COVID Integrated Data Repository (IDR) Navigation Guide

Contents

- 1) <u>Log In</u>
- 2) <u>Homepage</u>
- 3) <u>View Options for Data Searches</u>
- 4) Identify a Cohort
- 5) Get the Data
- 6) <u>Protocol Information and Data Dictionary</u>
- 7) <u>Table 1</u>
- 8) <u>User Permissions</u>

Log In

Any Duke faculty or staff member may submit a request via email to <u>CovidIDRandSpecimensRequest@duke.edu</u> to obtain access to the COVID-19 IDR dashboard. Once the request is processed, follow the steps below.

1	Navigate to https://covididr.duhs.duke.edu	▲ https://conditid ubinduktesition X + ✓ ✓ – □ × € → ♂ ○ ○ ○ conditid ubinduktesition LP ★ ○ ▶ □ ○ □ LP ★ ○ ▶ □ ○ □ € COVID-IDR Covid-IDR Lpgin
2	Click the Login button	Please log in to continue!
3	Click Sign in with Azure Active Directory	Sign in with Azure Active Directory
4	Select or type in your Duke NetID email address	Microsoft Pick an account
5	Log in using your Duke NetID credentials (including MFA if needed)	<image/>

Homepage

Navigate to desired areas via (1) left sidebar or (2) tabs in center of screen

Navigate with	n left sidebar or tabs in center of screen
COVIL DE	COVID-IDR The COVID integrated Reportiony (COVID-IDR) is a tool that hepp Dute researches search across the data and samples collected from multiple COVID-19 Buildes. A allowing you to see what is available from participants that fit you' research interests. The CR was developed to support the harmonization, sharing, and quality of
Pericipanta Per Study	To use COVID-IDR: Sign In using your NetiD Use the fitters to define your cohort of Interest Information about alvable data fitting your query will be displayed After finding data that fits your specifications, you can: Servi your canyon return to 1 state Download the destertified data for your leantified cohort Results the leadertified data for your leantified cohort Contact the R of a particular study to incuire about available biosecomes Protocol Kap: Protocol Kap: Pr
	Biogenimers - Ad Stader Biogenimers - Readuats (Hx0000010) Petrosen and Data Ductorary Petroperts Her Study Participants by Protocol

View Options for Data Searches

1	Users may select 1 of 3	Three views to search for data		
	views to search for data	Biospecimens - All Studies Biospecimens - Residuals (Pro00105316) Protocols and Data Dictionary Participants Per Study		
2	Select	COVID Integrated Data Repository Biospecimen Inventory		
	"Biospecimens	Filters Biospecimens		
	– All Studies"	Patients: Specimens: 10600 127461		
	to identify			
	patients or	Gender: Nasal Vash Nasal Swab Nasal Swab Nasal Swab Nasal Nasa		
	specimens	Race: Not Available Not Available Not Available PBMCs PBMCs Plasma Plasma 2 Spins RBCs RBCs RBCs RA, Total RBCs Saliva Saliva Stool RBCs Swab Nasal		
	across all	Ethnicity: Swab, Nasopharyngeal Tissue, FFPE Tissue, Fresh-Frozen Tissue, OCT Urine Turine Urine Wet Dry Web Blood		
	studies	Minimum Age:		
	included in the			
	COVID-19 IDR.	Naximum Age:		
	The survey	Tissue Type :		
	The graph	Condition Group:		
	displays	Condition:		
	specifien			
	types.			
L				

2	Select	COVID-19 Residual Data Repository Biospecimen Inventory		
	"Biospecimens-	Filters	Biospecimens	
	Residuals		Patients:	Specimens:
	(Pro00105316)	COVID Test Result:	4507	36474
	" to limit your	Gender:	BAL Blood Blood Lavage, Bronchoalvea Nasal Swab Nasal Swab Nasal Wash Nasal Wash Nasanhary Nasanhary	Lavage, Nasopharynge Nasal ngeal Swab Senum Nasopharyngeal Wash
	search to the	Race:	Swab, Nasopharyngeal	Whole Blood
	SOM COVID-19	Ethnicity:		
	Residual	Minimum Anet		
	Sample	remander rege-		
	Biorepository.	Maximum Age:		
		Tissue Type:		
	Note: The	Condition Groups:		
	"Biospecimens-	Condition:		
	Residuals	Minimum Collected Date: Select date		
	(Pro00105316)	Maximum Collected Date: Select date 🖹		
	" is identical to	Zip Code:		
	the other two			
	views, except it	Gendank Seq:		
	does not			
	include the			
	"Research			
	Protocol" filter.			
3	Select	Participants by Biospecimen Type and Protocol		
	"Participants	Filters	Biospecimens	
	Per Study" to		Patients:	Specimens:
	identify	COVID Test Result:	10600	127461
	patients or	Gender:	POLAGE_PANGOLIN_PRJ Pro0003	5974 Pro00100241 Pro00105249
	specimens			
	across all	Race!		
	studies	Ethnicity:		
	included in the	Minimum Age:		
	COVID-19 IDR.	Maximum Age:		
	The graph	Tissue Type:		
	displays counts	Condition Groups:		
	of specimens	Condition:		
	per protocols.			
		1		

Identify a Cohort

	1			
1 Select "Biospecimens - All Studies" to		Questions?View the full FAC	х Э.	
patients or specimens across all		Biospecimens - All Studies	Biospecimens - Residuals (Pro001	105316)
studies included in the COVID-19 IDR.		Participants by Protocol		
Select "Biospecimens- Residuals (Pro00105316)" to limit your search to the SOM COVID-19 Residual Sample Biorepository.				
2 Users will see this view to	COVID Integrated Data Repository Biospecimen Invento	yry	Biospecimens	
begin their search in "Biospecimens – All Studies". Note: The "Biospecimens- Residuals (Pro00105316)" is identical, except it does not include the "Research Protocol" filter.	COVID Test Result:		Patients: 10600 Masopharyngeal Bronchulw WashHaf None Net Available Sativa Sativa Sativa Boos RNA, Total Sativa Sativa Sativa Sativa Boos RNA, Total Sativa	Specimens: 127461

3	Use the filters to narrow your search to patients and/or specimens of interest.	COVID Integrated Data Repository Biospecimen Inventory Filters COVID Test Result: Gender:			
4a	Filter – COVID Test Result.	 COVID test results are pulled from the electronic health record (EHR). Individuals may have multiple test results over a period of time. Users may filter for Any positive test result Any negative test result No positive test result (i.e. only results in Duke data are negative) No negative test result (i.e. only results in Duke data are positive) There is no direct connection between a specimen and a positive /pegative test result 			
4b	Filter - Tissue Type	Users may specify tissue type for specimens of interest. "None" reflects that the field was blank (data issue).			
4c	Filters – Condition and Condition Group	Users may type conditions of interest and potential matches will populate for selection.			
4d	Filter – Research Protocols	Users may select a protocol to search for specimens only associated with protocols of interest.			
5	Data Preview	As filters are applied, the graph updates automatically. Depending on your level of access, some minimal de-identified data may be available for preview.			
		Patient ID Specimen ID Age C Race Sex Tissue Type			
		cd208f0a74fd0e2dd09ab80edc0 001d982b93c3d2390e2c2a913ac 64 BLACK OR AFRICAN AMERICAN male Whole Blood			
1		b298c1b95d58c2d37e7756ef56d 001f00ee27e75b434559797c0/a 62 CAUCASIAN/WHITE female Swab, Nasopharyngeal			
1		074e3a2dc51afcb05532372696a 0021c972a4c3Hed14x90b92546 33 CAUCASIAN/WHITE male Plasma			
		c22831bc62/e67d7da1a2b0980eb 00257c84715H268d033d678cefc 57 CAUCASIAN/WHITE male Plasma			
6	Save Search – You may save a search to execute again in the future.	Save Search Search name Save Save Save Save Save Save Save Sav			

Get the Data – For Users with Approved IRB Protocol

1	Download Data – You may download deidentified data from your search in .csv format	Search Download Data Patient Data Specimen Data Condition Data Observation Data		
2	Patient Data	 The following fields are available: "id","birth_date","gender","zip","ethnicity","race" "Birth date" is not the exact date, it has been "jittered" 		
3	Specimen Data	 The following fields are available: "id", "subject", "type", "protocol", "collected_date" "id" is de-identified specimen ID. Each one will be unique. "subject" is de-identified version of medical record number (MRN). You may see multiple results with this ID if the patient has more than one specimen 		
4	Condition Data	 The following fields are available: "id", "subject", "icd10", "icd9", "condition" "id" is condition ID. Each one will be unique. "subject" is de-identified version of medical record number (MRN). You may see multiple results with this ID if the patient has more than one condition Examples of conditions are 		
5	Observation Data	 The following fields are available: "id", "subject", "type", "value" "id" is observation ID. Each one will be unique. "subject" is de-identified version of medical record number (MRN). You may see multiple results with this ID if the patient has more than one condition Examples of observation types are BMI, smoking status, etc. Lab values should also be included here. 		
6	All Data	Users with IRB approval may send an email to <u>mailto:CovidIDRandSpecimensRequest@dm.duke.edu</u> to request more de-identified data and/or in a different format (e.g. FHIR) or identified data via an Honest Broker and delivered to PACE.		

Protocol Information and Data Dictionary

Provides a list of protocols included in the IDR with a brief description of each. Users may download a PDF version of the information.

1 Select "Protocols & Questions?View the full FAQ. Dictionary".							
		Biospecimens -	All Studies Biospecimens - R	esiduals (Pro0010)	5316) Protocols and Dat	a Dictionary Pa	rticipants Per Study
		Participants by	Protocol		Specimens by P	Protocol	
2	This view provides a list of protocols included in the IDR with a brief description of each.	Protocol #	Short Title	PI	Description		
		Pro00105316	Duke Shared Data and Specimen Repository regarding COVID Patients	Chris Woods	The aim of this data retaining of relevant Academic Medical C benefit is that centra rather than each tim with COVID-19 (or s have its own separat	and spe t patient lenters a alized, sh le a give uspecter te IRB ap	
3 Scroll down the page to view the data dictionary		Download PDF (Data Dictionary	Searc	h the data dictionary		
	Users may download a		Category ≑ 👻 Data Field	Data Lab	el 💠 🛛 FHIR Resource 🗘 👻	FHIR Value	Description
	PDF version of the full		Demographics MRN		Patient	identifienvalue	The unique ID of the patient's Duke N
	data dictionary.		Demographics PAT_ID		Patient	identifienvalue	The unique EPIC ID of the patient
	Users may also search for specific terms in the data dictionary						

Table 1

1	Select "Table 1" from left navigation menu	 Hor Bios Bios Prot Prot Table 	COVID-IDR me specimens - All Studies Biospecimens - Residuals (Pro00105316) tocols & Dictionary ticipants Per Study				
2	This view provides a "Table 1" to summarize all data across all protocols.		Table 1				
			Specimens		94,104		
			Patients		5,231		
			Specimen Types				
			RNA, Total			74	
			Plasma			26,012	
			Tissue, OCT			63	
			Saliva			437	

Permissions

Researchers may send an email to <u>mailto:CovidIDRandSpecimensRequest@dm.duke.edu</u> to request access to the IDR.

Researchers will only be granted access to view aggregated data unless they provide an RPR or protocol information. The staff will review the RPR or the protocol information in iRIS to confirm the appropriate authorization prior to adding to the IDR. Users will see the following based on their authorization level.

Role	Authorizations
Any faculty or staff having a valid Duke Net ID (excludes affiliate accounts)	 View aggregated data with counts of participants and specimens by protocol Define a cohort based on specific de-identified clinical parameters such as labs, diagnoses, and demographics; save cohort; view aggregated counts
Research Preparatory to Research	All authorizations listed above plus
(RPR)	 View columns of de-identified data (no download capabilities)
Researcher with IRB Protocol	All authorization listed above plus
Determined Exempt or Approved	 View and download de-identified data for individual participants
Researcher with Approved IRB	All authorization listed above plus
Protocol	View fully identified data
	 Receive download of fully identified data with Honest Broker assistance
Application Administrator	Full access to data for support to the user community and the application
Honest Broker	Full access to support the needs of the data management
Governance Committee	Elevated access to allow access to datasets subject to review by the committee