Cohort Profile: ReCoDID Consortium’s Harmonized Acute Febrile Illness Arbovirus

Meta-Cohort

ReCoDID Pooled Cohort Profile

Gustavo Gómez, Heather Hufstedler, Carlos Montenegro, Yannik Roell, Anyela Lozano, Adriana Tami, Tereza Magalhaes, Ernesto T.A. Marques, Angel Balmaseda, Guilherme Calvet, Eva Harris, Patricia Brasil, Victor Herrera, Luis Villar, Lauren Maxwell, Thomas Jaenisch, ReCoDID Arbovirus harmonization study group

1. Grupo de Epidemiología Clínica, Universidad Industrial de Santander, Bucaramanga, Colombia
2. Heidelberg Institute of Global Health (HIGH), Heidelberg University Hospital, Heidelberg, Germany
3. Sustainable Sciences Institute, Managua, Nicaragua
4. Center for Global Health, Colorado School of Public Health, Aurora, Colorado, United States of America
5. University of Groningen, University Medical Center Groningen, Department of Medical Microbiology and Infection Prevention, Groningen, The Netherlands
7. Center for Vector-Borne Infectious Diseases (CVID), Department of Microbiology, Immunology and Pathology (MIP), College of Veterinary Medicine and Biomedical Sciences, Fort Collins, CO, USA

NOTE: This preprint reports new research that has not been certified by peer review and should not be used to guide clinical practice.
8. Department of Preventive and Social Medicine, School of Medicine, Universidade Federal da Bahia, Salvador, BA, Brazil

9. Institute Aggeu Magalhães, Department of Virology and Experimental Therapeutics, Fundação Oswaldo Cruz, FIOCRUZ-PE, Recife Brazil

10. University of Pittsburgh, School of Public Health, Department of Infectious Diseases and Microbiology, Pittsburgh, PA, USA


12. Oswaldo Cruz Foundation (Fiocruz), Rio de Janeiro, Brasil

13. Division of Infectious Diseases, School of Public Health, University of California Berkeley, Berkeley, CA, USA

14. Centro de Atención y Diagnóstico de Enfermedades Infecciosas, Bucaramanga, Colombia

15. Section Clinical Tropical Medicine, Department for Infectious Diseases, Heidelberg University Hospital, Heidelberg, Germany

° Co-first authors, in alphabetical order

* Membership of the ReCoDID Arbovirus harmonization study group is provided in the acknowledgements
Abstract

Infectious disease (ID) cohorts are key to advancing public health surveillance, public policies and pandemic responses. Unfortunately, ID cohorts often lack funding to store and share clinical-epidemiological data (CE) and high-dimensional laboratory (HDL) data long-term, which is evident when the link between these data elements is not kept up to date. This becomes particularly apparent when smaller cohorts fail to successfully address the initial scientific objectives due to limited case numbers, which also limits the potential of pooling for these studies to monitor long-term cross-disease interactions within and across populations. To facilitate advancements in cross-population inference and reuse of cohort data, the European Commission (EC) and the Canadian Institutes of Health Research, Institute of Genetics (CIHR-IG) funded the ReCoDID (Reconciliation of Cohort Data for Infectious Diseases) Consortium to store and share harmonized and standardized CE and HDL data on a federated platform and also provide innovative statistical tools to conduct meta-analyses of the individual patient data. Here we describe the harmonization of CE data from nine arbovirus (arthropod-borne viruses) cohorts in Latin America, which serve as a starting point for the ReCoDID meta-cohort. CE data was retrospectively harmonized using Maelstrom’s methodology and standardized to Clinical Data Interchange Standards Consortium (CDISC).

This meta-cohort will facilitate various joint research projects, e.g., on immunological interactions between sequential flavivirus infections and for the evaluation of potential biomarkers for severe arboviral disease.
Introduction

The meta-cohort consists of participant-level data and descriptive metadata from nine studies from five countries (Brazil, Colombia, El Salvador, Nicaragua, and Venezuela). All studies were established to study arbovirus (arthropod-borne viruses) infections in the population, with some cohorts enrolling maternal-infant pairs, or either children or pregnant women.

Most studies did not recruit participants through study-initiated contacts (e.g., emails, calls, letters, etc.). Instead, the vast majority of participants in each cohort were directly referred to the study from the health unit where they were seeking care. A small fraction of participants contacted the cohorts directly to participate because participation was associated with access to more routine care or additional screenings, which can be seen as a benefit in resource-poor settings.

The interactions between immune responses caused by different patterns of exposure over time to the four dengue virus serotypes (DENV 1-4) and Zika virus (ZIKV) have attracted considerable attention - for example as a mechanism to explain the heterogeneity in severe dengue but also in the severe outcomes seen during and after the ZIKV epidemic in Latin America(1–5). The investigation of these interactions between closely related members of the flaviviridae requires large sample sizes and the inclusion of populations with exposures to different sequences of pathogens, resulting in heterogeneous immune profiles.

The ReCoDID Consortium, funded by the EC and CIHR-IG(6), aims to provide ID researchers with harmonized participant-level data and metadata resources as well as analysis tools to facilitate pooled analysis projects, i.e. to advance our knowledge on the effects of prior exposure on the immune response to subsequent epidemics at the population and individual level and to inform personalized medicine approaches to diagnosis and treatment of infections. To facilitate cross-
study inference in the context of emerging IDs, ReCoDID researchers created a platform to extract individual-level CE and HDL data from existing cohorts, and harmonize this data according to a specific standard. ReCoDID focuses current harmonization efforts on arbovirus and SARS-CoV-2 cohort data, but hopes to expand these services to other IDs in the future. This cohort profile provides an overview of the newly-created arbovirus meta-cohort from five countries: Brazil, Colombia, El Salvador, Nicaragua, and Venezuela. Acute and post-acute samples were collected from each study. Information extracted from samples vary from study to study and include DENV molecular tests, DENV serotype, DENV viral load, ZIKV molecular test, ZIKV viral load, CHIKV molecular tests and CHIKV viral load. Height, weight, birthdate, negative health outcomes associated with severe dengue (such as occurrence of bleeding, e.g.) and required interventions have also been collected. The ReCoDID Consortium aims to build a data sharing platform for linking CE data to HDL data (e.g., human and pathogen genomic data, human metabolomic and immunomics data) at the participant level that are collected from ID focused cohorts. While ReCoDID is working to share data related to other disease types, this paper describes the acute febrile illness (AFI) meta-cohort that includes, at the date of publishing this article, data from nine studies that have committed to sharing CE and HDL data from their arbovirus cohorts. All participating cohorts have submitted genomic sequences of the DENV virus to ReCoDID, except for the cohorts in Nicaragua and the Cohort of Symptomatic Pregnant Women in Brazil. Two studies— PHBDC and IDAMS cohorts— have also agreed to share genomic sequences for CHIKV and ZIKV viruses. Most participating cohorts collected and stored blood samples; Nicaragua also collected urine and saliva samples. Cohorts varied in their inclusion criteria— some admitting only patients who present with fever, some used rash or red eyes, while others admitted patients who presented with fever and/or rash. With the introduction of ZIKV and
CHIKV, the PHBDC study chose to adjust the inclusion criteria in order to admit patients who presented with rash, headache and/or fever. Altogether, the longitudinal data of the meta-cohort covers over 18,000 patients (pediatric and adults), in both inpatient and outpatient settings from five countries (Brazil, Colombia, El Salvador, Nicaragua, and Venezuela). Data collection start and end dates vary between cohorts, but ranged from 1998 to the present where the patients have been followed up during different intervals as it can be identified in the table 1 below:

**Table 1. Frequency of follow-up across included cohort studies**

<table>
<thead>
<tr>
<th>Study</th>
<th>Arm</th>
<th>Frequency of clinical observations</th>
<th>Frequency of laboratory investigations (not for diagnosis)</th>
<th>Frequency of laboratory Diagnosis (example 0,7,14)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Col: Diagnosis Cohort: Evaluation of the diagnostic accuracy and usefulness of rapid tests for early diagnosis of dengue</td>
<td>Outpatient/Inpatient</td>
<td>3-5, 4-6, 5-7, 15-17</td>
<td>3-5, 4-6, 5-7, 15-17</td>
<td>3-5, 4-6, 5-7, 15-17</td>
</tr>
<tr>
<td>Col: Prognosis: Immune Mechanisms of Pathogenesis in Patients with Dengue Infection</td>
<td>Outpatient/Inpatient</td>
<td>Once per day (Inpatient)</td>
<td>0-7 days, 3-6, and 24 weeks Note: Inpatients also provided samples every 48 hours during hospitalization (up to 4 samples)</td>
<td>0</td>
</tr>
<tr>
<td>Col: Piedecuesta’s household-based dynamic cohort. Identification of age groups to be prioritized for vaccination in a population of children and adolescents</td>
<td>Outpatient</td>
<td>Biweekly (All, telephone) Once per day (Incident Febrile cases)</td>
<td>Once per day (Incident Febrile Cases)</td>
<td>Annual (All) 0, 7 days (Incident Febrile cases)</td>
</tr>
<tr>
<td>Col:AEDES Cohorts Identification of prognostic markers of severity in dengue</td>
<td>Outpatient</td>
<td>Once per day</td>
<td>Once per day</td>
<td>0,14</td>
</tr>
<tr>
<td></td>
<td>Validation of a clinical definition for dengue and evaluation of its usefulness to identify early conditions associated with hospitalization</td>
<td>Outpatient</td>
<td>Once per day</td>
<td>Once per day</td>
</tr>
<tr>
<td></td>
<td>&quot;AEDES: Abordando Áreas Endémicas del Dengue para el Estudio de su Severidad&quot;. Colombian multicentric study.</td>
<td>Outpatient</td>
<td>Once per day</td>
<td>0, 6</td>
</tr>
<tr>
<td></td>
<td>Inpatient</td>
<td>Once per day</td>
<td>Once per day</td>
<td>0,10</td>
</tr>
<tr>
<td>CDY15: Efficacy and safety of a new tetravalent dengue vaccine in healthy children and adolescents aged 9 to 16 years in Latin America</td>
<td>Placebo</td>
<td>Enrollment, 6 y 12 (Months) (Incident Febrile Cases) Daily (Incident Febrile Cases)</td>
<td>0-5, 14 (Incident Febrile Cases)</td>
<td></td>
</tr>
<tr>
<td>Nica-Pediatric Dengue Hospital-based Study</td>
<td>Outpatient</td>
<td>Once per day</td>
<td>Once per day</td>
<td>0,1,2,14-28</td>
</tr>
<tr>
<td></td>
<td>Inpatient</td>
<td>More than once per day</td>
<td>Once per day</td>
<td>0,1,2,14-28</td>
</tr>
<tr>
<td>Nica-PDCS</td>
<td>Follow-up</td>
<td>2-3 times per week</td>
<td>2-3 times per week</td>
<td>0,14-21</td>
</tr>
<tr>
<td></td>
<td>Annual sample</td>
<td>Once per year.</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Evaluation of the diagnostic accuracy and usefulness of rapid tests for early diagnosis of dengue

(Diagnosis Cohort)

The diagnosis cohort, funded by E25bio, Inc., was aimed to determine the diagnostic usefulness of repeated NS1 rapid testing in clinical settings(7). This cohort enrolled and followed patients (≥2 years old) who had both clinically suspected dengue and a positive dengue rapid test (NS1 antigen) at the time of consultation or hospitalization. Participants were followed at 1, 2, and 10 days (convalescence) after recruitment to determine the incidence of dengue complications among confirmed cases. Dengue infection was defined as positive NS1 results (acute sample).

Immune Mechanisms of Pathogenesis in Patients with Dengue Infection (Prognosis Cohort)

The Prognosis Cohort began with the goal to prospectively validate the predictive accuracy of a pool of transcriptomics intended to predict severity among patients with confirmed DENV infection(8). This study was funded by the U.S. Department of Defense and Colombia’s Centro de Atención y Diagnóstico de Enfermedades Infecciosas (CDI, Colombia). This cohort enrolled patients (≥ 1 year old) with clinically suspected dengue and conducted follow-up at 1, 2, and 10 days (convalescence) after recruitment to determine the incidence of dengue complications among confirmed cases. Participants recruited in outpatient settings were clinically evaluated at enrolment.
and asked to provide additional blood samples between 3-6 days and 7-10 days if their first sample was obtained up to 2 days and 3-4 days after the onset of fever, respectively. Those who were recruited from inpatient settings underwent daily clinical evaluation, and blood samples were drawn every 48 hours during hospitalization (up to 4 samples). In all participants, additional samples were collected 3-6 and 24 weeks after onset of fever(9,10). Dengue infection was defined as positive PCR results (acute sample).

Piedecuesta’s Household-Based Dynamic Cohort (PHBDC)

Piedecuesta's household-based dynamic cohort sought to estimate age-specific dengue seroprevalence and identify age groups to be prioritized for vaccination among children and adolescents. This study was funded by the Colombian Science Ministry, Minciencias. PHBDC was a population-based cross-sectional study which began in 2014 and enrolled and evaluated healthy children and adults (15%) from Piedecuesta (a mid-size city with endemic DENV). The aim of the study was to estimate age-specific dengue seroprevalence in order to identify age groups to be prioritized for vaccination among children and adolescents. Based on the results of this seroprevalence study, a household-based dynamic cohort was initiated in 2015 with the aim of estimating the age-specific incidence of dengue in Piedecuesta (N=2,730). This cohort enrolled children (2-15 years old) and adults within the same household. The cohort followed up with participants on a biweekly basis through telephone contact to identify incident febrile cases. Cases were identified through clinical evaluation, and blood samples studied using Luminex ArboMIA to determine etiological diagnosis (annual [2016-2017] cumulative incidence: 6.0%). Additionally, this cohort conducted annual visits to the participants’ residences to collect blood samples to determine dengue seroconversion (losses to follow-up: 6.5% and 3.2% during the first and second
year, respectively), a strategy that allowed for the estimate of attack rates to be calculated for
chikungunya (22%) and ZIKV (34%) during the outbreaks of 2015 and 2016, respectively(11).

Abordando Áreas Endémicas del Dengue para el Estudio de su Severidad (AEDES) Cohorts

The AEDES cohorts’ data is an amalgamation of three independent studies, which were assembled
with the aim of developing and validating diagnostic and prognostic algorithms for DENV funded
by the same national agency above, Minciencias. The first two cohorts were initiated and
conducted during endemic periods (2003-2004, n=500; and 2006-2008, N=705) and a third one
during an epidemic (2009-2011, N=2,004). These studies shared similar enrolment protocols:
febrile patients with clinically suspected dengue were recruited at the point of care. Whereas the
first two cohorts were conducted in outpatient settings of Bucaramanga, the third (the AEDES
cohort—see Table 2 for clarification) was a multicenter study that enrolled and followed
individuals in both outpatient (Bucaramanga, Barranquilla, and Palmira) and inpatient
(Bucaramanga, Cali, Neiva, and Palmira) settings. Patients in the first two cohorts came in for
follow-up visits one to seven days via clinical and laboratory assessments (median follow-up: 4
and 3 days, respectively), with a convalescent blood sample taken approximately 2 weeks after
disease onset. The third study which enrolled participants from outpatient and inpatient settings
had median follow-up times of 3 and 2 days, respectively. Dengue infection was defined as an
ELISA IgM/IgG seroconversion or four-fold increase in titers in paired samples or virus isolation
(acute sample) in the first two cohorts, and as an ELISA IgM/IgG seroconversion or four-fold
increase in paired samples, or positive ELISA NS1 or PCR, or viral isolation (acute sample) in
AEDES(12).
Efficacy and Safety of a Novel Tetravalent Dengue Vaccine in Healthy Children and Adolescents
Aged 9 to 16 years in Latin America - placebo arm (CYD15)

Participants in the CYD15 were healthy children and adolescents (9-16 years old) who were recruited from a placebo arm of a randomized control trial funded by Sanofi Pasteur and had the primary goal of evaluating the efficacy of the chimeric yellow fever–dengue–tetravalent (CYF- DTV) dengue vaccine. The CYD15 study was a multicenter, placebo-controlled, randomized, observer-blind phase III DENV vaccine efficacy clinical trial, examining the efficacy of a vaccine to prevent symptomatic virologically-confirmed dengue cases in infants. Participants were randomly assigned, in a 2:1 ratio, to receive three doses of the Recombinant, live, attenuated, tetravalent dengue vaccine, CYD-TDV, (treatment group) or placebo (0.9% sodium chloride; control group) within 1 year. Participants included in this meta-cohort were healthy children and adolescents between 9-16 years old from the placebo arm living in Bucaramanga, Colombia. Volunteers were invited to participate through contacts with schools in the metropolitan area, referred by relatives of participants, or recruited by community leaders. Participants were followed through biweekly phone calls. In case of identifying any febrile episode, participants were asked to provide blood samples to perform virological confirmation and ELISA (NS1, IgM/IgG) testing for dengue infection as well as hemogram and hepatic function tests. Additional ELISA (IgM/IgG), hemogram and hepatic function tests were repeated in a convalescent sample collected up to 21 days after fever's onset(13).

Pediatric Dengue Cohort Study (PDCS)

The PDCS was established as a community-based cohort in District II of Managua, Nicaragua in 2004. The cohort was initially established to study DENV transmission and to characterize
It has since evolved to study the virologic and immunologic determinants of response to sequential DENV and ZIKV infections, epidemiological risk factors for infection and disease, and expanded to include other arboviruses, including CHIKV(14,15). The cohort has been funded by a variety of sources, including the US National Institutes of Health (NIH) National Institute of Allergy and Infectious Disease (NIAID), the Pediatric Dengue Vaccine Initiative of the International Vaccine Institute (IVI), the Bill and Melinda Gates Foundation, among others. PDCS is a community-based cohort study which enrolled 2-9-year-olds in August 2004. Participants were originally invited to remain in the study until their 12th birthday, but the restriction increased to 15 in 2007, and 18 in 2019(16). Each year, new two-year old children are enrolled, and additional replacement enrollment is performed as needed in the older age groups to maintain the cohort’s age structure. At any given time, there are roughly 3,800 - 4,100 children actively participating in the PDCS(17,18). With the introduction of chikungunya virus (CHIKV) and ZIKV into Latin America and specifically Nicaragua, CHIKV and ZIKV were added to the PDCS in 2014 and 2015, respectively. Visits are divided into 4 categories (A-D) based on symptomatology(19). Category A cases include fever plus symptoms and signs of suspected dengue (WHO’s case definition). Category B cases are undifferentiated febrile illnesses. Category C cases are fevers with a non-arboviral diagnosis (e.g., influenza, UTI), and category D cases are non-febrile cases. With the introduction of ZIKV, the D case category was divided into two subsets, D cases with ZIKV-like symptoms such as red eyes and rash and D cases without ZIKV-like symptoms(20).

Category A, B, and ZIKV-like D cases are tested for DENV, ZIKV, and CHIKV using RT-PCR and serological testing. Acute (1-4 days post-onset of symptoms) and convalescent (14-21 days)
samples are collected from all suspected cases. An additional sample is collected at day 4-6 from RT-PCR-confirmed DENV and ZIKV cases for immunological studies(21). Each year, a healthy blood sample (serum/plasma, and PBMCs prepared from a subset) is collected and used (a) to detect arbovirus infections that may not have been apparent but may have occurred throughout the year and (b) for additional immunological studies. Data on socioeconomic factors, demographics, and medical history are collected at enrollment and are updated annually. During clinical visits, detailed information on symptoms and symptom onset is collected(22).

**Pediatric Dengue Hospital-Based Study (PDHS)**

This cohort was founded in 1998 to investigate the clinical, immunological, and viral risk factors for severe DENV, assessing biomarkers, and studying immune responses over time. This study has been supported by the NIAID at NIH through various mechanisms. The Pediatric Hospital-Based Study began in 1998 and enrolls children ages 6 months to 14 years who present to the Hospital Infantil Manuel de Jesus Rivera (HIMJR) with suspected dengue (< 7 days from illness onset)(23). Both in-patient and out-patient suspected cases are eligible for enrollment. Upon enrollment, a complete physical exam is performed, and medical history is collected. Participants are followed throughout the acute phase of their illness and data including vital signs, symptoms, and treatment are recorded daily. Blood samples for complete blood counts, molecular, serological, and virological testing are collected daily for the first three days. An additional convalescent sample is collected 14-21 days after enrollment. A longitudinal arm of this study, for those participants who consent, collects samples 3-, 6-, 12-, and 18-months post-illness for immunological studies. The protocol was amended in 2014 to include CHIKV and again in 2016 to include ZIKV.
The primary objective of the IDAMS Cohort was to evaluate warning signs and predictors or biomarkers associated with progression to severe dengue in order to facilitate triage efforts. Funding was provided by the EC’s Seventh Framework Program. The IDAMS study (2011-2016) was a prospective multi-center acute febrile illness study conducted in Vietnam, Cambodia, Malaysia, Indonesia, Brazil, Venezuela, and El Salvador. Sites recruited participants who with a history of fever for $\leq$ 72-84 hours (site-dependent), presenting with clinical symptoms suggestive of dengue (in patients $>5$ years of age). Patients were excluded if a) they presented with severe dengue at enrolment, b) a clinician judged the patient was unlikely to attend daily outpatient follow-up visits, or c) the clinical presentation strongly suggested a diagnosis other than dengue (e.g., pneumonia, otitis, etc.). Only the data from Latin America were considered for the meta-cohort. However, as the data structure is homogenous, other study locations could be added in the future. The study design was described before in detail (24,25). In brief, patients with a history of fever for less than 72-84 hours (site dependent) and suggestive of DENV were recruited in outpatient clinics across the participating sites and followed up daily for a maximum of 6 days or until afebrile for 48h, with a final follow-up visit around day 10-14 of illness. Daily follow-up included physical examination as well as simple laboratory investigations such as full blood count. Dengue infection was defined as confirmed positive by either PCR or ELISA NS1 result in acute sample.
Cohort of Symptomatic Pregnant Women

In 2012, a prospective cohort for dengue surveillance in mother–infant pairs was established within the Manguinhos, Rio de Janeiro area. In 2015, however, most of these were later identified as ZIKV cases. To identify these ZIKV cases in the Rio de Janeiro population, the pregnancy cohort study was modified to enroll women who presented with a rash at any week of gestation. It was supported by the Department of Science and Technology (Departamento de Ciência e Tecnologia–DECIT) of the Brazilian Ministry of Health (Ministério da Saúde) and funded by the Coordination of the Improvement of Higher Level Personnel (Coordenação de Aperfeiçoamento de Pessoal de Nível Superior–CAPES); the Bill and Melinda Gates Foundation, Grand Challenges Explorations; and the National Institute of Allergy and Infectious Diseases (NIAID) of the National Institutes of Health. Brazil’s Symptomatic Women cohort offered enrollment between September 2015 to September 2016 to pregnant women who attended the acute febrile illness clinic at the Oswaldo Cruz Foundation and who presented with a rash that had developed in the previous 5 days, with or without an associated fever. Laboratory data was collected after enrollment(26). Weekly follow-ups occurred over the phone, and clinical and laboratory follow-up occurred within 30 days of enrollment and were referred for fetal ultrasonography follow-ups at three timepoints: before 20 weeks of gestation, between 20-30 weeks of gestation, and after 30 weeks of gestation. ZIKV infection was defined as positive RT-PCR, in acute samples of blood or urine.
<table>
<thead>
<tr>
<th>Country</th>
<th>Study</th>
<th>Years Study was Conducted</th>
<th>Included pediatric cohorts</th>
<th>Study Status</th>
<th>N</th>
<th>Mean age (yr)</th>
<th>Annual Loss to follow-up (%)</th>
<th>Mean age (yr)</th>
<th>Female (%)</th>
<th>Confirmed Dengue (%)</th>
<th>Mean age (yr)</th>
<th>Female (%)</th>
<th>Confirmed Dengue (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>COL</td>
<td>Diagnosis Cohort (Evaluation of the diagnostic accuracy and usefulness of rapid tests for early diagnosis of dengue)</td>
<td>2019-2021</td>
<td>x</td>
<td>Completed</td>
<td>28</td>
<td>20.5</td>
<td>10.7</td>
<td>21.1</td>
<td>60</td>
<td>40</td>
<td>8.0</td>
<td>33.3</td>
<td>33.3</td>
</tr>
<tr>
<td></td>
<td>Prognostic Cohort (Immune Mechanisms of Pathogenesis in Patients with Dengue Infection)</td>
<td>2019-2021</td>
<td>x</td>
<td>Completed</td>
<td>256</td>
<td>13.4</td>
<td>5.5</td>
<td>13.4</td>
<td>47.1</td>
<td>100</td>
<td>13.1</td>
<td>35.7</td>
<td>100</td>
</tr>
<tr>
<td></td>
<td>PHBDC (Piedecuesta’s household-based dynamic cohort: Identification of age groups to be prioritized for vaccination in a population of children and adolescents)</td>
<td>2015-2018</td>
<td>x</td>
<td>Completed</td>
<td>2730</td>
<td>12.6</td>
<td>10.1</td>
<td>14.6</td>
<td>51.2</td>
<td>6.0</td>
<td>12.3</td>
<td>48.5</td>
<td>NA</td>
</tr>
<tr>
<td></td>
<td>Identification of prognostic markers of severity in dengue</td>
<td>2003-2004</td>
<td>x</td>
<td>Completed</td>
<td>500</td>
<td>24.7</td>
<td>NA</td>
<td>24.7</td>
<td>51.2</td>
<td>39.0</td>
<td>NA</td>
<td>NA</td>
<td>NA</td>
</tr>
<tr>
<td></td>
<td>Validation of a clinical definition for dengue and evaluation of its usefulness to identify early conditions associated with hospitalization</td>
<td>2006-2008</td>
<td>x</td>
<td>Completed</td>
<td>705</td>
<td>26.5</td>
<td>NA</td>
<td>26.5</td>
<td>42.8</td>
<td>32.3</td>
<td>NA</td>
<td>NA</td>
<td>NA</td>
</tr>
<tr>
<td></td>
<td>AEDES (Abordando Áreas Endémicas del Dengue para el Estudio de su Severidad). Colombian multicentric study.</td>
<td>2009-2011</td>
<td>x</td>
<td>Completed</td>
<td>1008</td>
<td>21.7</td>
<td>NA</td>
<td>21.7</td>
<td>48.2</td>
<td>38.6</td>
<td>NA</td>
<td>NA</td>
<td>NA</td>
</tr>
<tr>
<td></td>
<td>Inpatient</td>
<td></td>
<td>Completed</td>
<td>996</td>
<td>14.8</td>
<td>NA</td>
<td>14.8</td>
<td>47.3</td>
<td>41.2</td>
<td>NA</td>
<td>NA</td>
<td>NA</td>
<td></td>
</tr>
<tr>
<td></td>
<td>CYD15 (Placebo Arm): Efficacy and safety of a new tetravalent dengue vaccine in healthy children and adolescents aged 9 to 16 years in Latin America</td>
<td>2011-2017</td>
<td>x</td>
<td>Completed</td>
<td>484</td>
<td>12.0</td>
<td>13.0</td>
<td>12.0</td>
<td>51.8</td>
<td>8.3</td>
<td>12.0</td>
<td>48.0</td>
<td>3.2</td>
</tr>
<tr>
<td>NICA</td>
<td>PDCS (Pediatric Dengue Cohort Study)</td>
<td>2004- ongoing</td>
<td>x</td>
<td>Ongoing</td>
<td>9,699</td>
<td>6</td>
<td>4.3*</td>
<td>7.9</td>
<td>49.00</td>
<td>11.5</td>
<td>NA</td>
<td>NA</td>
<td>NA</td>
</tr>
<tr>
<td></td>
<td>PDHIS (Prospective Hospital-Based Study of Dengue Classification, Case Management, and Diagnosis)</td>
<td>1998- ongoing</td>
<td>x</td>
<td>Ongoing</td>
<td>2,659</td>
<td>13</td>
<td>7.03</td>
<td>8.59</td>
<td>48.10</td>
<td>53.7</td>
<td>NA</td>
<td>NA</td>
<td>NA</td>
</tr>
<tr>
<td>IDAMS</td>
<td>IDAMS (International Research Consortium on Dengue Risk Assessment, Management and Surveillance)****</td>
<td>2012-2016</td>
<td>Completed</td>
<td>1625</td>
<td>19.3</td>
<td>2.7*</td>
<td>19.0</td>
<td>47.6</td>
<td>28.8</td>
<td>13.4</td>
<td>86.4</td>
<td>NA</td>
<td></td>
</tr>
<tr>
<td>Latin America</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>BRA</td>
<td>Cohort of Symptomatic Pregnant Women</td>
<td>2015-2016</td>
<td>Completed</td>
<td>383</td>
<td>6.27</td>
<td>13.6</td>
<td>29.77</td>
<td>86.40</td>
<td>0</td>
<td>27.56</td>
<td>13.60</td>
<td>0</td>
<td></td>
</tr>
</tbody>
</table>

* Excluded (not lost) because: (i) > 84 hours from disease onset (N=39); (ii) disease onset unknown (N=5); (iii) laboratory diagnosis not available (N=53). 7 patients fulfilled more than one criterion. NA: Not available.
**** Venezuela and Brazil sub-cohorts (Total N enrolled in all sites including Asia = 7428)

Average annual rate in longstanding cohort studies
Table 3. Summary of ethics approvals and consent among patients

<table>
<thead>
<tr>
<th>Study</th>
<th>Did you obtain ethics approval for this study?</th>
<th>Start date</th>
<th>End date</th>
<th>Type of informed consent obtained</th>
<th>Does the study include minors?</th>
<th>Was informed parental consent obtained for minors?</th>
<th>Was informed consent obtained for adults?</th>
</tr>
</thead>
<tbody>
<tr>
<td>Col: Diagnosis Cohort: Evaluation of the diagnostic accuracy and usefulness of rapid tests for early diagnosis of dengue</td>
<td>Yes</td>
<td>January 23, 2020</td>
<td>June 01, 2020</td>
<td>Written</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td>Col: Prognosis: Immune Mechanisms of Pathogenesis in Patients with Dengue Infection</td>
<td>Yes</td>
<td>March 18, 2019</td>
<td>February 20, 2020</td>
<td>Written</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td>Col: Piedecuesta’s household-based dynamic cohort. Identification of age groups to be prioritized for vaccination in a population of children and adolescents.</td>
<td>Yes</td>
<td>June 1, 2015</td>
<td>December 30, 2018</td>
<td>Written</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td>Identification of prognostic markers of severity in dengue</td>
<td>Yes. A copy of IRB approval obtained in 2003 was requested.</td>
<td>April 01, 2003</td>
<td>March 30, 2005</td>
<td>Written</td>
<td>Yes</td>
<td>Yes</td>
<td>No</td>
</tr>
<tr>
<td>Validation of a clinical definition for dengue and evaluation of its usefulness to identify early conditions associated with hospitalization</td>
<td>Yes. A copy of IRB approval obtained in 2005 was requested.</td>
<td>February 01, 2006</td>
<td>April 30, 2008</td>
<td>Written</td>
<td>Yes</td>
<td>Yes</td>
<td>NA</td>
</tr>
<tr>
<td>&quot;AEDES: Abordando Áreas Endémicas del Dengue para el Estudio de su Severidad&quot;. Colombian multicentric study.</td>
<td>Yes</td>
<td>May 01, 2009</td>
<td>May 30, 2011</td>
<td>Written</td>
<td>Yes</td>
<td>Yes</td>
<td>NA</td>
</tr>
<tr>
<td>CDY15: Efficacy and safety of a new tetravalent dengue vaccine in healthy children and adolescents aged 9 to 16 years in Latin America</td>
<td>yes</td>
<td>June 15, 2011</td>
<td>Feb 28, 2018</td>
<td>Written</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td>Study Description</td>
<td>Consent Required</td>
<td>Start Date</td>
<td>End Date</td>
<td>Consent Type</td>
<td>Consent Obtained</td>
<td>Consent Signed</td>
<td>Consent Withdrawn</td>
</tr>
<tr>
<td>---------------------------------------------------------------------------------</td>
<td>------------------</td>
<td>------------</td>
<td>------------</td>
<td>----------------------------------</td>
<td>------------------</td>
<td>----------------</td>
<td>------------------</td>
</tr>
<tr>
<td>Pediatric Dengue Cohort Study (PDCS) (Nicaragua)</td>
<td>Yes</td>
<td>August 31, 2004</td>
<td>ongoing</td>
<td>Informed consent, parental; verbal assent</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td>Pediatric Dengue Hospital-Based Study (PDHS) (Nicaragua)</td>
<td>Yes</td>
<td>August 4, 2005</td>
<td>ongoing</td>
<td>Informed consent; verbal assent; written assent</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td>IDAMS (Carabobo, Venezuela)</td>
<td>Yes</td>
<td>September 27, 2013</td>
<td>November 14, 2016</td>
<td>Written</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td>IDAMS (Rio de Janeiro, Brazil)</td>
<td>Yes</td>
<td>April 6th, 2015</td>
<td>May 9th, 2016</td>
<td>Written</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td>Cohort of Pregnant Women (Brazil)</td>
<td>Yes</td>
<td>September 1&lt;sup&gt;st&lt;/sup&gt;, 2015</td>
<td>May 30&lt;sup&gt;th&lt;/sup&gt;, 2016</td>
<td>Written</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
</tr>
</tbody>
</table>
Retrospective and Prospective Harmonization
Efforts of the Meta-Cohort

Individual patient data (IPD) meta-analyses (MA) are considered the gold standard for meta-analyses. The strength of conducting IPD-MAs, as opposed to standard meta-analyses using effect estimates, is the opportunity to control for baseline heterogeneity.

Knowing that we had a collective wealth of arbovirus data, we initially set out to find a common data dictionary for all cohorts within the ReCoDID consortium. This effort included trying to find common variables across all Zika cohorts, then across dengue cohorts, as well as any chikungunya patients. As mentioned in the literature, there has not been a gold standard method for data harmonization, but we eventually found the Maelstrom Guide to Rigorous Retrospective Harmonization. The first step of the guide is to define a research objective for harmonization, which allowed us to focus on the dengue studies based on a research question that aimed to identify future flavivirus clinical epidemiology in settings where we can infer the past history of infections.

Once we had a focused research question, the next step was implementing a well-defined structure of dimensions (endpoints/confounders/exposure) and domains to tackle from a medical perspective. These dimensions guided us in establishing a set of medical meetings where we identified, prioritized, and defined in a semantic manner the variables to include in the master data dictionary as indicated in the methodology used and explained in the Fig 1.

Fig 1. Harmonization process.

Then we compared each study data dictionary against the reconciled master data dictionary establishing a harmonization potential based on a predefined structure of conventions.
that gave us a complete set of variables with different levels of potential use for harmonization, this can be identified in the Fig 2.

Fig 2. Harmonization potential

Retrospective harmonization is generally resource-intensive (29,30), as we can attest to. In order to facilitate future IPD-MAs, we are in the process of creating a Case Report Form which we then standardized according to the CDISC data standard. The hope is that future cohorts conducting arbovirus research will be able to prospectively implement this CRF, decreasing, or removing entirely, the effort and cost of retrospectively harmonizing data to be pooled in an IPD-MA.

What has it found? Key findings and publications

Key findings include: 1) low serum 25(OH)D concentrations in patients predicted the progression of dengue fever(31); 2) the highest risk group for severe dengue being patients with preexisting anti-DENV antibody titers; the same study showing a preventative effect in those patients who have a (very) high level of antibody titers(32); and 3) the development, validation, and evaluation of the usefulness of a scale for the prediction of disease severity among confirmed cases of dengue (32%, 39%, and 41% for the three cohorts).
What are the main strengths and weaknesses of the combined meta-cohort effort?

A technical strength of the Consortium’s efforts is the retrospective harmonization of participant-level data which has been completed according to the Maelstrom group’s recommendations for best practice (28), which outlines steps 0 (define the research question) to 5 (disseminate and preserve final harmonization products). Maelstrom Research guidelines’ Step 0 is why the ReCoDID Consortium’s data outlined in this profile is primarily focused on dengue-related outcomes despite collecting data related to viruses. Another strength is the use of CDISC/SDTM standards as reference for the definition of the core variables harmonized (33), which enables interoperability between the meta-cohort and other studies that apply CDISC SDTM. The harmonization (processing data collected by studies und a common variable format), or standardization (use of a data standard (CDISC, e.g.) to define the core variables format to be generated, required to conduct individual-level meta-analyses can be extremely resource-demanding. Going forward, and in order to minimize this burden and improve individual-level patient data meta-analyses (IPD-Mas), the ReCoDID Consortium recommends the creation of a standard case report form (CRF) for acute viral syndrome (arbovirus) research, which includes the features of the overlapping clinical syndromes associated with the most important arboviruses (e.g., DENV, ZIKV, CHIKV, etc.), to be used by all partner studies in the future. Potentially, this CRF will include modules for different severe disease manifestations that can be adapted to the local situation (i.e., bleeding module, neurological module, liver pathology module).

Another strength of the combined cohort is that it covers different countries and partner sites across Latin America - each of which experienced slightly different histories of DENV serotypes,
CHIKV, and ZIKV infections over the last decades. The resulting ‘experiment of nature’ represents a population immune landscape that we now would like to prospectively follow with future cohorts, considering the potential of immunological interaction between related flaviviruses (e.g., DENV1-4 and ZIKV). In Fig 3 and Fig 4, created using R (4.0.5)(34), we present DENV, ZIKV, and CHIKV activity over time in the respective countries and partner sites. This combined cohort is a first step towards the direction of a multicentric arbovirus cohort, which needs to be complemented with advanced technology assessing immunological history, additional investments in future harmonization, and standardization.

**Fig 3.** DENV serotype distribution over time. **Fig 4.** Reported number of cases for DENV, ZIKV and CHIKV over time

| a) Pernambuco (Brazil) * | a) Pernambuco (Brazil) II, ∑,+ |
| b) Ceará (Brazil) Δ | b) Ceará (Brazil) II, + |
| c) Rio de Janeiro (Brazil) Δ | c) Rio de Janeiro (Brazil) II, |
| d) Nicaragua § | d) Nicaragua § |
| e) Colombia ∞ | e) Colombia ∞ |
| f) Venezuela # | f) Venezuela ** |
| g) El Salvador |

* Data of annual DENV serotypes from PE, provided by: the Central Laboratory of Pernambuco (Laboratório Central de Pernambuco-LACEN PE)
Δ Data of annual DENV serotypes from CE and RJ, retrieved from the national online database of the Brazilian Ministry of Health (https://datasus.saude.gov.br/informacoes-de-saude-tabnet/, accessed on 11/8/21 to 11/11/21)
∞ This data was derived from the Colombian National Institute of Health. http://portalsvigila.ins.gov.co/Paginas/Vigilancia-Rutinaria.aspx)
II Data of annual DENV, ZIKV and CHIKV, reported cases in the states of Ceará (CE), Pernambuco (PE) and Rio de Janeiro (RJ), retrieved from: the national online database of the Brazilian Ministry of Health (https://datasus.saude.gov.br/informacoes-de-saude-tabnet/, accessed on 11/8/21 to 11/11/21), except for CHIKV in 2015-2016 in CE and PE (see below). Data of CHIKV from CE and PE (2015-2016), taken from: Epidemiological Bulletins (EBs) published by the Brazilian Ministry of Health.
∑ Data for ZIKV from PE (2015), may be subject to reporting bias. At the time, ZIKV was largely unknown and therefore ZIKV cases might have been classified/notified as dengue cases, de Brito et al. (2016), Data from: Rev Soc Bras Med Trop 49(5):553-558, September-October, 2016 doi:10.1590/0037-8682-0245-2016.
+ Data of CHIKV from CE and PE (2015-2016), Taken from: Epidemiological Bulletins (EBs) published by the Brazilian Ministry of Health.
§ PDVS arboviral case data (Oct 2004 - Mar 2021). Image one contains yearly level data of PCR confirmed Dengue cases, and image two contains DENV, CHIKV and ZIKV confirmed symptomatic infections on a monthly basis.
Can I access the data? Where can I find out more?

Interested parties will be able to access data dictionaries that include information on variables across the datasets via BioStudies\(^{(35)}\). After consultation with each cohort, linked harmonized and curated human cohort data (CE and HDL) is planned to be made accessible through the European Genome-phenome Archive (EGA)\(^{(36)}\) platform to Data Users after their requests are evaluated by the ReCoDID Data Access Committee (DAC). Researchers will be able to openly access the descriptive cohort metadata related to the meta-cohort via the EMBL-EBI Cohort Browser\(^{(37)}\).

Cohorts sharing CE and HDL data within the ReCoDID Platform are responsible for obtaining regulatory and ethical approvals at the local, or, in the case of Brazil, national, level for data sharing. Where cohorts’ original informed consent forms did not include broad consent for future use of data, ReCoDID worked with cohorts to apply for a waiver of consent. The waivers of consent submitted to the Commission of Ethics in Research, CONEP, Brazil’s national ethics regulatory authority. The waivers of consent have been in process for one year and we expect we will need two years to complete the CONEP review. Samples will not be shared; however, the associated viral sequencing data can be shared and accessed openly at the European Nucleotide Archive (ENA)\(^{(38)}\). Any further associated and shared data types can be shared and retrieved from the relevant EMBL-EBI resources\(^{(39)}\).

The assessment of political, ethical, administrative, regulatory, and legal (PEARL) issues related to data sharing guided the implementation and activities of ReCoDID to promote ethical data
governance and sharing that carefully considers the perspective and context of Low and Middle Income Countries (LMICs). Empirical research, bibliography reviews, and stakeholders’ consultations guided ReCoDID members’ activities and informed the development of the ReCoDID Data Governance Framework (DGF). The ReCoDID DGF is a high-level normative, organizational, and technical document that describes the goals and principles by which the ReCoDID functions, among them the FAIR (findable, accessible, interoperable, and reusable) data principles. The ReCoDID DGF also implements international standards and best practices for data sharing to promote the public interest and advancement of science. It further describes how the ReCoDID Platform functions and complies with data protection and privacy legislation, mainly the European Union (EU) General Data Protection Regulation (GDPR), and how it relates to other countries’ legislations considering international data transfers. The ReCoDID DGF is centred on protecting the rights and interests of Data Subjects (patients and research participants) and different stakeholders that participate in the biomedical innovation ecosystem, including researchers and information technology (IT) professionals that engage in biomedical research and develop the infrastructure and tools that enable health data platforms.

ReCoDID acknowledges the challenges to promoting a biomedical innovation system that is transparent, equitable, and participatory, that incorporates LMICs’ context-specific perspectives. This is deeply rooted in identifying and overcoming the PEARL barriers related to data sharing mentioned above, as well as the enablers and different strategies to guarantee and implement ethical data sharing and governance. Addressing these PEARL barriers requires creating discussion forums, building research networks, and promoting best practices among academic communities that call for the de-colonization of global health practices and challenge power structures that support and perpetuate them. Therefore, the ReCoDID DGF and ReCoDID
Intellectual Property and Open Science Policy incorporate these elements and include issues related to benefit sharing, authorship, attribution, and recognition, as well as the mechanism to implement them.

For more information, please contact ena-path-collabs@ebi.ac.uk, with reference to the ReCoDID Project. To share associated human cohort data, please contact the University of Heidelberg (thomas.jaenisch@uni-heidelberg.de).

Funding

ReCoDID is supported by the European Union’s Horizon 2020 Research and Innovation Programme under Grant Agreement No. 825746 and the Canadian Institutes of Health Research, Institute of Genetics (CIHR-IG) under Grant Agreement N.01886-000. The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

Conflict of Interest

None declared.
Acknowledgements

This work was supported by the ReCoDID Arbovirus harmonization study group:

**Janet Achieng** (Heidelberg Institute of Global Health (HIGH), Heidelberg University Hospital, Heidelberg, Germany)

**Sonia Arguello** (Oswaldo Cruz Foundation (Fiocruz), Rio de Janeiro, Brasil)

**Gabriela Maria Marón Alfaro** (Hospital Nacional de Niños Benjamín Bloom, San Salvador, El Salvador)

**Till Bärnighausen** (Heidelberg Institute of Global Health (HIGH), Heidelberg University Hospital, Heidelberg, Germany)

**Bruno Souza Benevides** (University of the State of Ceara, Fortaleza, Brazil)

**Sarah Bethencourt** (Departamento de Estudios Clínicos, Facultad de Ciencias de la Salud, Universidad de Carabobo, Valencia, Venezuela)

**Andrea Caprara** (Departamento de Estudios Clínicos, Facultad de Ciencias de la Salud, Universidad de Carabobo, Valencia, Venezuela)

**Priscilla M.D.G. César** (McMaster University, Institute on Ethics & Policy for Innovation (IEPI))

**Guy Cochrane** (European Molecular Biology Laboratory, European Bioinformatics Institute, Wellcome Genome Campus, Hinxton, Cambridge CB10 1SD, UK)

**Monika Consuegra** (Centro de Atención y Diagnóstico de Enfermedades Infecciosas, Bucaramanga, Colombia)
Fabio Otero (Centro de Atención y Diagnóstico de Enfermedades Infecciosas, Bucaramanga, Colombia)

Kerstin Rosenberger (Heidelberg Institute of Global Health (HIGH), Heidelberg University Hospital, Heidelberg, Germany)

Anna M. Gajewski (Sustainable Sciences Institute, Managua, Nicaragua)

Nadim Rahman (European Molecular Biology Laboratory, European Bioinformatics Institute, Wellcome Genome Campus, Hinxton, Cambridge CB10 1SD, UK)

Bladimir Cruz (Hospital Nacional de Niños Benjamín Bloom, San Salvador, El Salvador)

Guillermo Barahona Escobar (Hospital Nacional de Niños Benjamín Bloom, San Salvador, El Salvador)

Maria I. Estupiñán (Centro de Atención y Diagnóstico de Enfermedades Infecciosas, Bucaramanga, Colombia)

Isabel Fortier (Research Institute, McGill University Health Centre, Montreal, Canada)

Rosa Margarita Gélvez Ramírez (Centro de Atención y Diagnóstico de Enfermedades Infecciosas, Bucaramanga, Colombia)

María Fernanda Vincenti Gonzalez (University of Groningen, University Medical Center Groningen, Department of Medical Microbiology and Infection Prevention, Groningen, The Netherlands)

Peter W. Harrison (European Molecular Biology Laboratory, European Bioinformatics Institute, Wellcome Genome Campus, Hinxton, Cambridge CB10 1SD, UK)

Suran Jayathilaka (European Molecular Biology Laboratory, European Bioinformatics Institute, Wellcome Genome Campus, Hinxton, Cambridge CB10 1SD, UK)

Guillermina Kuan (Sustainable Sciences Institute, Managua, Nicaragua)
Manish Kumar (European Molecular Biology Laboratory, European Bioinformatics Institute, Wellcome Genome Campus, Hinxton, Cambridge CB10 1SD, UK)

Luz Marina Leegstra (Heidelberg Institute of Global Health (HIGH), Heidelberg University Hospital, Heidelberg, Germany)

Erley Ferlipe Lizarazo (University of Groningen, University Medical Center Groningen, Department of Medical Microbiology and Infection Prevention, Groningen, The Netherlands)

Luigi Marongiu (Heidelberg Institute of Global Health (HIGH), Heidelberg University Hospital, Heidelberg, Germany)

Ágnes Molnár (Heidelberg Institute of Global Health (HIGH), Heidelberg University Hospital, Heidelberg, Germany)

Cesar Narvaez (Sustainable Sciences Institute, Managua, Nicaragua)

Federico Narvaez (Sustainable Sciences Institute, Managua, Nicaragua)

Sergio Ojeda (Sustainable Sciences Institute, Managua, Nicaragua)

Gabriele Rinck (European Molecular Biology Laboratory, European Bioinformatics Institute, Wellcome Genome Campus, Hinxton, Cambridge CB10 1SD, UK)

Ernesto Pleités Sandoval (Hospital Nacional de Niños Benjamín Bloom, San Salvador, El Salvador)

Frank Tobian (Heidelberg Institute of Global Health (HIGH), Heidelberg University Hospital, Heidelberg, Germany)

José Victor Zambrana (Sustainable Sciences Institute, Managua, Nicaragua)

16. Luana Damasceno (Oswaldo Cruz Foundation (Fiocruz), Rio de Janeiro, Brasil)
References


6. ReCoDID | Reconciliation of Cohort data in Infectious Diseases [Internet]. [cited 2022 Nov 7]. Available from: https://recodid.eu/


33. Standards | CDISC [Internet]. [cited 2022 Nov 8]. Available from: https://www.cdisc.org/standards


36. EGA European Genome-Phenome Archive [Internet]. [cited 2022 Dec 1]. Available from:
37. Pathogens [Internet]. [cited 2022 Dec 1]. Available from:
https://www.ebi.ac.uk/ena/pathogens/v2/cohorts

38. ENA Browser [Internet]. [cited 2022 Dec 1]. Available from:
https://www.ebi.ac.uk/ena/browser/home

39. An introduction to EMBL-EBI resources | EMBL-EBI Training [Internet]. [cited 2022 Dec 1].
Available from: https://www.ebi.ac.uk/training/events/introduction-embl-ebi-resources/
<table>
<thead>
<tr>
<th>Domain</th>
<th>Temp Variable Names</th>
<th>Definition</th>
<th>Name</th>
<th>Unit</th>
<th>Type</th>
<th>Studies</th>
<th>Variables</th>
<th>Has Potential Status per Variable</th>
<th>Has Potential Status detail per Variable</th>
<th>Comments on Harmonization Rule</th>
</tr>
</thead>
<tbody>
<tr>
<td>Endpoints</td>
<td>age</td>
<td>age of the patient at the day of measurement * reported by patient or family or official documents</td>
<td>age</td>
<td>years</td>
<td>Integer</td>
<td>Nica-Cohorte Arbovirus</td>
<td>exp</td>
<td>1</td>
<td>1.2</td>
<td>Transformation: Substract(fecha_consulta - exp)</td>
</tr>
<tr>
<td>Endpoints</td>
<td>age</td>
<td>age of the patient at the day of measurement * reported by patient or family or official documents</td>
<td>age</td>
<td>years</td>
<td>Integer</td>
<td>Nica-Estudio Hospitalario</td>
<td>Edad</td>
<td>1</td>
<td>1.2</td>
<td>Transformation: DD-MM-YYYY -&gt; YYYY-MM-DD</td>
</tr>
<tr>
<td>Endpoints</td>
<td>age</td>
<td>age of the patient at the day of measurement * reported by patient or family or official documents</td>
<td>age</td>
<td>years</td>
<td>Integer</td>
<td>IDAMS</td>
<td>dm_age</td>
<td>1</td>
<td>1.1</td>
<td>idéntico</td>
</tr>
<tr>
<td>Endpoints</td>
<td>age</td>
<td>age of the patient at the day of measurement * reported by patient or family or official documents</td>
<td>age</td>
<td>years</td>
<td>Integer</td>
<td>Col-AedesCohort</td>
<td>Edad</td>
<td>1</td>
<td>1.1</td>
<td>idéntico</td>
</tr>
<tr>
<td>Endpoints</td>
<td>age</td>
<td>age of the patient at the day of measurement * reported by patient or family or official documents</td>
<td>age</td>
<td>years</td>
<td>Integer</td>
<td>Col-Diagnosis Cohort</td>
<td>Edad</td>
<td>1</td>
<td>1.1</td>
<td>idéntico</td>
</tr>
<tr>
<td>Endpoints</td>
<td>age</td>
<td>age of the patient at the day of measurement * reported by patient or family or official documents</td>
<td>age</td>
<td>years</td>
<td>Integer</td>
<td>Col-Prognosis Cohort</td>
<td>Edad</td>
<td>1</td>
<td>1.1</td>
<td>idéntico</td>
</tr>
<tr>
<td>Endpoints</td>
<td>age</td>
<td>age of the patient at the day of measurement * reported by patient or family or official documents</td>
<td>age</td>
<td>years</td>
<td>Integer</td>
<td>Col-Seroprevalence Cohort</td>
<td>x</td>
<td>2</td>
<td>2.1</td>
<td>indisponible</td>
</tr>
<tr>
<td>Endpoints</td>
<td>age</td>
<td>age of the patient at the day of measurement * reported by patient or family or official documents</td>
<td>age</td>
<td>years</td>
<td>Integer</td>
<td>Brazil_cohort of symptomatic pregnant women</td>
<td>x</td>
<td>2</td>
<td>2.1</td>
<td>indisponible</td>
</tr>
</tbody>
</table>

Figure 2. Harmonization Potential
Figure 3. DENV Serootype Distribution
Figure 4. Reported DENV, ZIKV, CHKV Cases
Figure 1. Harmonization Process