

iDPCC Data Standard Reference for Subject Inclusion v1.1

	Project_Identifier	Fiscal_Year
Input Type	Text Field	Text Field
Definition	A unique Project Identifier generated by the iDPCC by combining the Center-generated Project Code and a random 4-digit number	The U.S. Government Federal fiscal year when the subject was enrolled in the study.
Format	Project_Code_XXXX Maximum length: 21 characters	FYXX Maximum length: 4 characters
Value List	None	None
Curation	The entry must be a Project Identifier value registered with the iDPCC.	Entry must be in the format FY, followed by two numbers, such as 16
Examples	20-0058	FY16
Notes	None	The Federal fiscal year begins October 1 and ends September 30, with identification of a fiscal year as the calendar year in which the fiscal year ends. For example: FY16 spans October 1, 2015 to
Dependent Fields		
Validation	Project_Identifier should be a valid project identifier.	Validate field length
Message Code	Error_9_PROJECT_NOT_FOUND	Error_70_INVALID_FIELD_LENGTH
Validation	Project_Identifier exists but user does not have permission to access or edit the project.	
Message Code	Error_4_DENIED_USER_ACCESS	

INSDC term (top level)	INSDC term (lower level)
Not applicable	
Missing	Not Collected
Missing	Not Provided
Missing	Restricted Access
http://www.ebi.ac.uk/ena/about/missing-values-reporting	

	Collection_Country	Subject_Unique_Identifier
Input Type	Text Field	Text Field
Definition	Country in which the original sample was collected, as listed in the iDPCC Country Codes list	An identifier assigned by the CEIRR surveillance sites to maintain subject privacy. If multiple samples are collected from a subject, then all must be linked to the same unique identifier.
Format	Text Maximum length: 3 characters	Text Allowed characters include alphanumeric, hyphen, and underscore: a-z, A-Z, 0-9, -, _ Maximum length: 50 characters
Value List	ISO 3166 Standard Country Code U	Text NA
Curation	The entry must be one and only one member of the Value List. The iDPCC will accept only ISO Country Code in all-caps.	None
Examples	VNM, or USA	1001F_C100F
Notes	The ISO 3166 Standard for country codes may be found on the iDPCC Data Standards Page.	None
Dependent Fields		
Validation	Field value should be a valid ISO three-letter country code or U	Validate field length
Message Code Validation	Error_1_INVALID_VALUE	Error_70_INVALID_FIELD_LENGTH
Message Code		

INSDC Definition
Information is inappropriate to report, can indicate that the standard itself fails to Information of an expected format was not given because it has not been collected Information of an expected format was not given, a value may be given at the later Information exists but can not be released openly because of privacy concerns

	Subject_Gender	Subject_Ethnicity
Input Type	Text Field	Text Field
Definition	The subject's self-definition of their gender from among the given choices.	Ethnicity of the subject, in the sense used by the US Census in which "Hispanic or Latino" is an ethnicity, as opposed to a race.
Format	Text Maximum length: 25 characters	Text Maximum length: 25 characters
Value List	M F Not Collected Not Provided Restricted Access	Hispanic or Latino Not Hispanic or Latino Not Collected Not Provided Restricted Access
Curation	The entry must be one and only one member of the Value List.	The entry must be one and only one member of the Value List.
Examples	M	Hispanic or Latino
Notes	M = Male F = Female	For clarification regarding ethnicity reporting requirement please see the following link: https://www.federalregister.gov/documents/1997/10/30/97-28653/revisions-to-the-standards-for-the-classification-of-federal-data-on-race-and-ethnicity
Dependent Fields		
Validation	Field value should be one of valid values as in list.	Field value should be one of valid values as in list.
Message Code Validation	Error_1_INVALID_VALUE	Error_1_INVALID_VALUE
Message Code		

	Subject_Race	Subject_Age
Input Type	Text Field	Text Field
Definition	The subject's self-definition of their race from among the given choices.	The subject's age at the time of enrollment in years.
Format	Text Maximum length: 150 characters	Text Maximum length: 17 characters
Value List	American Indian or Alaska Native Asian Black or African American Native Hawaiian or Other Pacific Islander White More Than One Race Not Collected Not Provided Restricted Access	Number Not Collected Not Provided Restricted Access
Curation	The entry must be one and only one member of the Value List.	Enter Subject_Age as a whole number in years.
Examples	White	19, or 40
Notes	For clarification regarding race reporting requirement please see the following link: https://www.federalregister.gov/documents/1997/10/30/97-28653/revisions-to-the-standards-for-the-classification-of-federal-data-on-race-and-ethnicity	Enter Subject_Age as a whole number in years. All ages greater than 90 must be entered as >90 to de-identify individual health information.
Dependent Fields		
Validation	Field value should be one or more of valid values as in list.	Validate field length
Message Code	Error_1_INVALID_VALUE	Error_70_INVALID_FIELD_LENGTH
Validation		All values greater than 90 must be entered as >90
Message Code		Error_90_INVALID_AGE