA Data Exchange Standard

→ Clear data standards

(2) Consortium Portal
Harmonist Hub

→ Managed research workflow

(3) Data Quality and Sharing Tool
Harmonist Data Toolkit

→ High quality data

The Harmonist Project



(1) Support for Data Model
leDEA Data Exchange Standard

→ Clear data standards

(2) Consortium Portal
Harmonist Hub

→ Managed research workflow

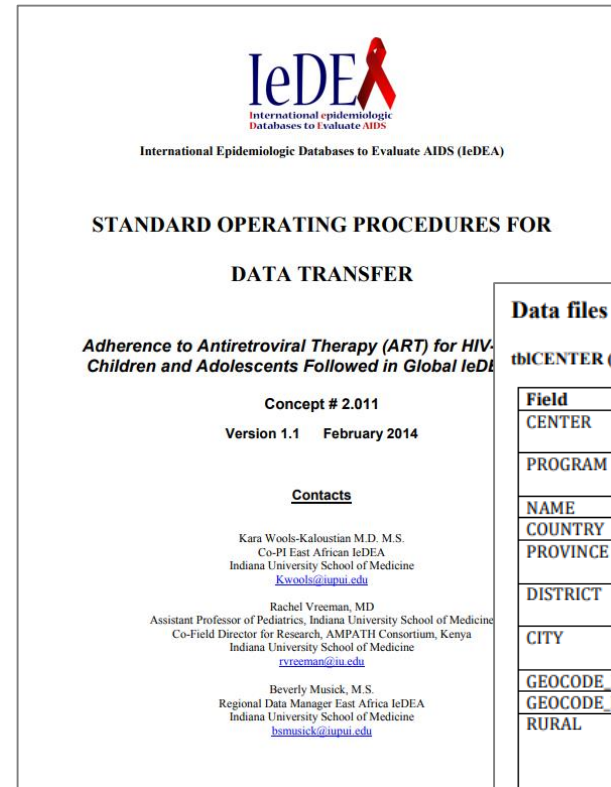
(3) Data Quality
Harmonist Data Toolkit

How can leDEA investigators
request data for projects?

ata

Before...

- IeDEA investigators with an approved concept sheet developed a “Data Request” (Word document).
- Data request documents and due dates were circulated via email.
- A review of the first 4 data requests of 2016 (by different people) found that **none aligned with the IeDEA DES.**



Original data request
(created by hand!)

Data files

tbICENTER (Facility-level information)

Field	Format	Description
CENTER	character	Code for Clinic/Centre/Hospital where patient is seen.
PROGRAM	character	Program with which the center is associated (if applicable)
NAME	character	Proper name to identify center
COUNTRY	character	3-letter ISO code
PROVINCE	character	(Optional) Proper name to identify province
DISTRICT	character	(Optional) Proper name to identify district
CITY	character	(Optional) Proper name to identify city
GEOCODE_LAT	numeric	Latitude
GEOCODE_LON	numeric	Longitude
RURAL	numeric: 1 = Urban 2 = Mostly urban 3 = Mostly rural 4 = Rural 9 = Unknown	Code for the site situation
LEVEL	numeric 1 = Health centre 2 = District hospital 3 = Regional, provincial or university hospital	Code for level of care
ADULTPED	character: "PED," "ADULT", or "BOTH"	Population the center serves
CLOSE_D	yyyy-mm-dd	Date of closing of dataset

International Epidemiologic Databases to Evaluate AIDS

[Home](#) >> [Concepts Overview](#) >> [Concept Status](#)

Concept Status

Lead Investigator
 Region
 Working Group

Concept Number	Title	Lead Investigator
2.042	Outcomes of children and adolescents treated with raltegravir in the IeDEA consortium	Mary David
2.041	Descriptions and outcomes of HIV infected patients treated for tuberculosis without microbiologic confirmation in HIV care programs within the IeDEA consortium	Kara Wool Kaloupek
2.040	Diagnostic, treatment and outcome of extra-pulmonary Tuberculosis in HIV-co-infected adults and children	Mari Balli
2.039	IeDEA-WHO collaboration 2016: Adolescent treatment outcomes in the IeDEA global consortium	Anne Sohr
2.038	IeDEA-WHO Collaboration: Global analysis of delays from ART eligibility to antiretroviral treatment initiation among adults	Denis Nash Ce...
2.037	IeDEA-WHO collaboration: global analysis of the pre-ART cascade and delay from diagnosis to start of antiretroviral therapy in HIV-infected children aged 0-19 years.	Vale Lero
2.036	IeDEA-WHO Collaboration: Global analysis of retention in care in initial HIV care and treatment program	Eliza Zani
2.035	Update of mortality estimates of HIV-infected patients for Spectrum	Cons Yian
2.034	Association between clinic-level factors and individual retention, engagement, and loss to follow-up following ART initiation in the IeDEA collaboration from 2009 to 2014	Bill V - I...


We did have a private consortium website, but

- It only indexed concept sheets, not data requests or due dates
- We couldn't tell which projects had received data and which ones were pending
- Nobody used it because we could never remember our passwords...


Designing the Harmonist Hub



We worked closely with leDEA administrators, PIs, and investigators to design a **web-based consortium management platform** that could

- Index concept sheets and publications
- Coordinate Executive Committee review and voting for concept sheets, abstracts, manuscripts, etc.
- Implement a workflow with dashboards and automatic emails
- Generate metrics based on leDEA's scientific portfolio
- Not a separate website, just another module installed on  **REDCap**
Research Electronic Data Capture

IeDEA Hub: Password-Free Login

IeDEA  **Harmonist Hub**

Access IeDEA Hub

IeDEA Hub (iedeahub.org) is the internal coordination website for the International Epidemiology Databases to Evaluate AIDS. Please visit iedea.org for our public site.

Request Access Link

Email Address

Type your email address.

Get Link

Mon 3/5/2018 1:38 PM
IeDEA Hub <harmonist@vanderbilt.edu>
Review Hub Access
da, Stephany Norah

Here is your link to access the IeDEA Hub:
<https://redcap.vanderbilt.edu/plugins/iedea/index.php?token=MQ>

****This link is unique to you and should not be forwarded to others.**
This link will expire in 7 days. You can request a new link at any time, which will invalidate the old link. If you are logging into the Hub from a public computer, please remember to log out of the Hub to invalidate the link.

If email address is in a list of active users, send an email with a URL that contains a one-time use token.

leDEA Hub: Home Page



Home

Requests **2**

Concepts

Publications

Data Hub

Admin **0**

CN Stephany Duda ▾

Home Page

🔔 The leDEA Hub updated to v2.10 on 2020-03-10. Contact us at harmonist@yumc.org if you encounter any unexpected website behavior. ⚙️

Open Requests **2**

Concept Sheets **1**

Abstracts

Manuscripts **1**

Posters

Fast Track

Other Items

Deadlines and Events ⚙️

Date		Text
2020-07-06	+7 days	AIDS 2020: Virtual begins
2020-07-06	+7 days	TB RePORT-leDEA Collaboration Workshop (register)
2020-08-04	+36 days	leDEA 4 FOA submission due date
2020-11-16	+140 days	HIV Pediatrics Workshop begins (postponed)

Active Data Calls **0**

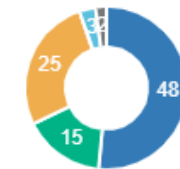
Check the list of active data calls for leDEA multi-regional projects.

[View Data Calls](#)

Hub Metrics [View more](#)

Requests

All leDEA Hub requests by category.



Quick Links ▾

My Requests [View more](#)

No requests available

Recent Activity [View more](#)

✓ Per v [redacted] submitted a **vote** for [Cervical Cancer Prevention and Care Cascade for wo...](#) 2 days ago

leDEA Hub – Data page

[Home](#)[Requests](#) 0[Concepts](#)[Publications](#)[Data Hub](#)[Admin](#) 1[CN](#) [Stephany Duda](#) ▼

Data Hub

The leDEA Hub provides a set of tools that allow you to **request, submit, and retrieve** leDEA data. The purpose of these tools is to make it easier to use the leDEA Data Exchange Standard (DES) and share standardized and quality-checked data in a secure way.



Explore the different types of leDEA data

Coming Soon



Request leDEA data for your approved concept

[Create Data Request](#)



Check and submit data for an active data call

[View Data Calls](#) 0



Retrieve data uploaded for your project

[Download Data](#)

Data News

[View more](#)

No data news to display

Recent Data Activity

[View more](#)

- [↑ Beverly](#) uploaded EA data for [MR180](#). 7 hours ago
- [↑ Karu](#) uploaded CN data for [MR180](#). 8 hours ago
- [↓ Belinda C](#) downloaded CN data for [MR186](#). 19 hours ago

Data Log

[↑ 213](#) [↓ 210](#)

Track uploads and downloads of leDEA patient-level datasets.

Harmonist Data Toolkit



Go to the Toolkit webpage (without a data submission request).

iedeades.org



Browse the leDEA Data Exchange

Creating a Data Request: Step 1



STEP 1: Setup

The **leDEA Data Request Builder** is going to walk you through four steps:

1. Identify the concept for which you want patient-level data,
2. Pick the variables you need from the current version of the [leDEA Data Exchange Standard](#),
3. Define your inclusion/exclusion criteria, suggested data submission date, and who will have permission to download the data, and
4. Preview your draft Data Request.

At the end, we'll guide you through **making your draft public** so you can share it with the leDEA Data Harmonization Working Group for feedback.

Select Your Concept:

MR116 - Harmonist Data Toolkit Development: Request for leDEA DES Datasets fr ▾

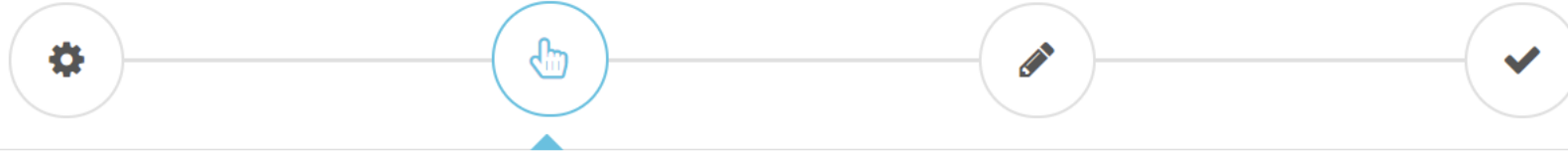
For test requests, select MR000.

Setup Type:

Create new data request Load draft

Save and continue

Creating a Data Request: Step 2



STEP 2: Choose Variables

MR116: Data Request #63

Select the [leDEA DES](#) tables and variables you need for your analysis.

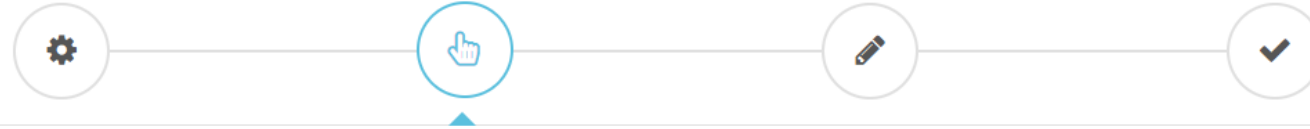
Keep in mind that the leDEA DES only specifies how data are transferred among regions; it does not define what data each region collects. Therefore **not all these variables will be available from leDEA regions**. Known variable availability (based on previous requests) is marked in the "Availability" column. The Data Harmonization Working Group will help you refine these selections.

Select	Field	Availability	Format	Description
<input checked="" type="checkbox"/>	PATIENT		Character or Numeric	Code to identify patient (Cohort Patient ID)
<input checked="" type="checkbox"/>	PROGRAM		Character	Direct one-to-one link from the patient to their program
<input type="checkbox"/>	BIRTH_D		YYYY-MM-DD	Birth date
<input type="checkbox"/>	BIRTH_D_A		Character < = Before this date D = Exact to the date M = Exact to the month Y = Exact to the year > = After the date	Date approximation for BIRTH_D

tbIBAS 2 Select All

Previous Save and continue

Creating a Data Request: Step 2.5



STEP 2: Choose Variables

Select the [leDEA DES](#) tables and variables you need for your analysis.

Keep in mind that the leDEA DES only specifies how data are transferred among regions; it does not define what data each region collects. Therefore **not all these variables will be available from leDEA regions**. Known variable availability (based on previous requests) is marked in the "Availability" column. The Data Harmonization Working Group will help you refine these selections.

tbIBAS	15	<input type="checkbox"/> Select All	▾
tbILTFU	8	<input type="checkbox"/> Select All	▾
tbIVIS	4	<input type="checkbox"/> Select All	▾
tbILAB_CD4	5	<input checked="" type="checkbox"/> Select All	▾
tbILAB_RNA	4	<input type="checkbox"/> Select All	▾
tbIART	7	<input type="checkbox"/> Select All	▾
tbIDIS	5	<input type="checkbox"/> Select All	▾
tbICENTER	14	<input type="checkbox"/> Select All	▾
tbIPROGRAM	2	<input checked="" type="checkbox"/> Select All	▾
tbIART_MUM		<input type="checkbox"/> Select All	▾
tbICANG		<input type="checkbox"/> Select All	▾

Previous

Save and continue

Step 3



STEP 3: Add Details

The form below collects some extra information needed to generate the Data Request.

Text Descriptions

Inclusion criteria *(list variable names if possible)*

↶	↷	B	<i>I</i>	<u>U</u>	☰	☰	☰	☰	🔗	🔗	Ω
☰	☰	<u>I</u>	x ₂	x ²	☰	☰	A	A	🏠	<>	

- Include all HIV-positive older adolescents and adults (>=16 years of age at enrollment) in the IeDEA cohorts from 1996 onward.
- Both ART-naïve and non-naïve patients at enrollment are eligible.
- Include all years of data from CENTERS that diagnose cryptococcal meningitis (CM) and provide data on CM/extrapulmonary cryptococcosis, even if the CENTER only started diagnosing CM recently. (e.g., If a clinic started in 2013, but CM diagnoses only began in 2015, include all data from 2013 onward.)
- Include both patients with CM and patients without CM from these CENTERS.

Exclusion criteria *(list variable names if possible)*

↶	↷	B	<i>I</i>	<u>U</u>	☰	☰	☰	☰	🔗	🔗	Ω
☰	☰	<u>I</u>	x ₂	x ²	☰	☰	A	A	🏠	<>	


- Exclude CENTERS that do not systematically collect any opportunistic infection (OI) data.
- Exclude CENTERS that have never diagnosed CM.
- Exclude CENTERS that do not collect or provide data on CM for IeDEA (This means exclude CENTERS that do not ever collect or provide DIS_ID = 'CRCO', see definition of CM below.)

Save & Stay

Step 3.5

New content guided by the leDEA Data Harmonization Working Group (not part of Word document Data Requests)






Data Format and Access








Due Date **required* 

Preferred File Format

- CSV
- SAS
- Stata
- SPSS
- R
- Other (describe)

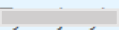
File Format Details

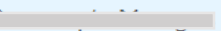
← → **B** *I* U      Ω


  x_2 x^2      <>

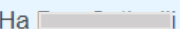
Choose your Data Downloaders **required* ▾

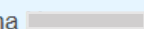
Judy Lewis

Olga 

Eva E 

Nicola 

Cam Ha 

Brenna 

Drag and Drop

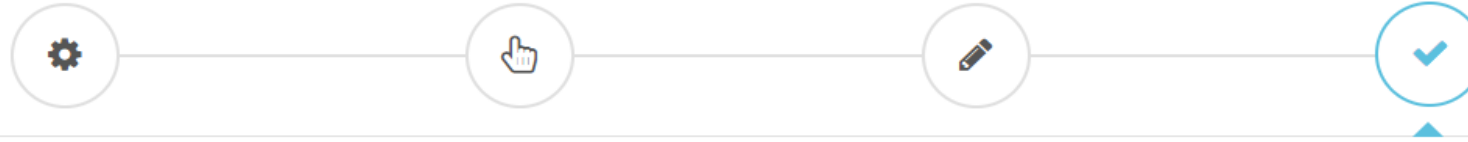
↔

Stephany Duda

Not sure about Data Downloaders yet.

[Save & Stay](#)

Step 4



STEP 4: Preview Data Request

Preview the document below. You can return to any previous step to edit the draft. The **next step will save the draft Data Request to your library and create a PDF**. You can still return edit the document afterwards and regenerate a new PDF.



International epidemiology Databases to Evaluate AIDS (IeDEA)

DATA TRANSFER REQUEST – MR131

Clinical Outcomes in HIV-positive patients with Cryptococcal Meningitis in the IeDEA global cohort consortium

Data Due: 17 October 2018

Research Contact(s)

Anna F.

user.email@example.com

Brenda C.

user.email@example.com

Data Contact

Previous


Save and create PDF


Steps Complete ✓

Your Data Request has been generated successfully. You can review and download the PDF below or download a ZIP file with an HTML and a PDF version. If you need to make changes, [you can go back to edit your data request](#).

The **next step is to make your draft public and request a review** by the leDEA Data Harmonization Working Group. The working group chairs may ask to meet with you to revise this document together, or may ask if you can discuss this data request on the next WG call.

[View in My Drafts](#)[Send for Review](#)[< Back to Edit Data Request](#)

Data Request Download ZIP  ▾



International epidemiology Databases to Evaluate AIDS (leDEA)

DATA TRANSFER REQUEST – MR131

Clinical Outcomes in HIV-positive patients with Cryptococcal Meningitis in the leDEA global cohort consortium



International epidemiology Databases to Evaluate AIDS (IeDEA)

DATA TRANSFER REQUEST – MR131

Clinical Outcomes in HIV-positive patients with Cryptococcal Meningitis in the IeDEA global cohort consortium

Data Due: 17 October 2018

Research Contact(s)

Anna [redacted]
user.email@example.com

Brenda [redacted]
user.email@example.com

Data Contact

Stephany Duda
user.email@example.com

Data Request Version: 07 March 2019

Automatically Generated Data Request

tbILTFU - Death and dropout

Table with 3 columns: Field, Format, Description. Rows include PATIENT, DROP_Y, DROP_D, DROP_D_A, DROP_RS, DEATH_Y, DEATH_D, and DEATH_D_A.



Why this matters

- No more manually assembled Word documents! Point-and-click data request builder **saves time**.
 - The list of requestable variables is pulled from the machine-readable leDEA DES. The tool automatically adds required variables if not selected
 - Therefore all data requests will be **compliant with the standard**.
 - One “source of truth” for data requests (on the Hub)
 - All resources in GitHub and the REDCapRepo
- And...
- Data requests are stored in a **machine-readable format** now too.

[< Back to Data](#)

Check and Submit Data

leDEA data is submitted securely through the **Harmonist Data Toolkit**. The Toolkit will

1. Scan your uploaded files to run data format and quality checks,
2. Auto-generate dataset reports for you to download, and
3. Allow data upload to the **secure Harmonist cloud** (for data transfer only).

Data files should be formatted according to the [leDEA Data Exchange Standard \(leDEA DES\)](#).

The following leDEA Concepts have active Data Requests. Please review the request details or select the request for which you wish to upload data.

Not started
 Partial Data
 Complete Data
 Data Not Available
 Region Not Requested
 Other Status

[View Data Activity Log](#) | [View Past Data Calls](#)

Open Data Calls

No Open Data Calls available

Completed Data Calls

Due Date	Data Request Details	Data Contact	AP	CA	CN	EA	NA	SA	WA	Actions
2021-09-08 -23 days	MR180 Infectious and non-infectious causes of liver disease among adults living with HIV from low and middle income countries, the Sentinel Research Network of leDEA Draft ID: 54 Data Request MR180 Concept 📄	Kare... (WA)	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input type="button" value="Upload Data"/> <input type="button" value="Change Status"/>
2021-09-15 -16 days	MR181 Diagnostic value of serological markers for liver steatosis and fibrosis in people living with HIV: a multicenter study from the leDEA Network Draft ID: 55 Data Request MR181 Concept 📄	H... (CN)	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input type="button" value="Upload Data"/> <input type="button" value="Change Status"/>
2021-09-20 -11 days	MR185 Associations of mental health and substance use disorders with virological failure by sex among aging adults living with HIV from low/middle-income countries, the Sentinel Research Network of leDEA Draft ID: 60 Data Request MR185 Concept 📄	Kath... (CA)	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input type="button" value="Upload Data"/> <input type="button" value="Change Status"/>

The Harmonist Project



(1) Support for Data Model
leDEA Data Exchange Standard

→ Clear data standards






(2) Consortium Portal
Harmonist Hub

→ Managed research workflow

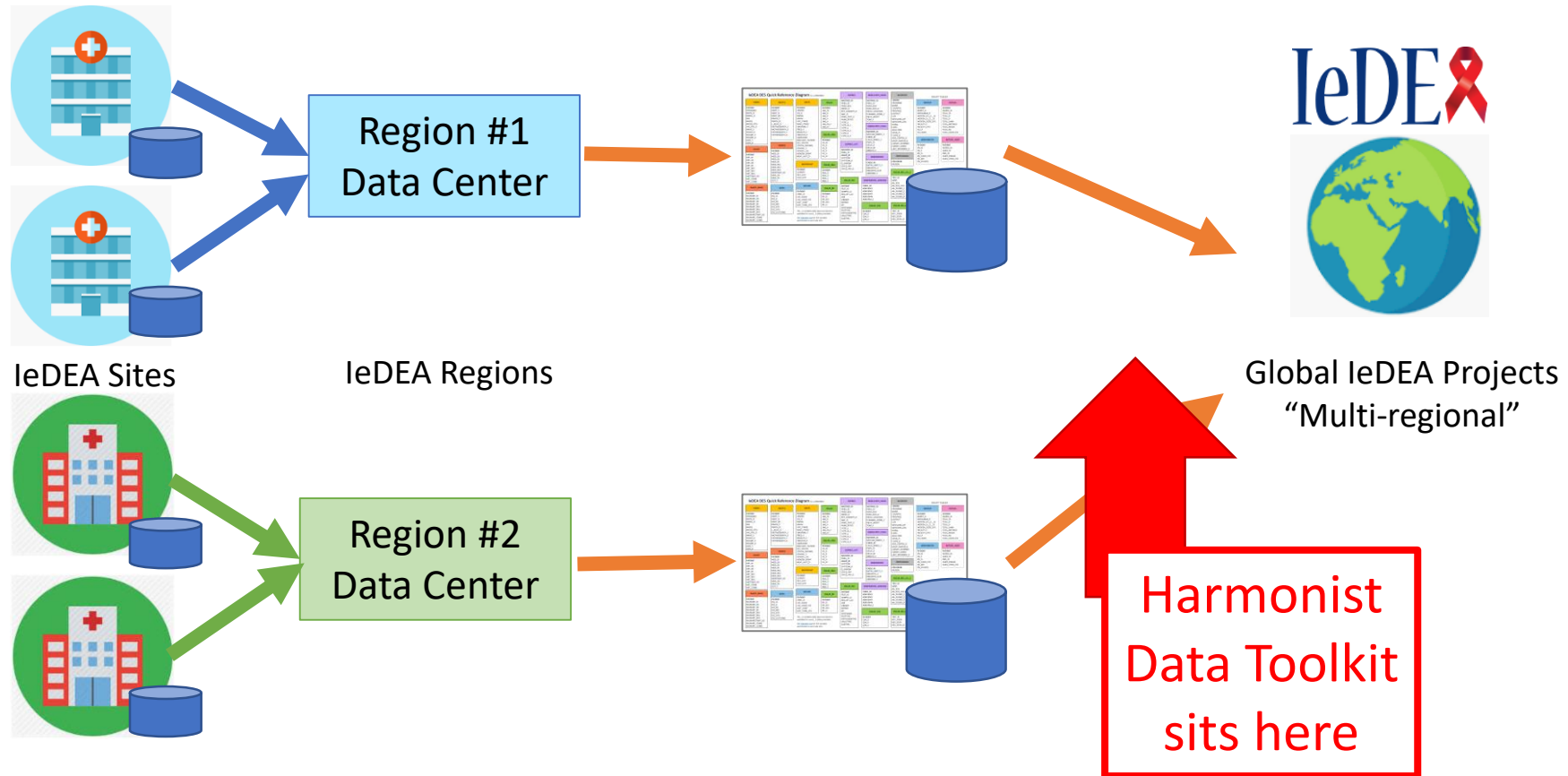
(3) Data Quality and Sharing Tool
Harmonist Data Toolkit

→ High quality data

leDEA Harmonist Data Toolkit

- Collaborative project with all seven **leDE**  regions
- Web application
- Developed with freely available/open source tools (R, Shiny, REDCap)
- Designed to evolve with data exchange standard
- Features:
 -  Ensures datasets conform to common data model
 -  Performs data quality checks
 -  Generates reproducible reports
 -  Submits approved datasets to secure cloud storage

Flow of leDEA Data for Global Projects



Workflow Begins in leDEA Project Portal (leDEA Hub)

The screenshot displays the leDEA Hub interface. The top navigation bar includes 'Home', 'Requests 2', 'Concepts', 'Publications', and 'Data'. The user 'Judy Lewis' is logged in. The main content area is titled 'Data' and provides instructions on how to use the leDEA Data Exchange Standard (DES). Below this, there are four cards: 'Explore the different types of leDEA data' (Coming Soon), 'Request leDEA data for your approved concept' (Request Data), 'Check and submit data for an active data call' (Submit Data), and 'Retrieve data uploaded for your project' (Retrieve Data). A blue arrow labeled '1' points to the 'Submit Data' button. Below the cards, there is a section titled 'Check and Submit Data' with instructions on how to submit data using the Harmonist Data Toolkit. A purple arrow labeled '2' points to the 'Upload Data' button in the 'Active Data Calls' table.

Check and Submit Data

leDEA data is submitted securely through the **Harmonist Data Toolkit**. The toolkit will:

1. Scan your uploaded files to run data format and quality checks,
2. Auto-generate dataset reports for you to download, and
3. Allow data upload to the **secure Harmonist cloud** (for data transfer only).

Data files should be formatted according to the [leDEA Data Exchange Standard \(leDEA DES\)](#).

The following leDEA Concepts have active Data Requests. Please review the request details or select the request for which you wish to upload data.

[View Upload History](#) | [View Past Data Calls](#)

Active Data Calls							
Due Date	Concept	Data Contact	Downloaders Assigned	Data Request	PDF	TT	Actions
2019-03-12 -3 days	MR116	Judy Lewis (TT)	2	2. Data Toolkit Practice Request A			Upload Data



1. Upload Files to IeDEA Toolkit

Introduction to Toolkit

ACTIONS

MR116

STEP 1: Upload files

STEP 2: Check data

STEP 3: Create summary

STEP 4: Submit data

TOOLS

Visualize data

Help

Provide feedback

STEP 1 Upload files

Choose the files containing your IeDEA tables to check for data quality. After files are uploaded, review the table summarizing uploaded files and variables.

MR116 Active Data Request

Title	Harmonist Data Toolkit Development: Request for IeDEA DES Datasets from All Regions
Hub Pages	MR116 on Hub , Data Specification
Requested Tables	tblBAS tblLTFU tblIVIS tblLAB_CD4 tblLAB_RNA tblCENTER tblPROGRAM
Requested Data Format	SAS
Contacts	<ul style="list-style-type: none">Judy Lewis (TT), Vanderbilt UniversityStephany Duda (CN), Vanderbilt UniversityJudy Lewis (TT) (Data contact), Vanderbilt University
Data Downloaders	<ul style="list-style-type: none">Stephany Duda (CN), Vanderbilt UniversityJudy Lewis (TT), Vanderbilt University

Select Data Files

Upload data in the [IeDEA Data Exchange Standard \(IeDEA DES\)](#) format. tblBAS is required.

Allowed file formats include **CSV, SAS, Stata, SPSS, or a ZIP containing multiple files** of this type.

Select a single ZIP file or multiple files with **Ctrl+Click**

Data files

No file selected




Use Sample Dataset

Launch the Toolkit with a sample dataset (fake data) for practice, testing, and demonstrations.

The sample dataset contains 48 intentionally error-filled records representing the following IeDEA DES tables: tblBAS tblLTFU tblIVIS

tblIART tblLAB tblLAB_BP

2. File Review/Data Quality Checks

 **Harmonist Data Toolkit**

Introduction to Toolkit

ACTIONS MR116


STEP 1: Upload files


STEP 2: Check data


STEP 3: Create summary

STEP 4: Submit data

TOOLS

 Visualize data

 Help

 Provide feedback

[Exit Data Toolkit](#)

STEP 1 Upload files

Choose the files containing your IeDEA tables to check for data quality. After files are uploaded, review the table summarizing uploaded files and variables.

MR116 Active Data Request +

Missing Variables

The following variables requested by MR116 were not found:

- tbIVIS** NEXT_VISIT_D, NEXT_VISIT_D_A
- tbIPROGRAM** Table missing (2 variables)

Summary of Uploaded IeDEA Tables


Table	Records	IeDEA DES Variables	Extra Variables
tbIBAS	28089	PATIENT, PROGRAM, BIRTH_D, BIRTH_D_A, ENROL_D, ENROL_D_A, GENDER, MODE, NAIVE_Y, PROPH_Y, RECART_Y, RECART_D, RECART_D_A, HAART_D, HAART_D_A, AIDS_Y, AIDS_D, AIDS_D_A	CENTER, YEAR_ENROLLED, SITE
tbILTFU	28089	PATIENT, DROP_Y, DROP_D, DROP_D_A, DROP_RS, DEATH_Y, DEATH_D, DEATH_D_A, L_ALIVE_D, L_ALIVE_D_A	CENTER, SITE
tbIVIS	1084237	PATIENT, CENTER, VIS_D, VIS_D_A, CDC_STAGE, WHO_STAGE, PREG_Y	GENDER_ID, SITE
tbILAB_CD4	298041	PATIENT, CD4_D, CD4_D_A, CD4_V, CD4_U	CENTER, SITE
tbILAB_RNA	159234	PATIENT, RNA_D, RNA_D_A, RNA_V	CENTER, SITE
tbIART	140435	PATIENT, ART_ID, ART_SD, ART_SD_A, ART_ED, ART_ED_A	CENTER, SITE
tbIDIS	3750	PATIENT, DIS_ID, DIS_D, DIS_D_A, DIS_ED, DIS_ED_A, DIS_OUTCOME	CENTER, SITE
tbICENTER	9	CENTER, PROGRAM, NAME, COUNTRY, PROVINCE, DISTRICT, CITY, GEOCODE_LAT, GEOCODE_LON, RURAL, LEVEL, ADULTPED, OPEN_D, CLOSE_D	REGION, ADD_CENTER_D, DROP_CENTER_D, SITE

Data Quality Checks

The toolkit is checking your dataset.

- ✓ Files read and formatted
- ✓ Checking numeric values
- ✓ Checking date logic and date format
- ✓ Checking for missing values
- ✓ Checking coded variables
- ✓ Checking lab values
- ✓ Checking tables for Patient IDs that don't exist in tblBAS
- ✓ Comparing all dates to BIRTH_D, DEATH_D, DROP_D, and L_ALIVE_D
- ✓ Checking for duplicate records in tables
- ✓ Checking for correct sequence for start dates and end dates
- ✓ Checking for possible typos in HEIGH: height values that decrease
- ✓ Checking for conflicting WHO_STAGE on the same date
- ⚙️ Checking for conflicting CDC_STAGE on the same date (*Quality check # 12 of 16*)

3. Data Quality Results

**Harmonist Data Toolkit**

Introduction to Toolkit

ACTIONS MR116

STEP 1: Upload files

STEP 2: Check data

STEP 3: Create summary

STEP 4: Submit data

TOOLS

Visualize data

Help

Provide feedback

Exit Data Toolkit

STEP 2 Check data

View interactive summary of errors and download detailed results of data quality checks to review offline.

Error Summary by Table

[Download error detail CSV](#)

tbIBAS **10** | tbiLTFU **14** | tbiVIS **2** | tbiLAB_CD4 **1** | tbiLAB_RNA **7** | tbiIART **85** | tbiDIS **✓** | tbiCENTER **✓** | Invalid Codes **28**

Show **10** entries Search:

Error description	Severity	Count	
Future date: ENROL_D	Error	1	View Detail
Invalid Code: MODE	Error	1	View Detail
Invalid Code: RECARD_D_A	Error	2	View Detail
Invalid Code: HAART_D_A	Error	2	View Detail
BIRTH_D before 1920	Warn	3	View Detail
Date before 1980: AIDS_D	Warn	1	View Detail

Showing 1 to 6 of 6 entries Previous **1** Next

Continue to Summary

Error checks completed

Your dataset contains **114 total errors in 12 error categories** including **28 invalid codes**


If you have already reviewed the content of the dataset, proceed to the next step to **generate a summary of the data**.

[Continue to Step 3](#)

Restart session

Start over and upload a **revised or different dataset**.

[Upload new dataset](#)



4. Reproducible Reports

STEP 3 Create summary

Generate and download customized reports summarizing uploaded dataset.

Customize Summary Report

File format for report

PDF

Select report content

- Summary statistics of tables
- Summary of data quality checks
- Histograms of dates

Data subgroup(s) for report

All

(Optional) Short title for report heading

Date Histogram Options

Choose years to include in histograms

Years: 2000 - present

Generate summary PDF report

IeDE Harmonist Data Toolkit Report

Dataset submitted from IeDEA Region: Harmonist Test

Report date: 2019-03-07

Dataset Summary

Total number of patients in dataset: 28089

Table 1: Table Summary

Table	Records	Patients	Age at Enrollment					
			0-4	5-9	10-14	15-19	20-24	Adults 25+
tblBAS	28089	28089	0	0	0	971	3342	23776
tblLTFU	28089	28089	0	0	0	971	3342	23776
tblVIS	1084237	26820	0	0	0	963	3270	22587
tblLAB_CD4	298041	27304	0	0	0	945	3273	23086
tblLAB_RNA	159234	15524	0	0	0	471	2010	13043
tblART	140435	25840	0	0	0	889	2971	21980
tblDIS	3750	3215	0	0	0	100	345	2770

Table 2: SITE in Dataset

SITE	Patients	tblLTFU	tblVIS	tblLAB_CD4	tblLAB_RNA	tblART	tblDIS
Hogwarts	594	594	537	557	461	536	36
Hufflepuff	11763	11763	11744	11447	239	11530	1933
Muggleton	5248	5248	5240	5145	4898	4397	650
Potterburg	3208	3208	3208	3168	3116	2723	69
Ravenclaw	1079						
Slytherin	458						
Snapetown	4099						
Wizardville	1640						

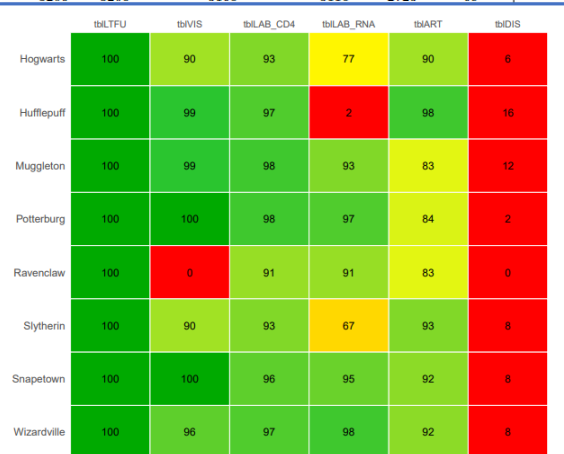


Figure 1: Percent of Patients from tblBAS Included

Table 3: Number of observations per year

Variable	< 2011	2011	2012	2013	2014	2015	2016	2017	2018	2019	Total
Enrolled	14728	2517	2527	2850	2391	1439	1332	295	0	0	28088
Visits	411772	115301	117341	124501	118630	120098	71094	5500	0	0	1084237
Deaths	1064	249	232	234	170	216	141	10	0	0	2316
Transfers Out	193	19	19	48	94	145	139	14	0	0	671
Viral Load	59549	15201	15740	16776	16726	17731	15799	1708	0	0	159230
CD4	122521	30414	30781	30465	33865	26465	20152	3378	0	0	298041

Histograms of important dates by SITE

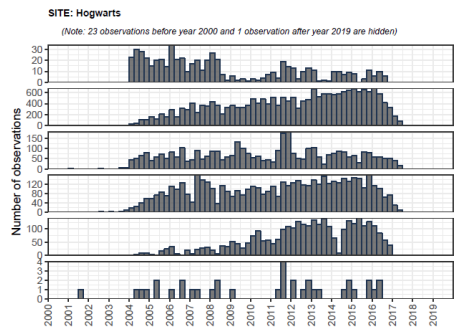



Table 4: Summary statistics from uploaded tables

Value	Count	Percent
Sex		
Male	17453	62.1
Female	10636	37.9
Missing	0	0.0
Deceased		
No	25773	91.8
Yes	2316	8.2
Missing	0	0.0
Treatment Naive at Enrollment		
No	2940	10.5
Yes	25149	89.5
Missing	0	0.0
Receive Antiretroviral Therapy (ART)		
Yes	25840	92.0
Missing	2249	8.0



Harmonist Data Toolkit

Introduction to Toolkit

ACTIONS MR116

- STEP 1: Upload files
- STEP 2: Check data
- STEP 3: Create summary
- STEP 4: Submit data**

TOOLS

- Visualize data
- Help
- Provide feedback

Exit Data Toolkit

STEP 4 Submit data

Submit dataset for selected concept.

Transfer Data for leDEA Concept

Ready to transfer data

Dataset summary:

- 28089 unique patient records included.
- 8 leDEA DES tables included.
- Missing 4 variables requested by MR116 across 2 tables.**
- 114 potential data quality issues** detected.


After transfer:

- Uploaded data will be **stored for 30 days.**
- Data will be automatically deleted after 30 days. You can manually delete your uploaded datasets via the leDEA Hub.
- Approved data downloaders will be able to retrieve your data through the Hub. (Downloaders: Stephany Duda, Judy Lewis)

Message to accompany your file upload (visible to Data Downloaders on the Hub):

Click below to submit your data to secure cloud storage to be retrieved by Judy Lewis

Submit Data



6. Recipient Downloads Data

leDE [Home](#) [Requests 2](#) [Concepts](#) [Publications](#) [Data](#) TT Judy Lewis ▾

[< Back to Data](#)

Download security:
Login with multifactor authentication

Retrieve Data

All leDEA data requests that you have access to are displayed here. Uncollapse the menus to see individual file downloads and details. Downloads expire after 30 days. If you expect to have access to datasets that are not listed here, you may not be listed as a permitted Data Downloader on that data request. Contact the project lead and the Harmonist team to request permission.



MR116 | Data Request #2 -10 days until due ↓ 2 ▾

Title: [Harmonist Data Toolkit Development: Request for leDEA DES Datasets from All Regions](#) | [Data Request #2](#)

Data Contact: Judy Lewis (judy [redacted])

Data Due: 12 March 2019

Upload Date	Region	Submitted By	Filename	PDF	Expires On	Actions
+ 2019-03-20 15:06:42	TT	Judy Lewis	MR116_TT_Lewis_201903201506.zip		21 April 2019 +30 days	Download
+ 2019-03-18 12:01:32	CN	Hilary [redacted]	MR116_CN_Vansell_201903181201.zip		21 April 2019 +30 days	Download

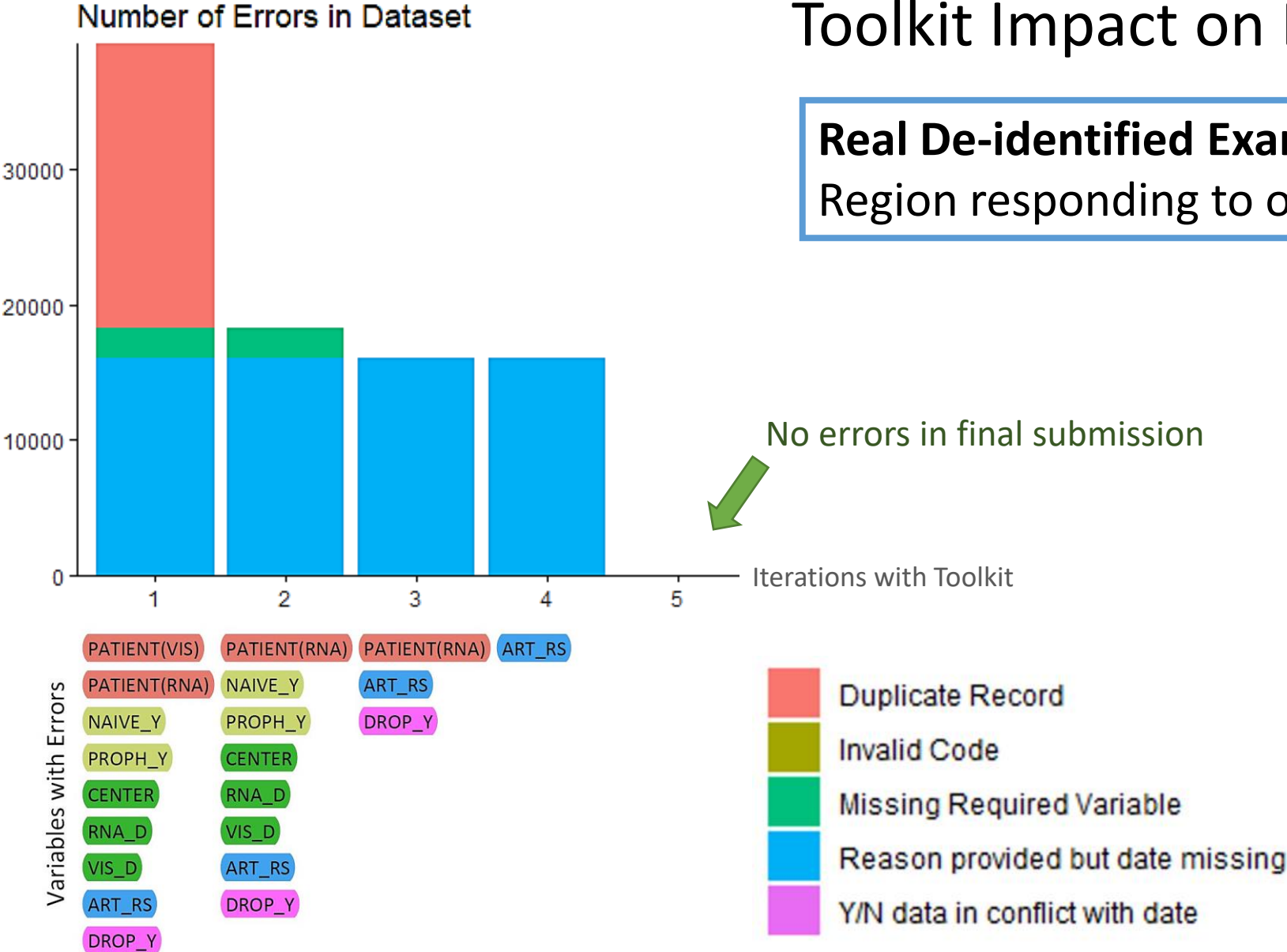


Toolkit Impact on Data Quality

- The Toolkit was **released in March 2019**
 - > 800 datasets have been checked and summarized
 - 1,800 to 986,089 patients per dataset
 - 81 secure file transfers for 17 multiregional leDEA projects
- Regional data managers **uploaded datasets and reviewed data quality results multiple times** before final transfer
- Data managers used Toolkit data quality reports to **improve datasets** before transfer
- The number and types of **errors decreased** with each iteration of Toolkit use, average of 61% decrease in errors

Toolkit Impact on Data Quality

Real De-identified Example:
Region responding to one data request



Why This Matters



- **High quality data** is essential to meaningful research.
- Tools like this can help:
 - **Improve adherence** to data model and standards
 - **Reduce time** for data preparation and checking
 - Highlight data **completeness** and **coding** problems
 - Increase **security** and **uniform workflow** for data exchange
- Generalized design allows software to be adapted to other domains.

Lessons Learned



- **Close collaboration** with stakeholders and users is key
 - Frequent contact with stakeholders (e.g., admins, data managers, working groups)
 - Structured testing and training exercises with users
 - In-person (now virtual) meetings to collaborate on design
- Using **REDCap as a platform** facilitates development and sharing
- Defining details in REDCap and using the REDCap API make it possible to design tools that **adapt with the changing data model**

Tools have been well received by leDEA; multiregional concept proposals and data requests have doubled.

Development & Opportunities



- Future tasks
 - Expand features
 - Gather insights from other consortia
 - Improve generalizability of tools



- We welcome all questions and input!
 - Contact us at harmonist@vumc.org



- Code available
 - github.com/leDEA/Harmonist
 - github.com/vanderbilt-redcap
 - redcap.vanderbilt.edu/consortium/modules/index.php

Thank You

- Harmonist and REDCap technical teams
- leDEA leadership, Data Harmonization Working Group, and collaborators
- Other consortia helping to make our processes generalizable!

Research reported in this publication was supported by the National Institute Of Allergy and Infectious Diseases of the National Institutes of Health under Award Number R24AI124872. The content is solely the responsibility of the authors and does not necessarily represent the official views of the National Institutes of Health.