

DPCC Data Standard Reference for Viral Isolate Data v2.1

	Project_Identifier	Contributing_Institution	Sample_Identifier	Embargo_End_Date
Input Type	Text Field	Text Field	Text Field	Date Field
Definition	A unique Project Identifier generated by the DPCC by combining the Center-generated Project Code and a random 4-digit number	The institution code for the group that PERFORMED THE ANALYSIS. Must be one of the CEIRS institution codes assigned by NIAID	Identifier initially assigned to each sample collected. If multiple samples are taken from the same host, each sample should have its own identifier.	In case an embargo of the information is needed, the date that the information should be released to the public databases by the DPCC
Format	Project_Code_XXXX Maximum length: 21 characters	Center three-letter code followed by three digits. Maximum length: 6 characters	Center-specific Allowed characters include alphanumeric, hyphen, and underscore: a-z, A-Z, 0-9, -, _ Maximum length: 50 characters	DD-Mon-YYYY DD-Mon-YY NA Maximum length: 11 characters
Value List	None	None	None	Date NA
Curation	The entry must be a Project Identifier value registered with the DPCC.	The entry must be an Institution Code value registered with the DPCC. Center 3-letter codes are case-sensitive and must be entered in all-caps.	The Sample_Identifier initially assigned to the surveillance sample must be provided.	1. Leading 0 in DD is optional. 2. Month must match the first three letters of the month. Month is NOT case sensitive. 3. Years may have two or four digits. 4. Date must conform to NIAID data release policies.
Examples	SJCProj02_4001	SJC101	22258468	3-Mar-2011, 03-Mar-2011, 03-MAR-2011, 3-MAR-11, or NA
Notes		A different Contributing_Institution may have performed the original collection of the surveillance sample that was analyzed.	Submissions without a corresponding surveillance submission will be listed as orphaned records.	An exact date must be provided. If Embargo_End_Date is NA, information will be released without delay. Embargo_End_Date cannot be more than 12 months after submission.
Dependent Fields				
Validation	Project_Identifier should be a valid project identifier.	Must be an Institution Code value registered with the DPCC	Validate field length	Embargo_End_Date must be exact date and cannot be a past date.
Message Code	Error_9_PROJECT_NOT_FOUND	Error_1_INVALID_VALUE	Error_70_INVALID_FIELD_LENGTH	Error_7_INVALID_EMBARGO_DATE
Validation	Project_Identifier exists but user does not have permission to access or edit the project.			Embargo_End_Date cannot be more than 12 months after submission.
Message Code	Error_4_DENIED_USER_ACCESS			Error_100_EMBARGO_DATEError_12M
Validation				
Message Code				
Validation				
Message Code				
Validation				
Message Code				
Validation				
Message Code				
Validation				
Message Code				

	Virus_Identifier	Influenza_Test_Type	Influenza_Test_Result	Influenza_Test_Interpretation
Input Type	Text Field	Text Field	Text Field	Text Field
Definition	A unique laboratory identifier assigned to this virus by the collector.	Assay and target used for virological assessment, as listed in the DPCC Data Dictionary	The numerical result(s) of Influenza_Test_Type	The Positive, Negative, or Inconclusive interpretation of the Influenza_Test_Result value(s)
Format	Text Allowed characters include alphanumeric, hyphen, and underscore: a-z, A-Z, 0-9, -, _ Maximum length: 50 characters	Assay/Analyte Maximum length: 200 characters	Text Maximum length: 50 characters	Text Maximum length: 50 characters
Value List	Text NA	DPCC Data Dictionary NA	Number U NA	P N U NA
Curation	None	The entry must be one or more comma-separated members of the Value List. If multiple tests are performed, separate each Assay/Analyte combination with a comma. For each assay performed, the target or analyte must be specified after the / character. Enter NA for serum samples.	For each assay listed under Influenza_Test_Type, provide a single numerical value. If multiple comma-separated tests are listed under Influenza_Test_Type, Influenza_Test_Result must list the same number of comma-separated test results (see notes for examples). Enter U when quantitative assay results are not known. Enter NA if the value under Influenza_Test_Type is NA.	-
Examples	100245	RRT-PCR/MP	24.5	P
Notes	Provides an additional layer of tracking if Strain_Name is not unique or multiple viruses were isolated from a single surveillance sample.	Please reference the DPCC Data Dictionary for the Assay and Analyte allowed values.	Example: if Influenza_Test_Type lists RRT-PCR/HA,VI/MDCK then Influenza_Test_Result would read 24.5,30.7	P = Positive N = Negative U = Unknown or inconclusive NA = Not Applicable Example: if Influenza_Test_Type lists RRT-PCR/MP,VI/MDCK then Influenza_Test_Interpretation would read P,N
Dependent Fields		Influenza_Test_Result, Influenza_Test_Interpretation, Definition_of_Positive_Sample	Influenza_Test_Type, Influenza_Test_Interpretation, Definition_of_Positive_Sample	Influenza_Test_Type, Influenza_Test_Result, Definition_of_Positive_Sample
Validation	Validate field length	Fields Influenza_Test_Type, Influenza_Test_Antigen,Influenza_Test_Result,Influenza_Test_Interpretation must have same number of values	Fields Influenza_Test_Type, Influenza_Test_Antigen,Influenza_Test_Result,Influenza_Test_Interpretation must have same number of values	Fields Influenza_Test_Type, Influenza_Test_Antigen,Influenza_Test_Result,Influenza_Test_Interpretation must have same number of values
Message Code	Error_70_INVALID_FIELD_LENGTH	Error_63_INVALID_NUMBER_ENTRIES	Error_63_INVALID_NUMBER_ENTRIES	Error_63_INVALID_NUMBER_ENTRIES
Validation		Validate field length	Validate field length	Field value should be one of valid values as in list.
Message Code		Error_70_INVALID_FIELD_LENGTH	Error_70_INVALID_FIELD_LENGTH	Error_1_INVALID_VALUE
Validation		Values must reference DPCC Data Dictionary.	Value must be number, U, or NA	
Message Code		Error_13_REFER_TO_DATA_DICTIONARY	Error_86_INVALID_NUM_U_NA	
Validation		Influenza_Test_Type should follow format assay/analyte		
Message Code		Error_101_TEST_TYPE_FORMAT		
Validation				
Message Code				
Validation				
Message Code				
Validation				
Message Code				

	Definition_of_Positive_Sample	Strain_Name	H5N1	Genbank_Accession_Numbers
Input Type	Text Field	Text Field	Text Field	Text Field
Definition	Provide the CT values that the investigator is classifying as a positive influenza sample.	The WHO strain name of the virus that was isolated from the sample	For highly pathogenic H5N1 viruses include WHO/OIE/FAO clade information.	The accession numbers of virus isolate sequences previously submitted to GenBank.
Format	Text Maximum length: 100 characters	Antigenic Type/Host of Origin/Geographical Origin/Strain Number/Year of Isolation (Subtype) Maximum length: 150 characters	Text Maximum length: 50 characters	Text Maximum length: 1000 characters
Value List	Text NA	Text U NA	Text U NA	Text NA
Curation	If multiple comma-separated tests are listed under Influenza_Test_Type, Definition_of_Positive_Sample must list the same number of comma-separated test results. Enter NA if the Definition_of_Positive_Sample is unavailable or Influenza_Test_Interpretation is negative or unknown.	Must only be provided when there is a viral isolate, otherwise use U for samples that are PCR positive but there is no viral isolate. WHO strain naming convention: Fields must be ordered as follows and separated with the '/' character: 1. The antigenic type (e.g., A, B, C, D) 2. The host of origin (e.g., swine, equine, chicken. For human-origin viruses, no host of origin designation is given.) 3. Geographical origin (e.g., Denver, Taiwan) 4. Strain number (e.g., 15, 7) 5. Year of isolation (e.g., 2009, 1934) 6. For influenza A viruses, the hemagglutinin and neuraminidase antigen description in parentheses (e.g., (H1N1), (H3N2))	None	Validate GenBank number with Strain_Name and gene segment. Enter NA for cases where no sequences are associated.
Examples	35	A/Hong Kong/1/1968 (H3N2), A/chicken/Fujian/4/2002 (H3N6), A/chicken/Fujian/4/2002 (HxNx), A/chicken/Fujian/4/2002 (mixed), A/swine/Iowa/233-56/2011 (H3N2v), A/duck/Alberta/35/1976 (H1N1), B/Hong Kong/432/2014, or C/Texas/19876/2011	1.1	HM006759
Notes		Reference: http://www.cdc.gov/flu/about/viruses/types.htm (HxNx) can be used in cases where a partial subtype has been determined (e.g., H5Nx). If there are mixed subtypes contained within a sample use A/chicken/Fujian/4/2002 (mixed) for Strain_Name or list multiple comma-separated Strain_Names: A/mallard/Alaska/2/2007 (H3N8),A/mallard/Alaska/2/2007 (H3N9).	Enter U if virus clade information is unknown. Enter NA if the isolate is not an HPAI H5N1 virus.	Enter all GenBank accession numbers associated with each virus isolate.
Dependent Fields	Influenza_Test_Type, Influenza_Test_Result, Influenza_Test_Interpretation			
Validation	Fields Influenza_Test_Type, Influenza_Test_Antigen, Influenza_Test_Result, Influenza_Test_Interpretation must have same number of values	Validate field length	Validate field length	Validate field length
Message Code	Error_63_INVALID_NUMBER_ENTRIES	Error_70_INVALID_FIELD_LENGTH	Error_70_INVALID_FIELD_LENGTH	Error_70_INVALID_FIELD_LENGTH
Validation	Validate field length	Host of origin element of strain name must be alphanumeric, underscore, dash, period, or single quote: a-z, A-Z, 0-9, _, -, ., ' , ,'		
Message Code	Error_70_INVALID_FIELD_LENGTH	Error_137_INVALID_HOST_SPECIES_STRAIN_TEXT		
Validation		Geographical origin element of strain name must be alphanumeric, underscore, dash, period, or single quote: a-z, A-Z, 0-9, _, -, ., ' , ,'		
Message Code		Error_138_INVALID_LOCATION_STRAIN_TEXT		
Validation		Strain number element of strain name must be alphanumeric, underscore, dash, or period: a-z, A-Z, 0-9, _, -, . , , ,		
Message Code		Error_139_INVALID_STRAIN_NUMBER_STRAIN_TEXT		
Validation		Year of isolation element of strain name must be 4-digit year.		
Message Code		Error_110_INVALID_STRAIN_YEAR		
Validation		If strain is influenza A, subtype must be present as text between parentheses.		
Message Code		Error_112_INVALID_STRAIN_SUBTYPE		
Validation		Brackets cannot be present if strain does not have rg-prefix.		
Message Code		Error_119_INVALID_STRAIN_BRACKETS		

	Comments
Input Type	Text Field
Definition	Text describing anything else of interest related to the submission
Format	Text Maximum length: 2000 characters
Value List	Text NA
Curation	None
Examples	NA
Notes	If there are no comments, enter NA.
Dependent Fields	
Validation	Validate field length
Message Code	Error_70_INVALID_FIELD_LENGTH
Validation	
Message Code	
Validation	
Message Code	
Validation	
Message Code	
Validation	
Message Code	
Validation	
Message Code	
Validation	
Message Code	
Validation	
Message Code	
Validation	
Message Code	
Validation	
Message Code	
Validation	
Message Code	