TABLE 1: USE CASE 1 - DATASET GOVERNANCE

[Legend: Blank cell in Table 1 = information not available/found; N/A = information confirmed to not exist]

	Governance	Dataset 1 - NHANES					Dataset 2 - NSDUH					Dataset 3 - MTF					Dataset 4 - AFCARS				
		Data Collection	Data Linkage	Data Sharing	Data Access	Data Use	Data Collection	Data Linkage	Data Sharing	Data Access	Data Use	Data Collection	Data Linkage	Data Sharing	Data Access	Data Use	Data Collection	Data Linkage	Data Sharing	Data Access	Data Use
1	Authorization/s																				
1.1	Assent	Assent from children authorizes data collection	Assent from children authorizes data linkage	Does not authorize/specify	Does not authorize/specify	Does not authorize/specify	Assent from children authorizes data collection					Assent from children authorizes data collection	Does not authorize/specify				N/A	N/A	N/A	N/A	N/A
1.2	Consent	Consent from adults authorizes data collection	Consent from adults authorizes data linkage	Does not authorize/specify	Does not authorize/specify	Consent from adults authorizes data use	Consent from adults authorizes data collection					Consent from parents authorizes data collection	Does not authorize/specify			Consent from adults authorizes data use	N/A	N/A	N/A	N/A	N/A
1.3	IRB/equivalent Privacy Board determination	NCHS ERB (IRB) approval authorizes data collection	NCHS ERB approval authorizes data linkage	NCHS Disclosure Review Board (DRB) authorizes data sharing	N/A	NCHS ERB (IRB) approval authorizes data use	RTI (DCC for NSDUH) IRB authorizes data collection					MTF (U-Mich) IRB approval authorizes data collection	Does not authorize/specify	Two IRBs authorize data sharing: 1. ICPSR (U-Mich) IRB 2. MTF (U-Mich) IRB			N/A	N/A	N/A	N/A	N/A
1.4	Local/state/federal law	Four federal laws authorizes data collection: 1. Section 306 of the Public Health Service Act (42 U.S.C. 242k) 2. National Nutrition Monitoring and Related Research Act of 1990 (P.L. 101-445), (October 22, 1990) 3. Food Quality Protection Act of 1996 (P.L. 104-170) 4. Federal Food, Drug, and Cosmetic Act (21 USC 393), Chapter 9		Confidential Information Protection and Statistical Efficiency Act (CIPSEA) authorizes data sharing through the NCHS RDC	N/A	Designated Agent Agreement (Non-Disclosure CIPSEA Agent Form) authorizes data use	Public Health Service Act Section 505 authorizes data collection						Does not authorize/specify	Family Educational Rights and Privacy Act (FERPA) authorize data sharing			Social Security Act (Section 479) authorizes data collection by states and Tribal agencies	N/A			
1.5	Institutional Certification												Does not authorize/specify				N/A	N/A	N/A	N/A	N/A
1.6	Data originator agreement			Two agreements authorize data sharing from NCHS/NHANES collaborators that are authorized to review pre-release data that they contributed: 1. NCHS non-disclosure affidavit 2. Data Sharing Agreement									Does not authorize/specify	N/A			N/A	N/A	N/A	N/A	N/A

TABLE 1: USE CASE 1 - DATASET GOVERNANCE

[Legend: Blank cell in Table 1 = information not available/found; N/A = information confirmed to not exist]

	Governance	Dataset 1 - NHANES					Dataset 2 - NSDUH					Dataset 3 - MTF					Dataset 4 - AFCARS				
		Data Collection	Data Linkage	Data Sharing	Data Access	Data Use	Data Collection	Data Linkage	Data Sharing	Data Access	Data Use	Data Collection	Data Linkage	Data Sharing	Data Access	Data Use	Data Collection	Data Linkage	Data Sharing	Data Access	Data Use
1.7	Repository agreements/policies				N/A	Data Use/Access Agreement (Rules of Behavior) authorizes data use				Two repository agrements authorize data access: 1. Data Access Agreement (DAA) form 2. For students, the SAMHSA RDC Student Data User Acknowledgement form signed by the student and their advisor	RDC Data Access Agreement (DAA) form authorizes data use		Does not authorize/specify		NAHDAP Restricted Data Use Agreement for Restricted Data in the Virtual Data Enclave (NAHDAP VDE RDUA) authorizes data access	NAHDAP Restricted Data Use Agreement for Restricted Data in the Virtual Data Enclave (NAHDAP VDE RDUA) authorizes data use		N/A	N/A	1. Contractual agreement between Children's Bureau and NDACAN 2. NDACAN Terms of Use Agreement	NDACAN Terms of Use Agreement authorizes data use
1.8	Other (specify)		DHANES/NHANES approval authorizes data linkage		Data Use/Access Agreement (Rules of Behavior) authorizes data access	Data Use/Access Agreement (Rules of Behavior) authorizes data use								MTF PI authorizes data sharing (through the ICPSR VDE)	MTF PI authorizes data access (through the ICPSR VDE)			Does not authorize/specify	Contractual agreement between NDACAN and Children's Bureau authorizes data sharing	Two repository agreements authorize data access: 1. Contractual agreement between Children's Bureau and NDACAN 2. NDACAN Terms of Use Agreement	
2	**Applicable Regulations/Policies**																				
2.1	Local regulations/policies	N/A		NHANES Protected Data Policy	1. NHANES Protected Data Policy 2. NCHS RDC policy	NCHS RDC policy	SAMHSA recommendation														
2.2	Tribal regulations/policies	N/A	N/A	N/A	N/A	N/A												N/A	N/A	N/A	N/A
2.3	State regulations/policies	N/A									Confidential Information Protection and Statistical Efficiency Act (CIPSEA)										
2.4	Federal regulations/policies	1. Section 306 of the Public Health Service Act (42 U.S.C. 242k) 2. National Nutrition Monitoring and Related Research Act of 1990 (P.L. 101-445), (October 22, 1990) 3. Food Quality Protection Act of 1996 (P.L. 104-170) 4. Federal Food, Drug, and Cosmetic Act (21 USC 393), Chapter 9		1. Section 308(d) Public Health Act 2. Confidential Information Protection and Statistical Efficiency Act (CIPSEA)	1. Section 308(d) of the Public Health Service Act 2. Confidential Information Protection and Statistical Efficiency Act (CIPSEA)	1. Section 308(d) of the Public Health Service Act 2. Confidential Information Protection and Statistical Efficiency Act (CIPSEA)	Confidential Information Protection and Statistical Efficiency Act (CIPSEA)		Confidential Information Protection and Statistical Efficiency Act (CIPSEA)	Confidential Information Protection and Statistical Efficiency Act (CIPSEA)				1. Grant of Confidentality from the U.S. Department of Justice 2. Family Educational Rights and Privacy Act (FERPA)		Grant of Confidentality from the U.S. Department of Justice	45 CFR 1355.40 D46				
2.5	International regulations/policies	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A

TABLE 1: USE CASE 1 - DATASET GOVERNANCE

[Legend: Blank cell in Table 1 = information not available/found; N/A = information confirmed to not exist]

	Governance	Dataset 1 - NHANES					Dataset 2 - NSDUH					Dataset 3 - MTF					Dataset 4 - AFCARS				
		Data Collection	Data Linkage	Data Sharing	Data Access	Data Use	Data Collection	Data Linkage	Data Sharing	Data Access	Data Use	Data Collection	Data Linkage	Data Sharing	Data Access	Data Use	Data Collection	Data Linkage	Data Sharing	Data Access	Data Use
2.6	Contractual obligations	N/A		N/A												N/A	N/A	N/A			
2.7	Repository policies	N/A			N/A	NCHS RDC policy				SAMHSA RDC policy	SAMHSA RDC policy				NAHDAP/ICPSR VDE policy	NAHDAP/ICPSR VDE policy	N/A	N/A		NDACAN policy	NDACAN policy
3	Data Linking/Sharing/Access/Use Governance Based on Authorizations and Applicable Regulations/Policies																				
3.1	Whether the data can be linked	Assent/Consent specifies that the data can be linked	Assent/Consent, NCHS ERB approval, and DHANES/HANES approval specify that the data can be linked	Does not authorize/specify	Does not authorize/specify	Does not authorize/specify	Does not authorize/specify					Does not authorize/specify	Does not authorize/specify	Does not authorize/specify	Does not authorize/specify	Does not authorize/specify	Does not authorize/specify		Does not authorize/specify	Does not authorize/specify	Does not authorize/specify
3.2	With what other data can it be linked or can it not be linked (scope of linkage)	Assent/Consent specifies that the data can be linked to vital statistics, health, nutrition, and other related records	Assent/Consent specifies that the data can be linked to vital statistics, health, nutrition, and other related records	Does not authorize/specify	Does not authorize/specify	Does not authorize/specify	Does not authorize/specify					Does not authorize/specify	Does not authorize/specify	Does not authorize/specify	Does not authorize/specify	Does not authorize/specify	Does not authorize/specify		Does not authorize/specify	Does not authorize/specify	Does not authorize/specify
3.3	Whether data can be shared	Does not authorize/specify	Does not authorize/specify	Confidential Information Protection and Statistical Efficiency Act (CIPSEA) specifies that data can be shared through NCHS RDC (authorized by CIPSEA to share the data)	Does not authorize/specify	Does not authorize/specify							Does not authorize/specify	ICPSR (U-Mich) IRB and MTF (U-Mich) IRB approval specify that data can be shared through ICPSR/NAHDAP	Does not authorize/specify	Does not authorize/specify			Contractual agreement between NDACAN and Children's Bureau specifies that AFCARS data can be shared through NDACAN.	Contractual agreement between NDACAN and Children's Bureau specifies that AFCARS data can be shared through NDACAN	
3.4	How data can be shared (de-identification status, disclosure review)	Does not authorize/specify	Does not authorize/specify	NCHS RDC specifies that data shared through RDC: 1. Must be de-identified of all direct identifiers and certain indirect identifiers (geography) may be included 2. Must undergo disclosure review by NCHS Disclosure Review Board/NCHS Confidentiality Officer prior to sharing the restricted-use data through the RDC, and then the RDC and DHANES reviews the output before releasing that output	NCHS RDC specifies that data shared through RDC: 1. Must be de-identified of all direct identifiers and certain indirect identifiers (geography) may be included 2. Must undergo disclosure review by NCHS Disclosure Review Board/NCHS Confidentiality Officer prior to sharing the restricted-use data through the RDC, and then the RDC and DHANES reviews the output before releasing that output	Does not authorize/specify							Does not authorize/specify	FERPA and ICPSR/NAHDAP specify that data shared through ICPSR VDE: 1. Must be fully de-identified (for MTF restricted-use data, this does not include state and zipcode) 2. Must undergo disclosure review by ICPSR/NAHDAP staff prior to sharing, and disclosure review of analysis outputs is required prior to removing output data from the VDE	ICPSR/NAHDAP specifies that disclosure review of analysis outputs is required prior to removing output data from the VDE.	Does not authorize/specify			Contractual agreement between NDACAN and Children's Bureau specifies that AFCARS data shared through NDACAN: 1. Must be de-identified of all 18 HIPAA identifiers 2. Must undergo disclosure review by NDACAN staff will conduct disclosure review of data prior to sharing (removing county FIPS code with >1,000 records, recode DoB to the 15th and adjust all other dates accordingly)	Contractual agreement between NDACAN and Children's Bureau specifies that AFCARS data shared through NDACAN: 1. Data must be de-identified of all 18 HIPAA identifiers 2. Data must undergo disclosure review of data prior to sharing (removing county FIPS code with >1,000 records, recode DoB to the 15th and adjust all other dates accordingly)	

TABLE 1: USE CASE 1 - DATASET GOVERNANCE

[Legend: Blank cell in Table 1 = information not available/found; N/A = information confirmed to not exist]

	Governance	Dataset 1 - NHANES					Dataset 2 - NSDUH					Dataset 3 - MTF					Dataset 4 - AFCARS				
		Data Collection	Data Linkage	Data Sharing	Data Access	Data Use	Data Collection	Data Linkage	Data Sharing	Data Access	Data Use	Data Collection	Data Linkage	Data Sharing	Data Access	Data Use	Data Collection	Data Linkage	Data Sharing	Data Access	Data Use
3.5	How data can be accessed (access type, data use agreement, data access committee/group approval, IRB LOD, etc.)	Does not authorize/specify	Does not authorize/specify	NHANES Protected Data Policy specifies that the restricted use data must be accessed through the NCHS RDC (on-site enclave)	NCHS RDC specifies that for data access, a user: 1. Must obtain review and approval by NHCS Confidentiality Officer, NCHS RDC, and DHANES/NHANES on the proposed research 2. Must execute Data Use/Access Agreement (Rules of Behavior) 3. Must sign Designated Agent Agreement (Non-Disclosure CIPSEA Agent Form) 4. Must undergo confidentiality training 5. Must access data through the NCHS RDC (on-site enclave)	Does not authorize/specify			SAMHSA RDC specifies that for data access, a user: 1. Submit RDC application 2. Obtain approval from SAMHSA staff on the research proposed in the RDC application 3. Sign and complete Designated Agent Form (DAF) 4. Sign and complete Data Access Agreement (DAA) form 5. Complete confidentiality training 6. Students and their advisors must also sign the SAMHSA RDC Student Data User Acknowledgement form 7. Access data only through RDC		Does not authorize/specify	Does not authorize/specify	ICPSR/NAHDAP specifies that for data access, a user: 1. Must execute NAHDAP VDE RDUA between ICPSR (U-Mich) and the researcher's institution 2. Must only access data through the ICPSR VDE (virtual enclave) 3. Must obtain IRB approval or exemption from the researcher's institution 4. Must obtain review and approval from NAHDAP on the proposed research	ICPSR/NAHDAP specifies that for data accessed through the ICPSR VDE, a user: 1. Must execute NAHDAP VDE RDUA between ICPSR (U-Mich) and the researcher's institution 2. Must only access data through the ICPSR VDE (virtual enclave) 3. Must obtain IRB approval or exemption from the researcher's institution 4. Must obtain review and approval from NAHDAP on the proposed research	Does not authorize/specify				Contractural agreement between NDACAN and Children's Bureau specifies that for data accessed: 1. User must execute of the NDACAN Terms of Use Agreement 2. User must obtain review and approval from NDACAN staff on the proposed research	
3.6	How data can be used (including data use limitations)	Assent/Consent specifies that the data can be used for statistical reporting and analysis (includes broad research)	Does not authorize/specify	Section 308(d) of the Public Health Service Act specifies that NCHS data containing personal information cannot be used for any purpose other than what was described to survey participants. CIPSEA specifies that NCHS data can only be used for statistical and research purposes.	Section 308(d) of the Public Health Service Act specifies that NCHS data containing personal information cannot be used for any purpose other than what was described to survey participants. CIPSEA specifies that NCHS data can only be used for statistical and research purposes.	1. Consent, Section 308(d) Public Health Service Act [42 U.S.C. 242m(d)], and Confidential Information Protection and Statistical Efficiency Act (CIPSEA) specify that the data can only be used for statistical reporting and analysis (includes broad research). 2. RDC DUA specifies that the data were collected with the assurance that they will be used only for health statistical reporting and analysis.					RDC DAA specifies that the data were collected with the assurance that they will be used only for health statistical reporting and analysis	Consent specifies that the data can be used for broad research	Does not authorize/specify	NAHDAP/ICPSR specifies that the data can be used for broad research	NAHDAP/ICPSR specifies that the data can be used for broad research	Consent and the NAHDAP/ICPSR RDUA specify that the data be used solely for research or statistical purposes			Does not authorize/specify	NDACAN Terms of Use Agreement specifies that the data can be used by researchers in accordance with their approved research described in Section I.1 of the Terms of Use Agreement	NDACAN Terms of Use Agreement specifies that the data can be used by researchers in accordance with their approved research described in Section I.1 of the Terms of Use Agreement

TABLE 1: USE CASE 1 - DATASET GOVERNANCE

[Legend: Blank cell in Table 1 = information not available/found; N/A = information confirmed to not exist]

	Governance	Dataset 1 - NHANES					Dataset 2 - NSDUH					Dataset 3 - MTF					Dataset 4 - AFCARS				
		Data Collection	Data Linkage	Data Sharing	Data Access	Data Use	Data Collection	Data Linkage	Data Sharing	Data Access	Data Use	Data Collection	Data Linkage	Data Sharing	Data Access	Data Use	Data Collection	Data Linkage	Data Sharing	Data Access	Data Use
3.7	Other (specify)	Does not authorize/specify	Does not authorize/specify	Does not authorize/specify		Does not authorize/specify	SAMHSA recommends against comparing 2020 NSDUH data with prior years due to methodological changes														

TABLE 2: USE CASE 1 - DATASET LINKAGE

Definitions: Limitations are restrictions on data linkage and use (e.g., dataset must only be linked with other disease-relevant data, dataset must be used in a physical enclave, etc.). **Controls** are processes established to ensure compliance with governance for data sharing, access, and use (e.g., user must access data in a physical enclave, user must sign data use agreement, user must receive data access committee approval, etc.). **Authorization gaps exist** when there is no explicit authorization or the authorization is not available or found in the information collected by the project team for various data life cycle stages (data collection, linking, sharing, access and use).

		Can the datasets be linked?	What limitations do the linked datasets inherit?	What controls do the linked dataset require?	What authorization gaps exist?
1	**NHANES (Dataset 1) and NSDUH (Dataset 2) linkage**	**Yes, NHANES and NSDUH can be linked provided:** **A. NCHS RDC staff:** 1. Shares de-identified data except for certain indirect identifiers (geography)for NHANES through the RDC [Control 1a] 2. Performs disclosure review by NCHS Disclosure Review Board/NCHS Confidentiality Officer prior to sharing the NHANES restricted-use data through the RDC [Control 1b] 3. Performs disclosure review of the output by RDC and DHANES before releasing that output [Control 1c] **B. The researcher/user:** 1. Obtains authorization for data linkage and sharing for NSDUH [Authorization gap 2a] - Assumption 2. Links NHANES dataset only to vital statistics, health, nutrition, and other related records [Limitation 1a] 3. Uses NHANES and NSDUH data within the NCHS RDC [Limitations 1b, 2a, Controls 1h, 2g] 4. Uses the linked NHANES and NSDUH data only for health statistical reporting and analysis [Limitations 1c, 2b] 5. Submits RDC application [Control 2a] 6. Obtains approvals from NCHS Confidentiality Officer, NCHS RDC, DHANES/NHANES, and SAMHSA staff on the proposed linkage [Controls 1d, 2b] 7. Signs the NHANES Data Use/Access Agreement, NHANES Designated Agent Agreement, NSDUH Designated Agent Form, NSDUH Data Access Agreement (DAA); and if the researcher/user is a student, signs the SAMHSA RDC Student Data User Acknowledgement form along with their advisor [Controls 1e, 1f, 2c, 2d, 2f] 8. Completes confidentiality training for NHANES and NSDUH data access [Controls 1g, 2e]	**Researchers/users:** 1a. Can only link NHANES data to vital statistics, health, nutrition, and other related records [NHANES] 1b. Must use NHANES data within the NCHS RDC (on-site enclave) [NHANES] 1c. Must use NHANES data only for health statistical reporting and analysis [NHANES] 2a. Must use NSDUH data within the NCHS RDC [NSDUH] 2b. Must use NSDUH data only for health statistical reporting and analysis [NSDUH]	**For sharing NHANES, NCHS RDC staff must:** 1a. Share NHANES data de-identified of all direct identifiers; certain indirect identifiers (geography) may be included [NHANES] 1b. Perform disclosure review prior to sharing the NHANES restricted-use data through the RDC (NCHS Disclosure Review Board/NCHS Confidentiality Officer) [NHANES] 1c. Perform disclosure review of the output before releasing it (RDC and DHANES) [NHANES] **For accessing NHANES, researchers/users must:** 1d. Obtain approvals from NCHS Confidentiality Officer, NCHS RDC, and DHANES/NHANES on the proposed research [NHANES] 1e. Execute Data Use/Access Agreement (Rules of Behavior) [NHANES] 1f. Sign Designated Agent Agreement [NHANES] 1g. Complete confidentiality training [NHANES] 1h. Access data only within NCHS RDC (on-site enclave) [NHANES] **For accessing NSDUH, researchers/users must:** 2a. Submit RDC application [NSDUH] 2b. Obtain approval from SAMHSA staff on the research proposed in the RDC application [NSDUH] 2c. Sign Designated Agent Form (DAF) [NSDUH] 2d. Sign Data Access Agreement (DAA) [NSDUH] 2e. Complete confidentiality training [NSDUH] 2f. Sign the SAMHSA RDC Student Data User Acknowledgement form and obtain advisor's signature, if researcher/user is a student [NSDUH] 2g. Access data within NCHS RDC [NSDUH]	2a. For NSDUH, information on authorizations for linkage and sharing is not available/found.

TABLE 2: USE CASE 1 - DATASET LINKAGE

Definitions: Limitations are restrictions on data linkage and use (e.g., dataset must only be linked with other disease-relevant data, dataset must be used in a physical enclave, etc.). **Controls** are processes established to ensure compliance with governance for data sharing, access, and use (e.g., user must access data in a physical enclave, user must sign data use agreement, user must receive data access committee approval, etc.). **Authorization gaps** exist when there is no explicit authorization or the authorization is not available or found in the information collected by the project team for various data life cycle stages (data collection, linking, sharing, access and use).

		Can the datasets be linked?	What limitations do the linked datasets inherit?	What controls do the linked dataset require?	What authorization gaps exist?
2	**NHANES (Dataset 1) and MTF (Dataset 3) linkage**	**Yes, NHANES and MTF can be linked provided:** **A. NCHS RDC staff:** 1. Shares de-identified data except for certain indirect identifiers (geography)for NHANES through the RDC [Control 1a] 2. Performs disclosure review by NCHS Disclosure Review Board/NCHS Confidentiality Officer prior to sharing the NHANES restricted-use data through the RDC [Control 1b] 3. Performs disclosure review of the output by RDC and DHANES before releasing that output [Control 1c] **B. ICPSR/NAHDAP staff:** 4. Shares fully de-identified data (for MTF restricted-use data, this does not include state and zipcode) [Control 3a] 5. Performs disclosure review prior to sharing [Control 3b] 6. Performs disclosure review of analysis outputs prior to removing output data from the VDE [Control 3c] **C. The researcher/user:** 1. Obtains authorization for linking MTF data [Authorization gap 3a] - Assumption	**Researchers/users:** 1a. Can only link NHANES data to vital statistics, health, nutrition, and other related records [NHANES] 1b. Must use NHANES data within the NCHS RDC (on-site enclave) [NHANES] 1c. Must use NHANES data only for health statistical reporting and analysis [NHANES] 3a. Must use MTF data for broad research or statistical purposes [MTF]	**For sharing NHANES, NCHS RDC staff must:** 1a. Share NHANES data de-identified of all direct identifiers; certain indirect identifiers (geography) may be included [NHANES] 1b. Perform disclosure review prior to sharing the NHANES restricted-use data through the RDC (NCHS Disclosure Review Board/NCHS Confidentiality Officer) [NHANES] 1c. Perform disclosure review of the output before releasing it (RDC and DHANES) [NHANES] **For sharing MTF, ICPSR/NAHDAP staff must:** 3a. Share fully de-identified data (for MTF restricted-use data, this does not include state and zipcode) [MTF] 3b. Perform disclosure review prior to sharing [MTF] 3c. Perform disclosure review of analysis outputs prior to removing output data from the VDE [MTF] **For accessing NHANES, researchers/users must:** 1d. Obtain approvals from NCHS Confidentiality Officer, NCHS RDC, and DHANES/NHANES on the proposed research [NHANES] 1e. Execute Data Use/Access Agreement (Rules of Behavior) [NHANES] 1f. Sign Designated Agent Agreement [NHANES]	3a. For MTF, information on authorizations for linkage is not available/found.

TABLE 2: USE CASE 1 - DATASET LINKAGE

Definitions: Limitations are restrictions on data linkage and use (e.g., dataset must only be linked with other disease-relevant data, dataset must be used in a physical enclave, etc.). **Controls** are processes established to ensure compliance with governance for data sharing, access, and use (e.g., user must access data in a physical enclave, user must sign data use agreement, user must receive data access committee approval, etc.). **Authorization gaps** exist when there is no explicit authorization or the authorization is not available or found in the information collected by the project team for various data life cycle stages (data collection, linking, sharing, access and use).

	Can the datasets be linked?	What limitations do the linked datasets inherit?	What controls do the linked dataset require?	What authorization gaps exist?
	2. Links NHANES dataset only to vital statistics, health, nutrition, and other related records [Limitation 1a] 3. Uses/access NHANES data within NCHS RDC and obtain approval from ICPSR/NAHDAP to export MTF data into NCHS RDC [Limitations 1b, Controls 1h, 3e] - Assumption for MTF 4. Uses the linked NHANES and MTF data only for statistical purposes [Limitations 1c, 3a] 5. Obtains approvals from NCHS Confidentiality Officer, NCHS RDC, DHANES/NHANES, and ICPSR/NAHDAP staff on the proposed linkage [Controls 1d, 3g] 6. Signs the NHANES Data Use/Access Agreement, NHANES Designated Agent Agreement, and NAHDAP VDE RDUA [Controls 1e, 1f, 3d] 7. Completes confidentiality training for NHANES data access [Control 1g] 8. Obtains IRB approval or exemption from their institution for accessing MTF [Control 3f]		1g. Complete confidentiality training [NHANES] 1h. Access data only within NCHS RDC (on-site enclave) [NHANES] **For accessing MTF, researchers/users must:** 3d. Execute NAHDAP VDE RDUA between ICPSR (U-Mich) and the researcher's institution [MTF] 3e. Access data only through the ICPSR VDE (virtual enclave) [MTF] 3f. Obtain IRB approval or exemption from the researcher's institution for data access [MTF] 3g. Obtain review and approval from NAHDAP on the proposed research [MTF]	

TABLE 2: USE CASE 1 - DATASET LINKAGE

Definitions: Limitations are restrictions on data linkage and use (e.g., dataset must only be linked with other disease-relevant data, dataset must be used in a physical enclave, etc.). **Controls** are processes established to ensure compliance with governance for data sharing, access, and use (e.g., user must access data in a physical enclave, user must sign data use agreement, user must receive data access committee approval, etc.). **Authorization gaps** exist when there is no explicit authorization or the authorization is not available or found in the information collected by the project team for various data life cycle stages (data collection, linking, sharing, access and use).

		Can the datasets be linked?	What limitations do the linked datasets inherit?	What controls do the linked dataset require?	What authorization gaps exist?
3	**NHANES (Dataset 1) and AFCARS (Dataset 4) linkage**	**Yes, NHANES and AFCARS can be linked provided:** **A. NCHS RDC staff:** 1. Shares de-identified data except for certain indirect identifiers (geography)for NHANES through the RDC [Control 1a] 2. Performs disclosure review by NCHS Disclosure Review Board/NCHS Confidentiality Officer prior to sharing the NHANES restricted-use data through the RDC [Control 1b] 3. Performs disclosure review of the output by RDC and DHANES before releasing that output [Control 1c] **B. NDACAN staff:** 4. Shares AFCARS data de-identified of all 18 HIPAA identifiers. [Control 4a] 5. Performs disclosure review of data prior to sharing (removing county FIPS code with >1,000 records, recode DoB to the 15th and adjust all other dates accordingly). [Control 4b] **C. The researcher/user:** 1. Obtains authorization for linking AFCARS data [Authorization gap 4a] - Assumption 2. Links NHANES dataset only to vital statistics, health, nutrition, and other related records	**Researchers/users:** 1a. Can only link NHANES data to vital statistics, health, nutrition, and other related records [NHANES] 1b. Must use NHANES data within the NCHS RDC (on-site enclave) [NHANES] 1c. Must use NHANES data only for health statistical reporting and analysis [NHANES] 4a. Must use AFCARS data in accordance with their approved research described in Section I.1 of the NDACAN Terms of Use Agreement [AFCARS]	**For sharing NHANES, NCHS RDC staff must:** 1a. Share NHANES data de-identified of all direct identifiers; certain indirect identifiers (geography) may be included [NHANES] 1b. Perform disclosure review prior to sharing the NHANES restricted-use data through the RDC (NCHS Disclosure Review Board/NCHS Confidentiality Officer) [NHANES] 1c. Perform disclosure review of the output before releasing it (RDC and DHANES) [NHANES] **For sharing AFCARS, NDACAN staff must:** 4a. Share AFCARS data de-identified of all 18 HIPAA identifiers. [AFCARS] 4b. Perform disclosure review of data prior to sharing (removing county FIPS code with >1,000 records, recode DoB to the 15th and adjust all other dates accordingly). [AFCARS] **For accessing NHANES, researchers/users must:** 1d. Obtain approvals from NCHS Confidentiality Officer, NCHS RDC, and DHANES/NHANES on the proposed research [NHANES] 1e. Execute Data Use/Access Agreement (Rules of Behavior) [NHANES] 1f. Sign Designated Agent Agreement [NHANES] 1g. Complete confidentiality training [NHANES]	4a. For AFCARS, information on authorizations for linkage is not available/found.

TABLE 2: USE CASE 1 - DATASET LINKAGE

Definitions: Limitations are restrictions on data linkage and use (e.g., dataset must only be linked with other disease-relevant data, dataset must be used in a physical enclave, etc.). **Controls** are processes established to ensure compliance with governance for data sharing, access, and use (e.g., user must access data in a physical enclave, user must sign data use agreement, user must receive data access committee approval, etc.). **Authorization gaps** exist when there is no explicit authorization or the authorization is not available or found in the information collected by the project team for various data life cycle stages (data collection, linking, sharing, access and use).

		Can the datasets be linked?	What limitations do the linked datasets inherit?	What controls do the linked dataset require?	What authorization gaps exist?
		[Limitation 1a] 3. Uses/access NHANES and AFCARS data within NCHS RDC [Limitation 1b, Control 1h] 4. Uses the linked NHANES and AFCARS data only for statistical reporting and analysis [Limitations 1c, 4a] 5. Obtains approval from NCHS Confidentiality Officer, NCHS RDC, DHANES/NHANES, and NDACAN staff on the proposed linkage [Controls 1d, 4d] 6. Signs the NHANES Data Use/Access Agreement, NHANES Designated Agent Agreement, NDACAN Terms of Use Agreement [Controls 1e, 1f, 4c] 7. Completes confidentiality training for NHANES data access [Controls 1g]		1h. Access data only within NCHS RDC (on-site enclave) [NHANES] **For accessing AFCARS, researchers/users must:** 4c. Execute of the NDACAN Terms of Use Agreement. [AFCARS] 4d. Obtain review and approval from NDACAN staff on the proposed research. [AFCARS]	

TABLE 2: USE CASE 1 - DATASET LINKAGE

Definitions: Limitations are restrictions on data linkage and use (e.g., dataset must only be linked with other disease-relevant data, dataset must be used in a physical enclave, etc.). **Controls** are processes established to ensure compliance with governance for data sharing, access, and use (e.g., user must access data in a physical enclave, user must sign data use agreement, user must receive data access committee approval, etc.). **Authorization gaps** exist when there is no explicit authorization or the authorization is not available or found in the information collected by the project team for various data life cycle stages (data collection, linking, sharing, access and use).

		Can the datasets be linked?	What limitations do the linked datasets inherit?	What controls do the linked dataset require?	What authorization gaps exist?
4	**NSDUH (Dataset 2) and MTF (Dataset 3) linkage**	**Yes, NSDUH and MTF can be linked provided:** **A. ICPSR/NAHDAP staff:** 1. Shares fully de-identified data (for MTF restricted-use data, this does not include state and zipcode) [Control 3a] 2. Performs disclosure review prior to sharing [Control 3b] 3. Performs disclosure review of analysis outputs prior to removing output data from the VDE [Control 3c] **B. The researcher/user:** 1. Obtains authorization for data linkage and sharing for NSDUH and authorization for linking [Authorization gap 2a, 3a] - ***Assumption*** 2. Uses NSDUH data within NCHS RDC and obtain approval from ICPSR/NAHDAP to export MTF data into NCHS RDC [Limitations 2a, Controls 2g, 3e] - ***Assumption for MTF*** 3. Uses the linked NSDUH and MTF data only for statistical purposes [Limitations 2b, 3a] 5. Submits RDC application [Control 2a] 6. Obtains approvals from SAMHSA staff and NAHDAP staff on the proposed linkage [Controls 2b, 3g] 7. Signs the NSDUH Designated Agent Form, NSDUH Data Access Agreement (DAA); and if the researcher/user is a student, signs the SAMHSA RDC Student Data User Acknowledgement form along with their advisors; and signs NAHDAP VDE RDUA [Controls 2c, 2d, 2f, 3d] 8. Completes confidentiality training for NSDUH data access [Controls 2e] 9. Obtains IRB approval or exemption from their institution for accessing MTF [Control 3f]	**Researchers/users:** 2a. Must use NSDUH data within the NCHS RDC [NSDUH] 2b. Must use NSDUH data only for health statistical reporting and analysis [NSDUH] 3a. Must use MTF data for broad research or statistical purposes [MTF]	**For sharing MTF, ICPSR/NAHDAP staff must:** 3a. Share fully de-identified data (for MTF restricted-use data, this does not include state and zipcode) [MTF] 3b. Perform disclosure review prior to sharing [MTF] 3c. Perform disclosure review of analysis outputs prior to removing output data from the VDE [MTF] **For accessing NSDUH, researchers/users must:** 2a. Submit RDC application [NSDUH] 2b. Obtain approval from SAMHSA staff on the research proposed in the RDC application [NSDUH] 2c. Sign Designated Agent Form (DAF) [NSDUH] 2d. Sign Data Access Agreement (DAA) [NSDUH] 2e. Complete confidentiality training [NSDUH] 2f. Sign the SAMHSA RDC Student Data User Acknowledgement form and obtain advisor's signature, if researcher/user is a student [NSDUH] 2g. Access data within the NCHS RDC [NSDUH] **For accessing MTF, researchers/users must:** 3d. Execute NAHDAP VDE RDUA between ICPSR (U-Mich) and the researcher's institution [MTF] 3e. Access data only through the ICPSR VDE (virtual enclave) [MTF] 3f. Obtain IRB approval or exemption from the researcher's institution [MTF] 3g. Obtain review and approval from NAHDAP on the proposed research [MTF]	2a. For NSDUH, information on authorizations for linkage and sharing is not available/found. 3a. For MTF, information on authorizations for linkage is not available/found.

TABLE 2: USE CASE 1 - DATASET LINKAGE

Definitions: Limitations are restrictions on data linkage and use (e.g., dataset must only be linked with other disease-relevant data, dataset must be used in a physical enclave, etc.). **Controls** are processes established to ensure compliance with governance for data sharing, access, and use (e.g., user must access data in a physical enclave, user must sign data use agreement, user must receive data access committee approval, etc.). **Authorization gaps** exist when there is no explicit authorization or the authorization is not available or found in the information collected by the project team for various data life cycle stages (data collection, linking, sharing, access and use).

		Can the datasets be linked?	What limitations do the linked datasets inherit?	What controls do the linked dataset require?	What authorization gaps exist?
5	**NSDUH (Dataset 2) and AFCARS (Dataset 4) linkage**	**Yes, NSDUH and AFCARS can be linked provided:** **A. NDACAN staff:** 1. Shares AFCARS data de-identified of all 18 HIPAA identifiers. [Control 4a] 2. Performs disclosure review of data prior to sharing (removing county FIPS code with >1,000 records, recode DoB to the 15th and adjust all other dates accordingly). [Control 4b] **B. The researcher/user:** 1. Obtains authorization for data linkage and sharing for NSDUH and obtains authorization for linking AFCARS [Authorization gaps 2a, 4a] - ***Assumption*** 2. Uses the NSDUH and AFCARS data within NCHS RDC [Limitation 2a, Control 2g] 3. Uses the linked NSDUH and AFCARS data only for health statistical reporting and analysis [Limitations 2b, 4a] 5. Submits RDC application [Control 2a] 6. Obtains approvals from SAMHSA staff and NDACAN staff on the proposed linkage [Controls 2b, 4d] 7. Signs the NSDUH Designated Agent Form, NSDUH Data Access Agreement (DAA); and if the researcher/user is a student, signs the SAMHSA RDC Student Data User Acknowledgement form along with their advisors; and signs NDACAN Terms of Use Agreement [Controls 2c, 2d, 2f, 4c] 8. Completes confidentiality training for NSDUH data access [Control 2e]	**Researchers/users:** 2a. Must use NSDUH data within the NCHS RDC [NSDUH] 2b. Must use NSDUH data only for health statistical reporting and analysis [NSDUH] 4a. Must use AFCARS data in accordance with their approved research described in Section I.1 of the NDACAN Terms of Use Agreement [AFCARS]	**For sharing AFCARS, NDACAN staff must:** 4a. Share AFCARS data de-identified of all 18 HIPAA identifiers. [AFCARS] 4b. Perform disclosure review of data prior to sharing (removing county FIPS code with >1,000 records, recode DoB to the 15th and adjust all other dates accordingly). [AFCARS] **For accessing NSDUH, researchers/users must:** 2a. Submit RDC application [NSDUH] 2b. Obtain approval from SAMHSA staff on the research proposed in the RDC application [NSDUH] 2c. Sign Designated Agent Form (DAF) [NSDUH] 2d. Sign Data Access Agreement (DAA) [NSDUH] 2e. Complete confidentiality training [NSDUH] 2f. Sign the SAMHSA RDC Student Data User Acknowledgement form and obtain advisor's signature, if researcher/user is a student [NSDUH] 2g. Access data within NCHS RDC [NSDUH] **For accessing AFCARS, researchers/users must:** 4c. Execute of the NDACAN Terms of Use Agreement. [AFCARS] 4d. Obtain review and approval from NDACAN staff on the proposed research. [AFCARS]	2a. For NSDUH, information on authorizations for linkage and sharing is not available/found. 4a. For AFCARS, information on authorizations for linkage is not available/found.

TABLE 2: USE CASE 1 - DATASET LINKAGE

Definitions: Limitations are restrictions on data linkage and use (e.g., dataset must only be linked with other disease-relevant data, dataset must be used in a physical enclave, etc.). **Controls** are processes established to ensure compliance with governance for data sharing, access, and use (e.g., user must access data in a physical enclave, user must sign data use agreement, user must receive data access committee approval, etc.). **Authorization gaps** exist when there is no explicit authorization or the authorization is not available or found in the information collected by the project team for various data life cycle stages (data collection, linking, sharing, access and use).

		Can the datasets be linked?	What limitations do the linked datasets inherit?	What controls do the linked dataset require?	What authorization gaps exist?
6	**MTF (Dataset 3) and AFCARS (Dataset 4) linkage**	**Yes, MTF and AFCARS can be linked provided:** **A. ICPSR/NAHDAP staff must:** 1. Shares fully de-identified data (for MTF restricted-use data, this does not include state and zipcode) [Control 3a] 2. Performs disclosure review prior to sharing [Control 3b] 3. Performs disclosure review of analysis outputs prior to removing output data from the VDE [Control 3c] **B. NDACAN staff:** 4. Shares AFCARS data de-identified of all 18 HIPAA identifiers. [Control 4a] 5. Performs disclosure review of data prior to sharing (removing county FIPS code with >1,000 records, recode DoB to the 15th and adjust all other dates accordingly). [Control 4b] **C. The researcher/user:** 1. Obtains authorization for linking MTF and AFCARS data [Authorization gaps 3a, 4a] - **Assumption** 2. Uses the linked data for broad research or statistical purposes [Limitation 3a] 3. Obtain approvals from ICPSR/NAHDAP and NDACAN staff on the proposed linkages [Limitation 4a, Controls 3g, 4d] 4. Uses/accesses the MTF and AFCARS data in the ICPSR VDE [Control 3e] 5. Signs NAHDAP VDE RDUA and NDACAN Terms of Use Agreement [Controls 3d and 4c] 6. Obtains IRB approval or exemption from their institution for accessing MTF [Control 3f]	**Researchers/users:** 3a. Must use MTF data for broad research or statistical purposes [MTF] 4a. Must use AFCARS data in accordance with their approved research described in Section I.1 of the NDACAN Terms of Use Agreement [AFCARS]	**For sharing MTF, ICPSR/NAHDAP staff must:** 3a. Share fully de-identified data (for MTF restricted-use data, this does not include state and zipcode) [MTF] 3b. Perform disclosure review prior to sharing [MTF] 3c. Perform disclosure review of analysis outputs prior to removing output data from the VDE [MTF] **For sharing AFCARS, NDACAN staff must:** 4a. Share AFCARS data de-identified of all 18 HIPAA identifiers. [AFCARS] 4b. Perform disclosure review of data prior to sharing (removing county FIPS code with >1,000 records, recode DoB to the 15th and adjust all other dates accordingly). [AFCARS] **For accessing MTF, researchers/users must:** 3d. Execute NAHDAP VDE RDUA between ICPSR (U-Mich) and the researcher's institution [MTF] 3e. Access data only through the ICPSR VDE (virtual enclave) [MTF] 3f. Obtain IRB approval or exemption from the researcher's institution [MTF] 3g. Obtain review and approval from NAHDAP on the proposed research [MTF] **For accessing AFCARS, researchers/users must:** 4c. Execute of the NDACAN Terms of Use Agreement. [AFCARS] 4d. Obtain review and approval from NDACAN staff on the proposed research. [AFCARS]	3a. For MTF, information on authorizations for linkage is not available/found. 4a. For AFCARS, information on authorizations for linkage is not available/found.

TABLE 2: USE CASE 1 - DATASET LINKAGE

Definitions: Limitations are restrictions on data linkage and use (e.g., dataset must only be linked with other disease-relevant data, dataset must be used in a physical enclave, etc.). **Controls** are processes established to ensure compliance with governance for data sharing, access, and use (e.g., user must access data in a physical enclave, user must sign data use agreement, user must receive data access committee approval, etc.). **Authorization gaps e**xist when there is no explicit authorization or the authorization is not available or found in the information collected by the project team for various data life cycle stages (data collection, linking, sharing, access and use).

		Can the datasets be linked?	What limitations do the linked datasets inherit?	What controls do the linked dataset require?	What authorization gaps exist?
7	**All datasets**	**Yes, NHANES, NSDUH, MTF, and AFCARS can be linked provided:** **NCHS RDC/ICPSR-NAHDAP/NDACAN staff:** 1. Shares NHANES, MTF and AFCARS data de-identified of all direct identifiers - for NHANES, certain indirect identifiers (geography) may be included - for MTF, state and zipcode may be included [Controls 1a, 3a, 4a] 2. Performs disclosure review prior to sharing of (a) NHANES data by NCHS Disclosure Review Board/NCHS Confidentiality Officer, (b) MTF data, and (c) AFCARS data (removing county FIPS code with >1,000 records, recode DoB to the 15th and adjust all other dates accordingly) [Controls 1b, 3b, 4b] 3. Reaches agreement with NHANES and MTF data sources on an approach to perform disclosure review that meets each data sources' requirements and appropriately addresses any increased re-identifiability risk introduced by linking these datasets [Controls 1c, 3c] **The researcher/user:** 1. Obtains authorizations for linkage and sharing for NSDUH and obtains authorization for linking	**Researchers/users:** 1a. Can only link NHANES data to vital statistics, health, nutrition, and other related records [NHANES] 1b. Must use NHANES data within the NCHS RDC (on-site enclave) [NHANES] 1c. Must use NHANES data only for health statistical reporting and analysis [NHANES] 2a. Must use NSDUH data within the NCHS RDC [NSDUH] 2b. Must use NSDUH data only for health statistical reporting and analysis [NSDUH] 3a. Must use MTF data for broad research or statistical purposes [MTF] 4a. Must use AFCARS data in accordance with their approved research described in Section I.1 of the NDACAN Terms of Use Agreement [AFCARS]	**For sharing NHANES, NCHS RDC staff must:** 1a. Share NHANES data de-identified of all direct identifiers; certain indirect identifiers (geography) may be included [NHANES] 1b. Perform disclosure review prior to sharing the NHANES restricted-use data through the RDC (NCHS Disclosure Review Board/NCHS Confidentiality Officer) [NHANES] 1c. Perform disclosure review of the output before releasing it (RDC and DHANES) [NHANES] **For sharing MTF, ICPSR/NAHDAP staff must:** 3a. Share fully de-identified data (for MTF restricted-use data, this does not include state and zipcode) [MTF] 3b. Perform disclosure review prior to sharing [MTF] 3c. Perform disclosure review of analysis outputs prior to removing output data from the VDE [MTF] **For sharing AFCARS, NDACAN staff must:** 4a. Share AFCARS data de-identified of all 18 HIPAA identifiers. [AFCARS] 4b. Perform disclosure review of data prior to sharing (removing county FIPS code with >1,000 records, recode DoB to the 15th and adjust all other dates accordingly). [AFCARS]	2a. For NSDUH, information on authorizations for linkage and sharing is not available/found. 3a. For MTF, information on authorizations for linkage is not available/found. 4a. For AFCARS, information on authorizations for linkage is not available/found.

TABLE 2: USE CASE 1 - DATASET LINKAGE

Definitions: Limitations are restrictions on data linkage and use (e.g., dataset must only be linked with other disease-relevant data, dataset must be used in a physical enclave, etc.). **Controls** are processes established to ensure compliance with governance for data sharing, access, and use (e.g., user must access data in a physical enclave, user must sign data use agreement, user must receive data access committee approval, etc.). **Authorization gaps** exist when there is no explicit authorization or the authorization is not available or found in the information collected by the project team for various data life cycle stages (data collection, linking, sharing, access and use).

		Can the datasets be linked?	What limitations do the linked datasets inherit?	What controls do the linked dataset require?	What authorization gaps exist?
		for MTF and AFCARS [Authorization gaps 2a, 3a, 4a] - Assumption 2. Ensures that NSDUH, MTF and AFCARS data used for linking to NHANES data must be either vital statistics, health, nutrition, or other related records [Limitation 1a] 3. Uses NHANES, NSDUH, MTF (after obtaining permission from ICPSR/NAHDAP staff) and AFCARS data within NCHS RDC [Limitations 1b, 2a, Controls 1h, 2g, 3e] - Assumption for MTF 4. Uses linked NHANES, NSDUH, MTF, and AFCARS data for statistical purposes only [Limitation 1c, 2b, 3a] 5. Submit RDC application for accessing NSDUH [Control 2a] 6. Obtains approval from NCHS Confidentiality Officer, NCHS RDC, and DHANES/NHANES, SAMHSA staff, NAHDAP staff, and NDACAN staff on the proposed linkage [Controls 1d, 2b, 3g, 4d] 7. Signs and completes the NHANES Data Use/Access Agreement, NHANES Designated Agent Agreement, NSDUH Designated Agent Form, NSDUH Data Access Agreement (DAA) , NAHDAP VDE RDUA, NDACAN Terms of Use Agreement [Controls 1e, 1f, 2c, 2d, 2f, 3d, 4c] 8. Completes confidentiality training for NHANES and NSDUH data access [Controls 1g, 2e]		**For accessing NHANES, researchers/users must:** 1d. Obtain approvals from NCHS Confidentiality Officer, NCHS RDC, and DHANES/NHANES on the proposed research [NHANES] 1e. Execute Data Use/Access Agreement (Rules of Behavior) [NHANES] 1f. Sign Designated Agent Agreement [NHANES] 1g. Complete confidentiality training [NHANES] 1h. Access data within NCHS RDC (on-site enclave) [NHANES] **For accessing NSDUH, researchers/users must:** 2a. Submit RDC application [NSDUH] 2b. Obtain approval from SAMHSA staff on the research proposed in the RDC application [NSDUH] 2c. Sign Designated Agent Form (DAF) [NSDUH] 2d. Sign Data Access Agreement (DAA) [NSDUH] 2e. Complete confidentiality training [NSDUH] 2f. Sign the SAMHSA RDC Student Data User Acknowledgement form and obtain advisor's signature, if researcher/user is a student [NSDUH] 2g. Access data within the NCHS RDC [NSDUH] **For accessing MTF, researchers/users must:** 3d. Execute NAHDAP VDE RDUA between ICPSR (U-Mich) and the researcher's institution [MTF]	

TABLE 2: USE CASE 1 - DATASET LINKAGE

Definitions: Limitations are restrictions on data linkage and use (e.g., dataset must only be linked with other disease-relevant data, dataset must be used in a physical enclave, etc.). **Controls** are processes established to ensure compliance with governance for data sharing, access, and use (e.g., user must access data in a physical enclave, user must sign data use agreement, user must receive data access committee approval, etc.). **Authorization gaps** exist when there is no explicit authorization or the authorization is not available or found in the information collected by the project team for various data life cycle stages (data collection, linking, sharing, access and use).

		Can the datasets be linked?	What limitations do the linked datasets inherit?	What controls do the linked dataset require?	What authorization gaps exist?
		9. Obtains IRB approval or exemption from their institution for accessing MTF [Control 3f]		3e. Access data only through the ICPSR VDE (virtual enclave) [MTF] 3f. Obtain IRB approval or exemption from the researcher's institution [MTF] 3g. Obtain review and approval from NAHDAP on the proposed research [MTF] **For accessing AFCARS, researchers/users must:** 4c. Execute of the NDACAN Terms of Use Agreement. [AFCARS] 4d. Obtain review and approval from NDACAN staff on the proposed research. [AFCARS]	

USE CASE 1 - GOVERNANCE INFORMATION				
Use Case 1: Effects of the COVID-19 pandemic on mental health of children. Are COVID-19 pandemic related mental health outcomes more severe for children in foster care?				
Dataset 1 - NHANES 'Mental Health - Depression Screener – Youth' dataset (2017-2020, limited)				
	Dataset Source	National Health and Nutrition Examination Survey (NHANES)		
	Dataset Source Agency	CDC		
	Dataset Type (Clinical, EHR, Survey, SDOH, etc.)	Survey/Clinical		
	Information Sources	Website; NCHS, DHANES, and RDC staff; NHANES linkage info document		
Dataset 1 - NHANES				
		Raw Language	**Interpretation**	**Source**
1	**Data Collection**			
1.1	**Authorizations and Applicable Regulations/Policies**			
1.1.1	**Authorizations**		1. Assent from children 2. Consent from adults 3. NCHS ERB (IRB) Approval 4. Section 306 of the Public Health Service Act (42 U.S.C. 242k) 5. National Nutrition Monitoring and Related Research Act of 1990 (P.L. 101-445), (October 22, 1990) 6. Food Quality Protection Act of 1996 (P.L. 104-170) 7. Federal Food, Drug, and Cosmetic Act (21 USC 393), Chapter 9	
1.1.1.1	Assent	Example 'Child Assent Form': "Your parents say that you can take part in this special survey. You have just read about the survey in this book. The survey tells us about the health of people. We will ask you to have an exam at our vans that are here in your town. This exam is a little like going to the doctor. Other kids and their families will be at the center. You do not have to do this if you do not want to. You can also stop at any time and you do not have to do any tests that you do not want to. If you take part, you will learn some things about yourself. You will help us learn a lot about other kids in the United States." [1]	Assent from children authorizes data collection	[1] https://wwwn.cdc.gov/nchs/data/nhanes/2019-2020/documents/2016-Child-Assent-7-11-Form-508.pdf (Accessed: 4/17/23)
1.1.1.2	Consent	Template for Home Interview Consent: You have been chosen to take part in the National Health and Nutrition Examination Survey (NHANES), conducted by the National Center for Health Statistics, part of the Centers for Disease Control and Prevention (CDC). This research tells us about the health and nutrition of people in the United States. It combines an interview with a health exam. You may take part in this survey or not. The choice is yours. You will not lose any benefits if you say no. If you choose to take part, you don't have to answer every question and you can stop the interview at any time. [1]	Consent from adults authorizes data collection	[1] https://wwwn.cdc.gov/nchs/data/nhanes/2019-2020/documents/2019-Home-Interview-Consent-English-508.pdf (Accessed: 4/17/23)

Dataset 1 - NHANES				
		Raw Language	**Interpretation**	**Source**
1.1.1.3	IRB/equivalent Privacy Board determination	NCHS Ethics Review Board (ERB) Approvals of NCHS Protocols: NHANES 2021-2022 Protocol #2021-05 NHANES 2019-2020 Protocol #2018-01 NHANES 2017-2018 Protocol #2018-01 (Effective beginning October 26, 2017) Continuation of Protocol #2011-17 (Effective through October 26, 2017). [1] PoCs stated that NCHS ERB approval ensures data collection is consistent with assent/consent [2]	NCHS ERB (IRB) approval authorizes data collection	[1] https://www.cdc.gov/nchs/nhanes/irba98.htm (Accessed: 4/17/23) [2] NCHS and DHANES Meeting
1.1.1.4	Local/state/federal law	Four public laws authorize or necessitate the collection of information about the health of the American people. These are: - Section 306 of the Public Health Service Act (42 U.S.C. 242k), which directs NCHS to collect statistics on subjects, such as the extent and nature of illness and disability of the population; environmental, social and other health hazards; determinants of health; health resources; and utilization of health care; - The National Nutrition Monitoring and Related Research Act of 1990 (P.L. 101-445), (October 22, 1990), which specifies that NHANES be maintained as a component of the comprehensive nutrition monitoring plan with continuous coverage of dietary and nutritional status for the population and high-risk subgroups; - The Food Quality Protection Act of 1996 (P.L. 104-170), which requires the implementation of surveys to collect data on food consumption patterns of infants and children and data on dietary exposure to pesticides among infants and children; and - The Federal Food, Drug, and Cosmetic Act (21 USC 393), Chapter 9, which authorizes the collection of information to support the Food and Drug Administration's objective to obtain current, timely, and policy-relevant consumer information to carry out its statutory functions. [1]	Four federal laws authorizes data collection: 1. Section 306 of the Public Health Service Act (42 U.S.C. 242k) 2. National Nutrition Monitoring and Related Research Act of 1990 (P.L. 101-445), (October 22, 1990) 3. Food Quality Protection Act of 1996 (P.L. 104-170) 4. Federal Food, Drug, and Cosmetic Act (21 USC 393), Chapter 9	[1] NHANES Linkage Info doc from NCHS staff
1.1.1.5	Institutional Certification	Information not available/found	Information not available/found	
1.1.1.6	Data originator agreement	Information not available/found	Information not available/found	
1.1.1.7	Repository agreements/policies	Information not available/found	Information not available/found	
1.1.1.8	Other (specify)	Information not available/found	Information not available/found	
1.1.2	**Applicable Regulations/Policies**			
1.1.2.1	Local regulations/policies	N/A - PoCs are not aware of any specific policies that apply. [1]	N/A	[1] NCHS and DHANES Meeting
1.1.2.2	Tribal regulations/policies	N/A - Tribal or local areas are not specifically targeted, so no specific agreements. [1]	N/A	[1] NCHS and DHANES Meeting
1.1.2.3	State regulations/policies	N/A - PoCs are not aware of any specific policies that apply. [1]	N/A	[1] NCHS and DHANES Meeting

Dataset 1 - NHANES				
		Raw Language	**Interpretation**	**Source**
1.1.2.4	Federal regulations/policies	Four public laws authorize or necessitate the collection of information about the health of the American people. These are: - Section 306 of the Public Health Service Act (42 U.S.C. 242k), which directs NCHS to collect statistics on subjects, such as the extent and nature of illness and disability of the population; environmental, social and other health hazards; determinants of health; health resources; and utilization of health care; - The National Nutrition Monitoring and Related Research Act of 1990 (P.L. 101-445), (October 22, 1990), which specifies that NHANES be maintained as a component of the comprehensive nutrition monitoring plan with continuous coverage of dietary and nutritional status for the population and high-risk subgroups; - The Food Quality Protection Act of 1996 (P.L. 104-170), which requires the implementation of surveys to collect data on food consumption patterns of infants and children and data on dietary exposure to pesticides among infants and children; and - The Federal Food, Drug, and Cosmetic Act (21 USC 393), Chapter 9, which authorizes the collection of information to support the Food and Drug Administration's objective to obtain current, timely, and policy-relevant consumer information to carry out its statutory functions. [1]	1. Section 306 of the Public Health Service Act (42 U.S.C. 242k) 2. National Nutrition Monitoring and Related Research Act of 1990 (P.L. 101-445), (October 22, 1990) 3. Food Quality Protection Act of 1996 (P.L. 104-170) 4. Federal Food, Drug, and Cosmetic Act (21 USC 393), Chapter 9	[1] NHANES Linkage Info doc from NCHS staff
1.1.2.5	International regulations/policies	N/A - The National Health and Nutrition Examination Survey (NHANES) is a program of studies designed to assess the health and nutritional status of adults and children in the United States. The sample for the survey is selected to represent the U.S. population of all ages. [1]	N/A	[1] https://www.cdc.gov/nchs/nhanes/about_nhanes.htm (Accessed: 4/17/23)
1.1.2.6	Contractual obligations	N/A - PoCs are not aware of any specific policies that apply. [1]	N/A	[1] NCHS and DHANES Meeting
1.1.2.7	Repository policies		N/A	
1.2	**Governance for data linkage, sharing, access, and use based on data collection authorization or applicable regulations/policies (i.e., the origin of the governance)**			
1.2.1	Whether the data can be linked	Child proxies 0-15 years – The parent/guardian of the child consents verbally as a proxy respondent for the child. Ages 16 and 17 years – Multiple consents/assents are required for this age group since children 16 – 17 years of age respond to their own interview. First, the parent/guardian verbally provides permission to the audio recording, interview administration, and linkage. If the parent/guardian grants permission (consent), then the child is asked to assent to the recording, interview, and linkage. [1]	Assent/Consent specifies that the data can be linked	[1] NHANES Linkage Info doc from NCHS staff
1.2.2	With what other data can it be linked or can it not be linked (scope of linkage)	Consent template for 'Home Interview Consent': Health research using NHANES can be enhanced by combining your survey records with other data sources. The data gathered are used to link your answers to vital statistics, health, nutrition, and other related records. "We can do additional health research by linking the interview and exam data of everyone listed under "SP NAME" in the gray box below to vital statistics, health, nutrition, and other related records. May we try to link these survey records with other records? Yes, No, N/A." [1]	Assent/Consent specifies that the data can be linked to vital statistics, health, nutrition, and other related records	[1] https://wwwn.cdc.gov/nchs/data/nhanes/2019-2020/documents/2019-Home-Interview-Consent-English-508.pdf (Accessed: 4/17/23)
1.2.3	Whether data can be shared	Does not authorize/specify	Does not authorize/specify	
1.2.4	How data can be shared (de-identification status, disclosure review)	Does not authorize/specify	Does not authorize/specify	

Dataset 1 - NHANES				
		Raw Language	**Interpretation**	**Source**
1.2.5	How data can be accessed (access type, data use agreement, data access committee/group approval, IRB LOD, etc.)	Does not authorize/specify	Does not authorize/specify	
1.2.6	How data can be used (data use limitations)	Consent template for 'Home Interview Consent': "Data gathered in this survey are used to study many health issues. We are required by law to use your information for statistical research only and to keep it confidential." [1] While the consent states that data may only be used for statistical purposes, PoC stated that this is a broad definition and the data may be used for general research purposes. [2]	Assent/Consent specifies that the data can be used for statistical reporting and analysis (includes broad research)	[1] https://wwwn.cdc.gov/nchs/data/nhanes/2019-2020/documents/2019-Home-Interview-Consent-English-508.pdf (Accessed: 4/17/23) [2] NCHS and DHANES Meeting
1.2.7	Other (specify)	Does not authorize/specify	Does not authorize/specify	
2	**Data Linkage**			
2.1	**Authorizations and Applicable Regulations/Policies**			
2.1.1	**Authorizations**		1. Assent from children 2. Consent from adults 3. NCHS ERB (IRB) approval 4. DHANES/NHANES approval	
2.1.1.1	Assent	Child proxies 0-15 years – The parent/guardian of the child consents verbally as a proxy respondent for the child. Ages 16 and 17 years – Multiple consents/assents are required for this age group since children 16 – 17 years of age respond to their own interview. First, the parent/guardian verbally provides permission to the audio recording, interview administration, and linkage. If the parent/guardian grants permission (consent), then the child is asked to assent to the recording, interview, and linkage. [1]	Assent from children authorizes data linkage	[1] NHANES Linkage Info doc from NCHS staff
2.1.1.2	Consent	Consent template for 'Home Interview Consent': Health research using NHANES can be enhanced by combining your survey records with other data sources. The data gathered are used to link your answers to vital statistics, health, nutrition, and other related records. "We can do additional health research by linking the interview and exam data of everyone listed under "SP NAME" in the gray box below to vital statistics, health, nutrition, and other related records. May we try to link these survey records with other records? Yes, No, N/A." [1]	Consent from adults authorizes data linkage	[1] https://wwwn.cdc.gov/nchs/data/nhanes/2019-2020/documents/2019-Home-Interview-Consent-English-508.pdf (Accessed: 4/17/23)
2.1.1.3	IRB/equivalent Privacy Board determination	Individual level linkage (1:1 linkage) requires IRB approval. [1]	NCHS ERB approval authorizes data linkage	[1] NCHS and DHANES Meeting
2.1.1.4	Local/state/federal laws	Information not available/found	Information not available/found	
2.1.1.5	Institutional Certification	Information not available/found	Information not available/found	
2.1.1.6	Data originator agreement	Information not available/found	Information not available/found	
2.1.1.7	Repository agreements/policies	Information not available/found	Information not available/found	
2.1.1.8	Other (specify)	RDC is not the owner of NHANES data. When RDC receives an application to link data with an external dataset, it is sent to the data owner (DHANES/NHANES) for their review and approval. [1]	DHANES/NHANES approval authorizes data linkage	[1] NCHS RDC Meeting
2.1.2	**Applicable Regulations/Policies**			
2.1.2.1	Local regulations/policies	Information not available/found	Information not available/found	
2.1.2.2	Tribal regulations/policies	N/A - Tribal or local areas are not specifically targeted, so no specific agreements. [1]	N/A	[1] NCHS and DHANES Meeting

Dataset 1 - NHANES				
		Raw Language	**Interpretation**	**Source**
2.1.2.3	State regulations/policies	Information not available/found	Information not available/found	
2.1.2.4	Federal regulations/policies	Information not available/found	Information not available/found	
2.1.2.5	International regulations/policies	N/A - The National Health and Nutrition Examination Survey (NHANES) is a program of studies designed to assess the health and nutritional status of adults and children in the United States. The sample for the survey is selected to represent the U.S. population of all ages. [1]	N/A	[1] https://www.cdc.gov/nchs/nhanes/about_nhanes.htm (Accessed: 4/17/23)
2.1.2.6	Contractual obligations	Information not available/found	Information not available/found	
2.1.2.7	Repository policies	Information not available/found	Information not available/found	
2.2	**Governance for data linkage, sharing, access, and use based on data linkage authorization or applicable regulations/ policies (i.e., the origin of the governance)**			
2.2.1	Whether the data can be linked	Child proxies 0-15 years – The parent/guardian of the child consents verbally as a proxy respondent for the child. Ages 16 and 17 years – Multiple consents/assents are required for this age group since children 16 – 17 years of age respond to their own interview. First, the parent/guardian verbally provides permission to the audio recording, interview administration, and linkage. If the parent/guardian grants permission (consent), then the child is asked to assent to the recording, interview, and linkage. [1] Individual level linkage (1:1 linkage) requires IRB approval. [2] RDC is not the owner of NHANES data. When RDC receives an application to link data with an external dataset, it is sent to the data owner (DHANES/NHANES) for their review and approval. [3]	Assent/Consent, NCHS ERB approval, and DHANES/HANES approval specify that the data can be linked	[1] NHANES Linkage Info doc from NCHS staff [2] NCHS and DHANES Meeting [3] NCHS RDC Meeting
2.2.2	With what other data can it be linked or can it not be linked (scope of linkage)	Consent template for 'Home Interview Consent': Health research using NHANES can be enhanced by combining your survey records with other data sources. The data gathered are used to link your answers to vital statistics, health, nutrition, and other related records. "We can do additional health research by linking the interview and exam data of everyone listed under "SP NAME" in the gray box below to vital statistics, health, nutrition, and other related records. May we try to link these survey records with other records? Yes, No, N/A." [1]	Assent/Consent specifies that the data can be linked to vital statistics, health, nutrition, and other related records	[1] https://wwwn.cdc.gov/nchs/data/nhanes/2019-2020/documents/2019-Home-Interview-Consent-English-508.pdf (Accessed: 4/17/23)
2.2.3	Whether data can be shared	Does not authorize/specify	Does not authorize/specify	
2.2.4	How data can be shared (de-identification status, disclosure review)	Does not authorize/specify	Does not authorize/specify	
2.2.5	How data can be accessed (access type, data use agreement, data access committee/group approval, IRB LOD, etc.)	Does not authorize/specify	Does not authorize/specify	
2.2.6	How data can be used (data use limitations)	Does not authorize/specify	Does not authorize/specify	
2.2.7	Other (specify)	Does not authorize/specify	Does not authorize/specify	
3	**Data Sharing**			
3.1	**Authorizations and Applicable Regulations/Policies**			

Dataset 1 - NHANES				
		Raw Language	**Interpretation**	**Source**
3.1.1	**Authorizations**		1. NCHS Disclosure Review Board (DRB) 2. Confidential Information Protection and Statistical Efficiency Act (CIPSEA) 3. NCHS non-disclosure affidavit (for NCHS/NHANES collaborators) 4. Data Sharing Agreement (for NCHS/NHANES collaborators)	
3.1.1.1	Assent	Does not authorize/specify	Does not authorize/specify	
3.1.1.2	Consent	Does not authorize/specify	Does not authorize/specify	
3.1.1.3	IRB/equivalent Privacy Board determination	IRB/equivalent Privacy Board determination: NCHS has a Disclosure Review Board (DRB) which helps determine NCHS data sharing policy and public data sharing approvals. [1]	NCHS Disclosure Review Board (DRB) authorizes data sharing	[1] NCHS RDC Email Communication
3.1.1.4	Local/state/federal laws	RDC is the data dissemination arm for NCHS restricted use data. NHANES collects their data under CIPSEA which authorizes NCHS to share the data with vetted/approved researchers (who then become agents of CIPSEA to use the restricted use data). Data collected under CIPSEA has a statement/promise of confidentiality to participants that the RDC will not release any data that could be used to reidentify the person or entity that provided the data. [1]	Confidential Information Protection and Statistical Efficiency Act (CIPSEA) authorizes data sharing through the NCHS RDC	[1] NCHS RDC Meeting
3.1.1.5	Institutional Certification	Information not available/found	Information not available/found	
3.1.1.6	Data originator agreement	Additionally, all NCHS/NHANES collaborators (i.e., FDA, NIH, USDA, etc.) are given the minimal amount of data pre-public release required to assess quality control for their contributed data collections. As a condition for sharing confidential data, the collaborator must also sign the NCHS non-disclosure affidavit and data are shared through a Data Sharing Agreement. [1]	Two agreements authorize data sharing from NCHS/NHANES collaborators that are authorized to review pre-release data that they contributed: 1. NCHS non-disclosure affidavit 2. Data Sharing Agreement	[1] NHANES Linkage Info doc from NCHS staff
3.1.1.7	Repository agreements/policies	Information not available/found	Information not available/found	
3.1.1.8	Other	Information not available/found	Information not available/found	
3.1.2	**Applicable Regulations/Policies**			
3.1.2.1	Local regulations/policies	Several NHANES datasets are not released to the public. However, secure, controlled access is granted through the NCHS Research Data Center (RDC) to guarantee confidentiality of the survey participants. A comprehensive list of the NHANES data files available through the RDC is available at, http://wwwn.cdc.gov/Nchs/Nhanes/Search/DataPage.aspx?Component=NonPublic. The NHANES data files released only through the RDC fall into two general categories: 1) the data file could possibly disclose participation in the survey (e.g., due to small sample sizes or rare combination of characteristics) or 2) the data file contains sensitive information (e.g., sexual behavior) and have been deemed to be NHANES Protected Data. A list of NHANES Protected Data Files from the 1999-2020 data collection cycles is included in the Appendix. [1]	NHANES Protected Data Policy	[1] https://www.cdc.gov/nchs/data/nhanes/NHANES-Protected-Data-Policy-Aug-2022.pdf (Accessed: 4/17/23)
3.1.2.2	Tribal regulations/policies	N/A - Tribal or local areas are not specifically targeted, so no specific agreements. [1]	N/A	[1] NCHS and DHANES Meeting
3.1.2.3	State regulations/policies	Information not available/found	Information not available/found	

Dataset 1 - NHANES				
		Raw Language	**Interpretation**	**Source**
3.1.2.4	Federal regulations/policies	RDC is the data dissemination arm for NCHS restricted use data. NHANES collects their data under CIPSEA which authorizes NCHS to share the data with vetted/approved researchers (who then become agents of CIPSEA to use the restricted use data). Data collected under CIPSEA has a statement/promise of confidentiality to participants that the RDC will not release any data that could be used to reidentify the person or entity that provided the data. [1] There are two laws that govern the NCHS RDC: Section 308(d) of the Public Health Service Act and the Confidential Information Protection and Statistical Efficiency Act (CIPSEA). The Public Health Service Act protects confidentiality and states that the only people who can access confidential data are NCHS staff and Designated Agents. Therefore, researchers wishing to access confidential data must become Designated Agents. CIPSEA stipulates that the penalty for willfully violating confidentiality is a class E felony with up to 5 years in prison or a $250,000 fine or both. The Freedom of Information Act does not apply to data collected under CIPSEA. [2]	1. Section 308(d) Public Health Act 2. Confidential Information Protection and Statistical Efficiency Act (CIPSEA)	[1] NCHS RDC Meeting [2] https://www.cdc.gov/rdc/data/b4/Disclosure-Manual-v2.5.pdf (Accessed: 4/17/23)
3.1.2.5	International regulations/policies	N/A - The National Health and Nutrition Examination Survey (NHANES) is a program of studies designed to assess the health and nutritional status of adults and children in the United States. The sample for the survey is selected to represent the U.S. population of all ages. [1]	N/A	[1] https://www.cdc.gov/nchs/nhanes/about_nhanes.htm (Accessed: 4/17/23)
3.1.2.6	Contractual obligations	N/A	N/A	
3.1.2.7	Repository policies	Information not available/found	Information not available/found	
3.2	**Governance for data linkage, sharing, access, and use based on data sharing authorization or applicable regulations/policies (i.e., the origin of the governance)**			
3.2.1	Whether the data can be linked	Does not authorize/specify	Does not authorize/specify	
3.2.2	With what other data can it be linked or can it not be linked (scope of linkage)	Does not authorize/specify	Does not authorize/specify	
3.2.3	Whether data can be shared	RDC is the data dissemination arm for NCHS restricted use data. NHANES collects their data under CIPSEA which authorizes NCHS to share the data with vetted/approved researchers (who then become agents of CIPSEA to use the restricted use data). Data collected under CIPSEA has a statement/promise of confidentiality to participants that the RDC will not release any data that could be used to reidentify the person or entity that provided the data. [1]	Confidential Information Protection and Statistical Efficiency Act (CIPSEA) specifies that data can be shared through NCHS RDC (authorized by CIPSEA to share the data)	[1] NCHS RDC Meeting

Dataset 1 - NHANES				
		Raw Language	**Interpretation**	**Source**
3.2.4	How data can be shared (de-identification status, disclosure review)	Direct identifiers (name, social security number, address) cannot be accessed through the RDC. Indirect Identifiers (geography) may be available through the RDC. • All geography below the national level is restricted for continuous NHANES, prior to that all geography below the regional level is restricted. • The exact date of interview and exam are restricted for all years. NHANES does not provide a variable name for exact date of interview and exam in the limited access documentation. [1] Deductive disclosure performed by the NCHS Disclosure Review Board/NCHS Confidentiality Officer before the data are released. [2] Researchers that wish to access data in the RDC must submit an application where they specify what data they need, what their research questions are, and their desired output. This application is reviewed and approved by the NCHS confidentiality officer (who chairs the NCHS DRB). If application is approved, the RDC and data owners (DHANES/NHANES) will then jointly review the output to ensure it conforms to their approved application and cannot reidentify an individual (as per CIPSEA) before that output is allowed to leave the RDC. [3]	NCHS RDC specifies that data shared through RDC: 1. Must be de-identified of all direct identifiers and certain indirect identifiers (geography) may be included 2. Must undergo disclosure review by NCHS Disclosure Review Board/NCHS Confidentiality Officer prior to sharing the restricted-use data through the RDC, and then the RDC and DHANES reviews the output before releasing that output	[1] "National Health and Nutrition Examination Survey (NHANES) Restricted Variables" webpage: https://www.cdc.gov/rdc/b1datatype/Dt1222.htm (Accessed: 4/17/23) [2] NCHS and DHANES Meeting [3] NCHS RDC Meeting
3.2.5	How data can be accessed (access type, data use agreement, data access committee/group approval, IRB LOD, etc.)	Several NHANES datasets are not released to the public. However, secure, controlled access is granted through the NCHS Research Data Center (RDC) to guarantee confidentiality of the survey participants. A comprehensive list of the NHANES data files available through the RDC is available at, http://wwwn.cdc.gov/Nchs/Nhanes/Search/DataPage.aspx?Component=NonPublic. The NHANES data files released only through the RDC fall into two general categories: 1) the data file could possibly disclose participation in the survey (e.g., due to small sample sizes or rare combination of characteristics) or 2) the data file contains sensitive information (e.g., sexual behavior) and have been deemed to be NHANES Protected Data. A list of NHANES Protected Data Files from the 1999-2020 data collection cycles is included in the Appendix. [1]	NHANES Protected Data Policy specifies that the restricted use data must be accessed through the NCHS RDC (on-site enclave)	[1] https://www.cdc.gov/nchs/data/nhane/NHANES-Protected-Data-Policy-Aug-2022.pdf (Accessed: 4/17/23)
3.2.6	How data can be used (data use limitations)	Section 308(d) of the Public Health Service Act (42 U.S.C. 242m(d)) includes NCHS's authority to collect information and prohibits anyone from using any personal information for any purpose other than what was described to survey participants. [1] In general, NCHS data can only be used for statistical purposes and research purposes as per CIPSEA and used for the purpose it was collected for as per Section 308(d) of the Public Health Service Act. [2]	Section 308(d) of the Public Health Service Act specifies that NCHS data containing personal information cannot be used for any purpose other than what was described to survey participants. CIPSEA specifies that NCHS data can only be used for statistical and research purposes.	[1] https://www.cdc.gov/rdc/Data/b4/section308.pdf (Accessed: 4/17/23) [2] NCHS RDC Email Communication
3.2.7	Other (specify)	Does not authorize/specify	Does not authorize/specify	
4	**Data Access**			
4.1	**Authorizations and Applicable Regulations/Policies**			
4.1.1	**Authorizations**		1. Data Use/Access Agreement (Rules of Behavior)	
4.1.1.1	Assent	Does not authorize/specify	Does not authorize/specify	
4.1.1.2	Consent	Does not authorize/specify	Does not authorize/specify	
4.1.1.3	IRB/equivalent Privacy Board determination	N/A	N/A	

Dataset 1 - NHANES				
		Raw Language	**Interpretation**	**Source**
4.1.1.4	Local/state/federal laws	N/A	N/A	
4.1.1.5	Institutional Certification	Information not available/found	Information not available/found	
4.1.1.6	Data originator agreement	Information not available/found	Information not available/found	
4.1.1.7	Repository agreements/policies	N/A	N/A	N/A
4.1.1.8	Other (specify)	Researchers that wish to access data in the RDC must submit an application where they specify what data they need, what their research questions are, and their desired output. This application is reviewed and approved by the NCHS confidentiality officer (who chairs the NCHS Disclosure Review Board). If application is approved, the RDC and data owners (DHANES/NHANES) will then jointly review the output to ensure it conforms to their approved application and cannot reidentify an individual (as per CIPSEA) before that output is allowed to leave the RDC. Once application is approved, researchers must also undergo confidentially training, data use/access agreement (similar to a Rules of Behavior form), sign a non-disclosure CIPSEA agent form, which makes them agents under the statue and legally liable for proper use of the data and the RDC can prosecute researchers under CIPSEA if the data is misused. [1]	Data Use/Access Agreement (Rules of Behavior) authorizes data access	[1] NCHS RDC Meeting
4.1.2	**Applicable Regulations/Policies**			
4.1.2.1	Local regulations/policies	Several NHANES datasets are not released to the public. However, secure, controlled access is granted through the NCHS Research Data Center (RDC) to guarantee confidentiality of the survey participants. A comprehensive list of the NHANES data files available through the RDC is available at, http://wwwn.cdc.gov/Nchs/Nhanes/Search/DataPage.aspx?Component=Non Public. The NHANES data files released only through the RDC fall into two general categories: 1) the data file could possibly disclose participation in the survey (e.g., due to small sample sizes or rare combination of characteristics) or 2) the data file contains sensitive information (e.g., sexual behavior) and have been deemed to be NHANES Protected Data. A list of NHANES Protected Data Files from the 1999-2020 data collection cycles is included in the Appendix. [1]	1. NHANES Protected Data Policy 2. NCHS RDC policy	[1] https://www.cdc.gov/nchs/data/nhanes/NHANES-Protected-Data-Policy-Aug-2022.pdf (Accessed: 4/17/23)
4.1.2.2	Tribal regulations/policies	N/A - Tribal or local areas are not specifically targeted, so no specific agreements. [1]	N/A	[1] NCHS and DHANES Meeting
4.1.2.3	State regulations/policies	Information not available/found	Information not available/found	

Dataset 1 - NHANES				
		Raw Language	**Interpretation**	**Source**
4.1.2.4	Federal regulations/policies	There are two laws that govern the NCHS RDC: Section 308(d) of the Public Health Service Act and the Confidential Information Protection and Statistical Efficiency Act (CIPSEA). The Public Health Service Act protects confidentiality and states that the only people who can access confidential data are NCHS staff and Designated Agents. Therefore, researchers wishing to access confidential data must become Designated Agents. CIPSEA stipulates that the penalty for willfully violating confidentiality is a class E felony with up to 5 years in prison or a $250,000 fine or both. The Freedom of Information Act does not apply to data collected under CIPSEA. [1] Designated Agent Agreement – NCHS Research Data Center (RDC) I, (name) , do solemnly swear (or affirm) I will observe all policies and procedures that protect the confidential data I access from unauthorized disclosures. The data that I will access in the RDC is described in my RDC proposal. I will not disclose this confidential data, either while as an agent or after project conclusion, whether in data files, lists, or reports created using the confidential data, as specified under section 308 (d) of the Public Health Service Act and under penalties* set forth in §3572(f) of the Confidential Information Protection and Statistical Efficiency Act of 2018 (44 USC 3561 – 3583). [2]	1. Section 308(d) of the Public Health Service Act 2. Confidential Information Protection and Statistical Efficiency Act (CIPSEA)	[1] https://www.cdc.gov/rdc/data/b4/DisclosureManual-v2.5.pdf (Accessed: 4/17/23) [2] https://www.cdc.gov/rdc/data/b4/DesignatedAgent-321.pdf (Accessed: 4/17/23)
4.1.2.5	International regulations/policies	N/A - The National Health and Nutrition Examination Survey (NHANES) is a program of studies designed to assess the health and nutritional status of adults and children in the United States. The sample for the survey is selected to represent the U.S. population of all ages. [1]	N/A	[1] https://www.cdc.gov/nchs/nhanes/about_nhanes.htm (Accessed: 4/17/23)
4.1.2.6	Contractual obligations	Information not available/found	Information not available/found	
4.1.2.7	Repository policies	N/A - the RDC does not have a repository agreement with NHANES. The RDC has a data use agreement with NHANES. The RDC does not archive NHANES data nor is the RDC a repository for NHANES data.	N/A	[1] NCHS RDC Email Communication
4.2	**Governance for data linkage, sharing, access, and use based on data access authorization or applicable regulations/policies (i.e., the origin of the governance)**			
4.2.1	Whether the data can be linked	Does not authorize/specify	Does not authorize/specify	
4.2.2	With what other data can it be linked or can it not be linked (scope of linkage)	Does not authorize/specify	Does not authorize/specify	
4.2.3	Whether data can be shared	Does not authorize/specify	Does not authorize/specify	

Dataset 1 - NHANES				
		Raw Language	**Interpretation**	**Source**
4.2.4	How data can be shared (de-identification status, disclosure review)	Direct identifiers (name, social security number, address) cannot be accessed through the RDC. Indirect Identifiers (geography) may be available through the RDC. • All geography below the national level is restricted for continuous NHANES, prior to that all geography below the regional level is restricted. • The exact date of interview and exam are restricted for all years. NHANES does not provide a variable name for exact date of interview and exam in the limited access documentation. [1] Deductive disclosure performed by the NCHS Disclosure Review Board/NCHS Confidentiality Officer before the data are released. [2] Researchers that wish to access data in the RDC must submit an application where they specify what data they need, what their research questions are, and their desired output. This application is reviewed and approved by the NCHS confidentiality officer (who chairs the NCHS DRB). If application is approved, the RDC and data owners (DHANES/NHANES) will then jointly review the output to ensure it conforms to their approved application and cannot reidentify an individual (as per CIPSEA) before that output is allowed to leave the RDC. [3]	NCHS RDC specifies that data shared through RDC: 1. Must be de-identified of all direct identifiers and certain indirect identifiers (geography) may be included 2. Must undergo disclosure review by NCHS Disclosure Review Board/NCHS Confidentiality Officer prior to sharing the restricted-use data through the RDC, and then the RDC and DHANES reviews the output before releasing that output	[1] "National Health and Nutrition Examination Survey (NHANES) Restricted Variables" webpage: https://www.cdc.gov/rdc/b1datatype/Dt1222.htm (Accessed: 4/17/23) [2] NCHS and DHANES Meeting [3] NCHS RDC Meeting
4.2.5	How data can be accessed (access type, data use agreement, data access committee/group approval, IRB LOD, etc.)	Researchers that wish to access data in the RDC must submit an application where they specify what data they need, what their research questions are, and their desired output. This application is reviewed and approved by the NCHS confidentiality officer (who chairs the NCHS Disclosure Review Board). If application is approved, the RDC and data owners (DHANES/NHANES) will then jointly review the output to ensure it conforms to their approved application and cannot reidentify an individual (as per CIPSEA) before that output is allowed to leave the RDC. Once application is approved, researchers must also undergo confidentially training, data use/access agreement (similar to a Rules of Behavior form), sign a non-disclosure CIPSEA agent form, which makes them agents under the statue and legally liable for proper use of the data and the RDC can prosecute researchers under CIPSEA if the data is misused. [1]	NCHS RDC specifies that for data access, a user: 1. Must obtain review and approval by NHCS Confidentiality Officer, NCHS RDC, and DHANES/NHANES on the proposed research 2. Must execute Data Use/Access Agreement (Rules of Behavior) 3. Must sign Designated Agent Agreement (Non-Disclosure CIPSEA Agent Form) 4. Must undergo confidentiality training 5. Must access data through the NCHS RDC (on-site enclave)	[1] NCHS RDC Meeting
4.2.6	How data can be used (data use limitations)	Section 308(d) of the Public Health Service Act (42 U.S.C. 242m(d)) includes NCHS's authority to collect information and prohibits anyone from using any personal information for any purpose other than what was described to survey participants. [1] In general, NCHS data can only be used for statistical purposes and research purposes as per CIPSEA and used for the purpose it was collect for as per Section 308(d) of the Public Health Service Act. [2]	Section 308(d) of the Public Health Service Act specifies that NCHS data containing personal information cannot be used for any purpose other than what was described to survey participants. CIPSEA specifies that NCHS data can only be used for statistical and research purposes.	[1] https://www.cdc.gov/rdc/Data/b4/section308.pdf (Accessed: 4/17/23) [2] NCHS RDC Email Communication
4.2.7	Other (specify)	Information not available/found		
5	**Data Use**			
5.1	**Authorizations and Applicable Regulations/Policies**			
5.1.1	**Authorizations**		1. Consent from adults 2. NCHS ERB (IRB) approval 3. Designated Agent Agreement (Non-Disclosure CIPSEA Agent Form) 4. Data Use/Access Agreement (Rules of Behavior)	
5.1.1.1	Assent	Does not authorize/specify	Does not authorize/specify	

Dataset 1 - NHANES				
		Raw Language	**Interpretation**	**Source**
5.1.1.2	Consent	Consent template for 'Home Interview Consent': "Data gathered in this survey are used to study many health issues. We are required by law to use your information for statistical research only and to keep it confidential." [1] While the consent states that data may only be used for statistical purposes, PoC stated that this is a broad definition and the data may be used for general research purposes. [2]	Consent from adults authorizes data use	[1] https://wwwn.cdc.gov/nchs/data/nhanes/2019-2020/documents/2019-Home-Interview-Consent-English-508.pdf (Accessed: 4/17/23) [2] NCHS and DHANES Meeting
5.1.1.3	IRB/equivalent Privacy Board determination	ERB impacts/authorizes how the data can be used, for example, providing authorization for the addition of the linkage question to the NHANES consent forms. [1]	NCHS ERB (IRB) approval authorizes data use	[1] NCHS and DHANES Meeting
5.1.1.4	Local/state/federal laws	Researchers that wish to access data in the RDC must submit an application where they specify what data they need, what their research questions are, and their desired output. This application is reviewed and approved by the NCHS confidentiality officer (who chairs the NCHS Disclosure Review Board). If application is approved, the RDC and data owners (DHANES/NHANES) will then jointly review the output to ensure it conforms to their approved application and cannot reidentify an individual (as per CIPSEA) before that output is allowed to leave the RDC. Once application is approved, researchers must also undergo confidentially training, data use/access agreement (similar to a Rules of Behavior form), sign a non-disclosure CIPSEA agent form, which makes them agents under the statue and legally liable for proper use of the data and the RDC can prosecute researchers under CIPSEA if the data is misused. [1]	Designated Agent Agreement (Non-Disclosure CIPSEA Agent Form) authorizes data use	[1] NCHS RDC Meeting
5.1.1.5	Institutional Certification	Information not available/found	Information not available/found	
5.1.1.6	Data originator agreement	Information not available/found	Information not available/found	
5.1.1.7	Repository agreements/policies	N/A - the RDC does not have a repository agreement with NHANES. The RDC has a data use agreement with NHANES. The RDC does not archive NHANES data nor is the RDC a repository for NHANES data.	N/A	[1] NCHS RDC Email Communication
5.1.1.8	Other (specify)	Researchers that wish to access data in the RDC must submit an application where they specify what data they need, what their research questions are, and their desired output. This application is reviewed and approved by the NCHS confidentiality officer (who chairs the NCHS Disclosure Review Board). If application is approved, the RDC and data owners (DHANES/NHANES) will then jointly review the output to ensure it conforms to their approved application and cannot reidentify an individual (as per CIPSEA) before that output is allowed to leave the RDC. Once application is approved, researchers must also undergo confidentially training, data use/access agreement (similar to a Rules of Behavior form), sign a non-disclosure CIPSEA agent form, which makes them agents under the statue and legally liable for proper use of the data and the RDC can prosecute researchers under CIPSEA if the data is misused. [1]	Data Use/Access Agreement (Rules of Behavior) authorizes data use	[1] NCHS RDC Meeting
5.1.2	**Applicable Regulations/Policies**			

Dataset 1 - NHANES				
		Raw Language	**Interpretation**	**Source**
5.1.2.1	Local regulations/policies	Researchers that wish to access data in the RDC must submit an application where they specify what data they need, what their research questions are, and their desired output. This application is reviewed and approved by the NCHS confidentiality officer (who chairs the NCHS Disclosure Review Board). If application is approved, the RDC and data owners (DHANES/NHANES) will then jointly review the output to ensure it conforms to their approved application and cannot reidentify an individual (as per CIPSEA) before that output is allowed to leave the RDC. Once application is approved, researchers must also undergo confidentially training, data use/access agreement (similar to a Rules of Behavior form), sign a non-disclosure CIPSEA agent form, which makes them agents under the statue and legally liable for proper use of the data and the RDC can prosecute researchers under CIPSEA if the data is misused. [1]	NCHS RDC policy	[1] NCHS RDC Meeting
5.1.2.2	Tribal regulations/policies	N/A - Tribal or local areas are not specifically targeted, so no specific agreements. [1]	N/A	[1] NCHS and DHANES Meeting
5.1.2.3	State regulations/policies	Information not available/found	Information not available/found	
5.1.2.4	Federal regulations/policies	NCHS survey data are protected by Federal confidentiality laws including Section 308(d) Public Health Service Act [42 U.S.C. 242m(d)] and the Confidential Information Protection and Statistical Efficiency Act or CIPSEA [Pub. L. No. 115-435, 132 Stat. 5529 § 302]. These confidentiality laws state the data collected by NCHS may be used only for statistical reporting and analysis. [1]	1. Section 308(d) of the Public Health Service Act 2. Confidential Information Protection and Statistical Efficiency Act (CIPSEA)	[1] https://www.cdc.gov/nchs/data_access/restrictions.htm (Accessed: 4/17/23)
5.1.2.5	International regulations/policies	N/A - The National Health and Nutrition Examination Survey (NHANES) is a program of studies designed to assess the health and nutritional status of adults and children in the United States. The sample for the survey is selected to represent the U.S. population of all ages. [1]	N/A	[1] https://www.cdc.gov/nchs/nhanes/about_nhanes.htm (Accessed: 4/17/23)
5.1.2.6	Contractual obligations	Information not available/found	Information not available/found	
5.1.2.7	Repository policies	N/A - the RDC does not have a repository agreement with NHANES. The RDC has a data use agreement with NHANES. The RDC does not archive NHANES data nor is the RDC a repository for NHANES data.	N/A	[1] NCHS RDC Email Communication
5.2	**Governance for data linkage, sharing, access, and use based on data access authorization or applicable regulations/policies (i.e., the origin of the governance)**			
5.2.1	Whether the data can be linked	Does not authorize/specify	Does not authorize/specify	
5.2.2	With what other data can it be linked or can it not be linked (scope of linkage)	Does not authorize/specify	Does not authorize/specify	
5.2.3	Whether data can be shared	Does not authorize/specify	Does not authorize/specify	
5.2.4	How data can be shared (de-identification status, disclosure review)	Does not authorize/specify	Does not authorize/specify	
5.2.5	How data can be accessed (access type, data use agreement, data access committee/group approval, IRB LOD, etc.)	Does not authorize/specify	Does not authorize/specify	

Dataset 1 - NHANES				
		Raw Language	**Interpretation**	**Source**
5.2.6	How data can be used (data use limitations)	Consent template for 'Home Interview Consent': "Data gathered in this survey are used to study many health issues. We are required by law to use your information for statistical research only and to keep it confidential." [1] NCHS survey data are protected by Federal confidentiality laws including Section 308(d) Public Health Service Act [42 U.S.C. 242m(d)] and the Confidential Information Protection and Statistical Efficiency Act or CIPSEA [Pub. L. No. 115-435, 132 Stat. 5529 § 302]. These confidentiality laws state the data collected by NCHS may be used only for statistical reporting and analysis. [2] RDC Data Use Agreement: These data were collected with the assurance that they will be used only for health statistical reporting and analysis. [3]	1. Consent, Section 308(d) Public Health Service Act [42 U.S.C. 242m(d)], and Confidential Information Protection and Statistical Efficiency Act (CIPSEA) specify that the data can only be used for statistical reporting and analysis (includes broad research). 2. RDC DUA specifies that the data were collected with the assurance that they will be used only for health statistical reporting and analysis.	[1] https://wwwn.cdc.gov/nchs/data/nhanes/2019-2020/documents/2019-Home-Interview-Consent-English-508.pdf (Accessed: 4/17/23) [2] https://www.cdc.gov/nchs/data_access/restrictions.htm (Accessed: 4/17/23) [3] https://www.cdc.gov/rdc/data/b4/rdc-data-b4-accessagreement.pdf (Accessed: 4/17/23)
5.2.7	Other (specify)	Does not authorize/specify	Does not authorize/specify	
6	**PII Elements**			
6.1	PII elements collected	Interviewer Procedures Manual refers to collecting last name, middle name, first name, SSN, DOB, sex at birth, gender, address. [1]	Last name, middle name, first name, SSN, DOB, sex at birth, gender, and address are collected from participants	[1] https://wwwn.cdc.gov/nchs/data/nhanes/2021-2023/manuals/2022-Interviewer-Procedures-508.pdf (Accessed: 4/17/23)
6.2	PII elements holder (i.e., party that holds the PII)	PII is stripped before coming to the RDC and resides with the data collector (DHANES/NHANES). [1] What types of confidential variables can I access through the RDC? Direct identifiers (name, social security number, address) cannot be accessed through the RDC. Indirect Identifiers (geography) may be available through the RDC. [2]	Data collector (DHANES/NHANES)	[1] NCHS RDC Meeting [2] https://www.cdc.gov/rdc/b1datatype/Dt100.htm (Accessed: 4/17/23)
6.3	Use of common data model, if any, for data collection	N/A	N/A	
7	**Prior Data Linkages**			
7.1	**Dataset linked with other datasets**			
7.1.1	Name of other linked dataset	NCHS is currently linking various NCHS surveys with administrative data from the following: National Death Index (NDI) Centers for Medicare and Medicaid Services (CMS) Medicare Medicaid/CHIP United States Renal Data System (USRDS) Social Security Administration (SSA) Department of Housing and Urban Development (HUD) Department of Veterans Affairs (VA). [1]	NCHS is currently linking NHANES with the following: National Death Index (NDI) Centers for Medicare and Medicaid Services (CMS) Medicare Medicaid/CHIP United States Renal Data System (USRDS) Social Security Administration (SSA) Department of Housing and Urban Development (HUD) Department of Veterans Affairs (VA)	[1] https://www.cdc.gov/nchs/data-linkage/index.htm (Accessed: 4/17/23)

Dataset 1 - NHANES				
		Raw Language	**Interpretation**	**Source**
7.1.2	Other dataset type (clinical, EHR, survey, claims, SDOH, etc.)	NCHS is currently linking various NCHS surveys with administrative data from the following: National Death Index (NDI) Centers for Medicare and Medicaid Services (CMS) Medicare Medicaid/CHIP United States Renal Data System (USRDS) Social Security Administration (SSA) Department of Housing and Urban Development (HUD) Department of Veterans Affairs (VA). [1]	Administrative	[1] https://www.cdc.gov/nchs/data-linkage/index.htm (Accessed: 4/17/23)
7.1.3	Other dataset source(s)	NCHS is currently linking various NCHS surveys with administrative data from the following: National Death Index (NDI) Centers for Medicare and Medicaid Services (CMS) Medicare Medicaid/CHIP United States Renal Data System (USRDS) Social Security Administration (SSA) Department of Housing and Urban Development (HUD) Department of Veterans Affairs (VA). [1]	Sources of data for NHANES linkages: CDC, CMS, USRDS, SSA, HUD and VA	[1] https://www.cdc.gov/nchs/data-linkage/index.htm (Accessed: 4/17/23)
7.1.4	Linking methodology (PPRL or non-PPRL); linkage technology	Method for NHANES-NDI linkage: The linkage between the NCHS survey data and the NDI was based on both deterministic and probabilistic approaches. The probabilistic approach performs weighting and link adjudication following the Fellegi-Sunter method (9). The Fellegi-Sunter method is the foundational methodology used for record linkage. It estimates the likelihood that each pair is a match before selecting the most probable match between a survey record and NDI record. Following these approaches, a selection process was implemented with the goal of selecting pairs believed to represent the same individual between the data sources. The three main steps taken to link the NCHS survey data to the NDI are as follows: 1. Deterministic linkage was conducted, joining on exact SSN, and validated by comparison of other identifying fields. 2. Probabilistic linkage was conducted, identifying likely matches, or links, between all records. All deterministic matched pairs (from Step 1) were assigned a probabilistic match probability of 1; other records were linked and scored as follows (note that SSN is excluded from the analysis for this step): a. Pairs were formed via blocking. b. Potential matches were scored based on the concurrence of first name, middle initial, last name or father's surname, year of birth, month of birth, day of birth, state of birth, state of residence, race, and sex. c. Match probabilities were estimated through a model which assigned the estimated probability that pairs are matches.	Non-PPRL method was used for NHANES linkages with various data sources	[1] https://www.cdc.gov/nchs/data/datalinkage/2019NDI-Linkage-Methods-and-Analytic-Considerations-508.pdf (Accessed: 4/17/23)

Dataset 1 - NHANES				
		Raw Language	**Interpretation**	**Source**
		3. Pairs were selected which were believed to represent the same individual between the data sources. The pair having the highest estimated match probability was kept as long as it was above the linkage cut-off (see Appendix I). The linkage algorithm was developed with custom code (using SAS 9.4) and was tailored to perform these specific linkages, in order to produce high-quality matches with a low degree of linkage error. More detailed descriptions of the linkage methodology can be found in Appendix I of this report. [1]		
7.1.5	PII elements used for the linkage	For linkage with NDI: The primary identifiers used in the linkages were: SSN9 or SSN4 (depending on the survey year or cycle of the survey), first name, middle initial, last name or father's surname, month of birth, day of birth, year of birth, state of birth, state of residence, race, and sex. [1]	For linkage with NDI, the following PIIs were used: -SSN9 or SSN4 (depending on the survey year or cycle of the survey) - first name - middle initial - last name or father's surname - month of birth - day of birth - year of birth - state of birth - state of residence - race - sex	[1] https://www.cdc.gov/nchs/data/datalinkage/2019NDI-Linkage-Methods-and-Analytic-Considerations-508.pdf (Accessed: 4/17/23)
7.1.6	Entity resolver (data originator or data linker or third party)	NCHS performed the entity resolution	NCHS performed the entity resolution	[1] https://www.cdc.gov/nchs/data/datalinkage/2019NDI-Linkage-Methods-and-Analytic-Considerations-508.pdf (Accessed: 4/17/23)
7.1.7	Party performing the linkages	NCHS performed the data linkage	NCHS performed the data linkage	[1] https://www.cdc.gov/nchs/data/datalinkage/2019NDI-Linkage-Methods-and-Analytic-Considerations-508.pdf (Accessed: 4/17/23)
7.1.8	Linkage quality assessment	An external data source was used to assess the quality of the 2015 LMF and 2019 LMF. Based on the analysis, the 2019 LMF shows a slightly higher concordance with the external benchmark than the 2015 LMF, especially during the years when only SSN4 was collected. These analyses show that the new algorithm has improved linkage accuracy. [1]	An external data source was used to assess the quality of the 2015 LMF and 2019 LMF.	[1] https://www.cdc.gov/nchs/data/datalinkage/2019NDI-Linkage-Methods-and-Analytic-Considerations-508.pdf (Accessed: 4/17/23)
7.1.9	Linked data sharing method (linkage maps or pre-linked dataset)	Additionally, the linked data files are made available in secure facilities for approved research projects. Researchers who want to access the restricted-use 2019 LMF must submit a research proposal to the NCHS Research Data Center (RDC) to obtain permission to access the restricted use files. [1]	Pre-linked data is made available through RDC	[1] https://www.cdc.gov/nchs/data/datalinkage/2019NDI-Linkage-Methods-and-Analytic-Considerations-508.pdf (Accessed: 4/17/23)

USE CASE 1 - GOVERNANCE INFORMATION				
Use Case 1: Effects of the COVID-19 pandemic on mental health of children. Are COVID-19 pandemic related mental health outcomes more severe for children in foster care?				
Dataset 2 - National Survey on Drug Use and Health (NSDUH)				
	Dataset Source	National Survey on Drug Use and Health (NSDUH)		
	Dataset Source Agency	SAMHSA		
	Dataset Type (Clinical, EHR, Survey, SDOH, etc.)	Survey (2017-2020)		
	Information Sources	Meeting with SAMHSA staff; Website		
Dataset 2 - National Survey on Drug Use and Health (NSDUH)				
		Raw Language	**Interpretation**	**Source**
1	**Data Collection**			
1.1	**Authorizations and Applicable Regulations/Policies**			
1.1.1	**Authorizations**			
1.1.1.1	Assent	You must always obtain informed consent from a respondent (and gain permission from a parent/guardian before speaking to a youth respondent aged 12-17 about the study, then obtain informed consent from the parent and youth to participate in the study) by reading the Introduction and Informed Consent script verbatim and providing the Study Description. [1]	Assent from children authorizes data collection	[1] https://www.samhsa.gov/data/sites/default/files/reports/rpt23074/NSDUHmrbFIManual2019.pdf (Accessed: 4/20/23)
1.1.1.2	Consent	You must always obtain informed consent from a respondent (and gain permission from a parent/guardian before speaking to a youth respondent aged 12-17 about the study, then obtain informed consent from the parent and youth to participate in the study) by reading the Introduction and Informed Consent script verbatim and providing the Study Description. [1]	Consent from adults authorizes data collection	[1] https://www.samhsa.gov/data/sites/default/files/reports/rpt23074/NSDUHmrbFIManual2019.pdf (Accessed: 4/20/23)
1.1.1.3	IRB/equivalent Privacy Board determination	RTI International takes great precautions to ensure all its surveys, including NSDUH, uphold the rights of our participants. Like other research studies RTI conducts, NSDUH was reviewed by one of RTI's Institutional Review Boards (IRBs) before any interviews were conducted. The IRB reviews the study's protocol following guidelines from the U.S. Department of Health and Human Service's Office for Human Research Protections.	RTI (DCC for NSDUH) IRB authorizes data collection	[1] https://nsduhweb.rti.org/respweb/confidentiality.html (Accessed: 4/20/23)
1.1.1.4	Local/state/federal law	The National Survey on Drug Use and Health (NSDUH) provides national and state-level data on the use of tobacco, alcohol, illicit drugs (including non-medical use of prescription drugs) and mental health in the United States. NSDUH began in 1971 and is currently conducted on an annual basis. NSDUH is authorized by Section 505 of the Public Health Service Act, which requires annual surveys to collect data on the level and patterns of substance use.	Public Health Service Act Section 505 authorizes data collection	[1] https://nsduhweb.rti.org/respweb/faq.html (Accessed: 4/20/23)
1.1.1.5	Institutional Certification	Information not available/found	Information not available/found	
1.1.1.6	Data originator agreement	Information not available/found	Information not available/found	

Dataset 2 - National Survey on Drug Use and Health (NSDUH)				
		Raw Language	**Interpretation**	**Source**
1.1.1.7	Repository agreements/policies	Information not available/found	Information not available/found	
1.1.1.8	Other (specify)	Information not available/found	Information not available/found	
1.1.2	**Applicable Regulations/Policies**			
1.1.2.1	Local regulations/policies	Because of the coronavirus disease 2019 (COVID-19) pandemic, major changes were made to the methods used in data collection in 2020. There is no way to separate out the true changes in behavior from the changes due to the new methodology. The main methodological changes were: Almost no data collection from mid-March through September 2020, Introduction of web data collection in October 2020 with very limited in-person data collection, and Additions to the questionnaire beginning in October 2020. The 2020 NSDUH is missing two quarters of data. Tests of data from before 2020 show that estimates based on just quarters 1 and 4 are not comparable to estimates based on the entire year. This indicates that 2020 estimates should not be compared to previous years. Repeated analyses have showed that web responses are not comparable to in-person responses, and that the comparability is not consistent in a way that we can fully account for. For these reasons, it is not recommended to compare any estimates from 2020 to estimates from 2019 or earlier. [1]	SAMHSA recommendation	[1] https://www.samhsa.gov/data/faq/concerns-about-trend-comparability/why-does-samhsa-caution-against-comparing-2020-estimates-estimates-prior-years (Accessed: 9/12/23)
1.1.2.2	Tribal regulations/policies	Information not available/found	Information not available/found	
1.1.2.3	State regulations/policies	Information not available/found	Information not available/found	
1.1.2.4	Federal regulations/policies	NSDUH collects highly-sensitive information from individual respondents, including data on substance use and mental health concerns. Ensuring respondent confidentiality is crucial to encouraging participation in the survey and responses are protected under the Confidential Information Protection and Statistical Efficiency Act (CIPSEA). [1]	Confidential Information Protection and Statistical Efficiency Act (CIPSEA)	[1] https://www.datafiles.samhsa.gov/get-help/public-vs-restricted-use/what-difference-between-nsduh-public-and-restricted-use-data (Accessed: 4/20/23)
1.1.2.5	International regulations/policies	N/A - The National Survey on Drug Use and Health (NSDUH), conducted annually by the Substance Abuse and Mental Health Services Administration (SAMHSA), provides nationally representative data on the use of tobacco, alcohol, and illicit drugs; substance use disorders; receipt of substance use treatment; mental health issues; and the use of mental health services among the civilian, noninstitutionalized population aged 12 or older in the United States. [1]	N/A	[1] https://www.samhsa.gov/data/data-we-collect/nsduh-national-survey-drug-use-and-health (Accessed: 4/20/23)
1.1.2.6	Contractual obligations	Information not available/found	Information not available/found	
1.1.2.7	Repository policies	Information not available/found	Information not available/found	
1.2	**Governance for data linkage, sharing, access, and use based on data collection authorization or applicable regulations/policies (i.e., the origin of the governance)**			

Dataset 2 - National Survey on Drug Use and Health (NSDUH)				
		Raw Language	**Interpretation**	**Source**
1.2.1	Whether the data can be linked	Does not authorize/specify - linkage is not specified in any authorization. [1]	Does not authorize/specify	[1] NDACAN Meeting
1.2.2	With what other data can it be linked or can it not be linked (scope of linkage)	Does not authorize/specify - linkage is not specified in any authorization. [1]	Does not authorize/specify	[1] NDACAN Meeting
1.2.3	Whether data can be shared	Information not available/found	Information not available/found	
1.2.4	How data can be shared (de-identification status, disclosure review)	Information not available/found	Information not available/found	
1.2.5	How data can be accessed (access type, data use agreement, data access committee/group approval, IRB LOD, etc.)	Information not available/found	Information not available/found	
1.2.6	How data can be used (data use limitations)	Information not available/found	Information not available/found	
1.2.7	Other (specify)	Because of the coronavirus disease 2019 (COVID-19) pandemic, major changes were made to the methods used in data collection in 2020. There is no way to separate out the true changes in behavior from the changes due to the new methodology. The main methodological changes were: Almost no data collection from mid-March through September 2020, Introduction of web data collection in October 2020 with very limited in-person data collection, and Additions to the questionnaire beginning in October 2020. The 2020 NSDUH is missing two quarters of data. Tests of data from before 2020 show that estimates based on just quarters 1 and 4 are not comparable to estimates based on the entire year. This indicates that 2020 estimates should not be compared to previous years. Repeated analyses have showed that web responses are not comparable to in-person responses, and that the comparability is not consistent in a way that we can fully account for. For these reasons, it is not recommended to compare any estimates from 2020 to estimates from 2019 or earlier. [1]	SAMHSA recommends against comparing 2020 NSDUH data with prior years due to methodological changes	[1] https://www.samhsa.gov/data/faq/concerns-about-trend-comparability/why-does-samhsa-caution-against-comparing-2020-estimates-estimates-prior-years (Accessed: 9/12/23)
2	**Data Linkage**			
2.1	**Authorizations and Applicable Regulations/Policies**			
2.1.1	**Authorizations**			
2.1.1.1	Assent	Information not available/found	Information not available/found	
2.1.1.2	Consent	Information not available/found	Information not available/found	

Dataset 2 - National Survey on Drug Use and Health (NSDUH)				
		Raw Language	**Interpretation**	**Source**
2.1.1.3	IRB/equivalent Privacy Board determination	Information not available/found	Information not available/found	
2.1.1.4	Local/state/federal laws	Information not available/found	Information not available/found	
2.1.1.5	Institutional Certification	Information not available/found	Information not available/found	
2.1.1.6	Data originator agreement	Information not available/found	Information not available/found	
2.1.1.7	Repository agreements/policies	Information not available/found	Information not available/found	
2.1.1.8	Other (specify)	Information not available/found	Information not available/found	
2.1.2	**Applicable Regulations/Policies**			
2.1.2.1	Local regulations/policies	Information not available/found	Information not available/found	
2.1.2.2	Tribal regulations/policies	Information not available/found	Information not available/found	
2.1.2.3	State regulations/policies	Information not available/found	Information not available/found	
2.1.2.4	Federal regulations/policies	Information not available/found	Information not available/found	
2.1.2.5	International regulations/policies	N/A - The National Survey on Drug Use and Health (NSDUH), conducted annually by the Substance Abuse and Mental Health Services Administration (SAMHSA), provides nationally representative data on the use of tobacco, alcohol, and illicit drugs; substance use disorders; receipt of substance use treatment; mental health issues; and the use of mental health services among the civilian, noninstitutionalized population aged 12 or older in the United States. [1]	N/A	[1] https://www.samhsa.gov/data/data-we-collect/nsduh-national-survey-drug-use-and-health (Accessed: 4/20/23)
2.1.2.6	Contractual obligations	Information not available/found	Information not available/found	
2.1.2.7	Repository policies	Information not available/found	Information not available/found	
2.2	**Governance for data linkage, sharing, access, and use based on data linkage authorization or applicable regulations/policies (i.e., the origin of the governance)**			
2.2.1	Whether the data can be linked	Information not available/found	Information not available/found	
2.2.2	With what other data can it be linked or can it not be linked (scope of linkage)	Information not available/found	Information not available/found	
2.2.3	Whether data can be shared	Information not available/found	Information not available/found	
2.2.4	How data can be shared (de-identification status, disclosure review)	Information not available/found	Information not available/found	
2.2.5	How data can be accessed (access type, data use agreement, data access committee/group approval, IRB LOD, etc.)	Information not available/found	Information not available/found	
2.2.6	How data can be used (data use limitations)	Information not available/found	Information not available/found	
2.2.7	Other (specify)	Information not available/found	Information not available/found	
3	**Data Sharing**			
3.1	**Authorizations and Applicable Regulations/Policies**			
3.1.1	**Authorizations**			

Dataset 2 - National Survey on Drug Use and Health (NSDUH)				
		Raw Language	**Interpretation**	**Source**
3.1.1.1	Assent	Information not available/found	Information not available/found	
3.1.1.2	Consent	Information not available/found	Information not available/found	
3.1.1.3	IRB/equivalent Privacy Board determination	Information not available/found	Information not available/found	
3.1.1.4	Local/state/federal laws	Information not available/found	Information not available/found	
3.1.1.5	Institutional Certification	Information not available/found	Information not available/found	
3.1.1.6	Data originator agreement	Information not available/found	Information not available/found	
3.1.1.7	Repository agreements/policies	Information not available/found	Information not available/found	
3.1.1.8	Other	Information not available/found	Information not available/found	
3.1.2	**Applicable Regulations/Policies**			
3.1.2.1	Local regulations/policies	Information not available/found	Information not available/found	
3.1.2.2	Tribal regulations/policies	Information not available/found	Information not available/found	
3.1.2.3	State regulations/policies	Information not available/found	Information not available/found	
3.1.2.4	Federal regulations/policies	NSDUH collects highly-sensitive information from individual respondents, including data on substance use and mental health concerns. Ensuring respondent confidentiality is crucial to encouraging participation in the survey and responses are protected under the Confidential Information Protection and Statistical Efficiency Act (CIPSEA). [1]	Confidential Information Protection and Statistical Efficiency Act (CIPSEA)	[1] https://www.datafiles.samhsa.gov/get-help/public-vs-restricted-use/what-difference-between-nsduh-public-and-restricted-use-data (Accessed: 4/20/23)
3.1.2.5	International regulations/policies	N/A - The National Survey on Drug Use and Health (NSDUH), conducted annually by the Substance Abuse and Mental Health Services Administration (SAMHSA), provides nationally representative data on the use of tobacco, alcohol, and illicit drugs; substance use disorders; receipt of substance use treatment; mental health issues; and the use of mental health services among the civilian, noninstitutionalized population aged 12 or older in the United States. [1]	N/A	[1] https://www.samhsa.gov/data/data-we-collect/nsduh-national-survey-drug-use-and-health (Accessed: 4/20/23)
3.1.2.6	Contractual obligations	Information not available/found	Information not available/found	
3.1.2.7	Repository policies	Information not available/found	Information not available/found	
3.2	**Governance for data linkage, sharing, access, and use based on data sharing authorization or applicable regulations/policies (i.e., the origin of the governance)**			
3.2.1	Whether the data can be linked	Information not available/found	Information not available/found	
3.2.2	With what other data can it be linked or can it not be linked (scope of linkage)	Information not available/found	Information not available/found	
3.2.3	Whether data can be shared	Information not available/found	Information not available/found	
3.2.4	How data can be shared (de-identification status, disclosure review)	Information not available/found	Information not available/found	
3.2.5	How data can be accessed (access type, data use agreement, data access committee/group approval, IRB LOD, etc.)	Information not available/found	Information not available/found	

Dataset 2 - National Survey on Drug Use and Health (NSDUH)				
		Raw Language	**Interpretation**	**Source**
3.2.6	How data can be used (data use limitations)	Information not available/found	Information not available/found	
3.2.7	Other (specify)	Information not available/found	Information not available/found	
4	**Data Access**			
4.1	**Authorizations and Applicable Regulations/Policies**			
4.1.1	**Authorizations**			
4.1.1.1	Assent	Information not available/found	Information not available/found	
4.1.1.2	Consent	Information not available/found	Information not available/found	
4.1.1.3	IRB/equivalent Privacy Board determination	Information not available/found	Information not available/found	
4.1.1.4	Local/state/federal laws	Information not available/found	Information not available/found	
4.1.1.5	Institutional Certification	Information not available/found	Information not available/found	
4.1.1.6	Data originator agreement	Information not available/found	Information not available/found	
4.1.1.7	Repository agreements/policies	All analysts entering the RDC must sign the DAA (Data Access Agreement) form. For students wanting to access NSDUH RUF, both students and their advisors must also sign the SAMHSA RDC Student Data User Acknowledgement form. [1]	Two repository agrements authorize data access: 1. Data Access Agreement (DAA) form 2. For students, the SAMHSA RDC Student Data User Acknowledgement form signed by the student and their advisor	[1] https://www.samhsa.gov/data/data-we-collect/samhsa-rdc (Accessed: 4/20/23)
4.1.1.8	Other (specify)	Information not available/found	Information not available/found	
4.1.2	**Applicable Regulations/Policies**			
4.1.2.1	Local regulations/policies	Information not available/found	Information not available/found	
4.1.2.2	Tribal regulations/policies	Information not available/found	Information not available/found	
4.1.2.3	State regulations/policies	Information not available/found	Information not available/found	
4.1.2.4	Federal regulations/policies	All researchers on the project must complete confidentiality training and sign and complete the Designated Agent Form (DAF) with a notary signature. Signing the DAF allows researchers to become designated agents to access CIPSEA protected data. In addition, all analysts entering the RDC must sign the DAA (Data Access Agreement) form. The training certificate and signed DAA and DAF must be submitted to be considered a complete package. For students wanting to access NSDUH RUF, both students and their advisors must also sign the SAMHSA RDC Student Data User Acknowledgement form. [1]	Confidential Information Protection and Statistical Efficiency Act (CIPSEA)	[1] https://www.samhsa.gov/data/data-we-collect/samhsa-rdc (Accessed: 4/20/23)

Dataset 2 - National Survey on Drug Use and Health (NSDUH)				
		Raw Language	**Interpretation**	**Source**
4.1.2.5	International regulations/policies	N/A - The National Survey on Drug Use and Health (NSDUH), conducted annually by the Substance Abuse and Mental Health Services Administration (SAMHSA), provides nationally representative data on the use of tobacco, alcohol, and illicit drugs; substance use disorders; receipt of substance use treatment; mental health issues; and the use of mental health services among the civilian, noninstitutionalized population aged 12 or older in the United States. [1]	N/A	[1] https://www.samhsa.gov/data/data-we-collect/nsduh-national-survey-drug-use-and-health (Accessed: 4/20/23)
4.1.2.6	Contractual obligations	Information not available/found	Information not available/found	
4.1.2.7	Repository policies	Prospective researchers must submit an RDC application, also known as an RDC proposal, that will be reviewed by the SAMHSA RDC team. The proposal must be approved before any other procedures can happen. [1] All researchers on the project must complete confidentiality training and sign and complete the Designated Agent Form (DAF) with a notary signature. Signing the DAF allows researchers to become designated agents to access CIPSEA protected data. In addition, all analysts entering the RDC must sign the DAA (Data Access Agreement) form. The training certificate and signed DAA and DAF must be submitted to be considered a complete package. For students wanting to access NSDUH RUF, both students and their advisors must also sign the SAMHSA RDC Student Data User Acknowledgement form. [1]	SAMHSA RDC policy	[1] https://www.samhsa.gov/data/data-we-collect/samhsa-rdc (Accessed: 4/20/23)
4.2	**Governance for data linkage, sharing, access, and use based on data access authorization or applicable regulations/policies (i.e., the origin of the governance)**			
4.2.1	Whether the data can be linked	Information not available/found	Information not available/found	
4.2.2	With what other data can it be linked or can it not be linked (scope of linkage)	Information not available/found	Information not available/found	
4.2.3	Whether data can be shared	Information not available/found	Information not available/found	
4.2.4	How data can be shared (de-identification status, disclosure review)	Information not available/found	Information not available/found	

Dataset 2 - National Survey on Drug Use and Health (NSDUH)				
		Raw Language	**Interpretation**	**Source**
4.2.5	How data can be accessed (access type, data use agreement, data access committee/group approval, IRB LOD, etc.)	Prospective researchers must submit an RDC application, also known as an RDC proposal, that will be reviewed by the SAMHSA RDC team. The proposal must be approved before any other procedures can happen. SAMHSA review is to make sure that all requirements as specified in "Guidelines for SAMHSA RDC Data Users" and "RDC sample proposal" are carefully followed. In addition to the format and completeness, the following two aspects are of particular importance for the application to pass SAMHSA review: The feasibility of existing data to the project, that is, whether it is possible for the research to be conducted with the available information. On occasion, it is clear from the outset that the sample will not support the intended analysis. For instance, NSDUH does not allow for individual-level record linkage. The risk of disclosure of restricted information, that is, whether the analysis can be conducted without compromising the confidentiality promised to all respondents (children, adults, households, neighborhoods). We may ask the researchers or the data users to provide additional clarifications and revisions if it is deemed necessary. The application will be approved if all requirements are met. Approval of the proposal does not constitute endorsement by SAMHSA of the substantive, methodological, theoretical, policy relevance, or scientific aspects of the proposed research. [1] All researchers on the project must complete confidentiality training and sign and complete the Designated Agent Form (DAF) with a notary signature. Signing the DAF allows researchers to	SAMHSA RDC specifies that for data access, a user: 1. Submit RDC application 2. Obtain approval from SAMHSA staff on the research proposed in the RDC application 3. Sign and complete Designated Agent Form (DAF) 4. Sign and complete Data Access Agreement (DAA) form 5. Complete confidentiality training 6. Students and their advisors must also sign the SAMHSA RDC Student Data User Acknowledgement form 7. Access data only through RDC	[1] https://www.samhsa.gov/data/data-we-collect/samhsa-rdc (Accessed: 4/20/23)

Dataset 2 - National Survey on Drug Use and Health (NSDUH)				
		Raw Language	**Interpretation**	**Source**
		become designated agents to access CIPSEA protected data. In addition, all analysts entering the RDC must sign the DAA (Data Access Agreement) form. The training certificate and signed DAA and DAF must be submitted to be considered a complete package. For students wanting to access NSDUH RUF, both students and their advisors must also sign the SAMHSA RDC Student Data User Acknowledgement form. [1] The research data center (RDC) program provides a mechanism for data users to access NSDUH restricted-use data files in a secure, confidentiality-compliant manner. SAMHSA RDC does not have our own RDC sites. SAMHSA RDC collaborates with the National Center for Health Statistics (NCHS) RDC and the Federal Statistical Research Data Centers (FSRDC) to carry out the NSDUH RDC program. All SAMHSA RDC users should carefully read "Guidelines for SAMHSA RDC Data Users" before accessing RUF data. [1]		
4.2.6	How data can be used (data use limitations)	Information not available/found	Information not available/found	
4.2.7	Other (specify)	Information not available/found	Information not available/found	
5	**Data Use**			
5.1	**Authorizations and Applicable Regulations/Policies**			
5.1.1	**Authorizations**			
5.1.1.1	Assent	Information not available/found	Information not available/found	
5.1.1.2	Consent	Information not available/found	Information not available/found	
5.1.1.3	IRB/equivalent Privacy Board determination	Information not available/found	Information not available/found	
5.1.1.4	Local/state/federal laws	Information not available/found	Information not available/found	
5.1.1.5	Institutional Certification	Information not available/found	Information not available/found	
5.1.1.6	Data originator agreement	Information not available/found	Information not available/found	
5.1.1.7	Repository agreements/policies	All analysts entering the RDC must sign the DAA (Data Access Agreement) form. [1]	RDC Data Access Agreement (DAA) form authorizes data use	[1] https://www.samhsa.gov/data/data-we-collect/samhsa-rdc (Accessed: 4/20/23)
5.1.1.8	Other (specify)	Information not available/found	Information not available/found	
5.1.2	**Applicable Regulations/Policies**			
5.1.2.1	Local regulations/policies	Information not available/found	Information not available/found	
5.1.2.2	Tribal regulations/policies	Information not available/found	Information not available/found	

Dataset 2 - National Survey on Drug Use and Health (NSDUH)				
		Raw Language	**Interpretation**	**Source**
5.1.2.3	State regulations/policies	Study description: This study, sponsored by the U.S. Department of Health and Human Services, collects information for research and program planning by asking about: • tobacco, alcohol, and drug use or non-use, • knowledge and attitudes about drugs, • mental health, and • other health issues. You cannot be identified through any information you give us. Your name and address will never be connected to your answers. Also, federal law requires us to keep all of your answers confidential. Any data that you provide will only be used by authorized personnel for statistical purposes according to the Confidential Information Protection and Statistical Efficiency Act of 2002. [1]	Confidential Information Protection and Statistical Efficiency Act (CIPSEA)	[1] https://www.samhsa.gov/data/sites/default/files/reports/rpt23074/NSDUHmrbFIManual2019.pdf (Accessed: 4/20/23)
5.1.2.4	Federal regulations/policies	Information not available/found	Information not available/found	
5.1.2.5	International regulations/policies	N/A - The National Survey on Drug Use and Health (NSDUH), conducted annually by the Substance Abuse and Mental Health Services Administration (SAMHSA), provides nationally representative data on the use of tobacco, alcohol, and illicit drugs; substance use disorders; receipt of substance use treatment; mental health issues; and the use of mental health services among the civilian, noninstitutionalized population aged 12 or older in the United States. [1]	N/A	[1] https://www.samhsa.gov/data/data-we-collect/nsduh-national-survey-drug-use-and-health (Accessed: 4/20/23)
5.1.2.6	Contractual obligations	Information not available/found	Information not available/found	
5.1.2.7	Repository policies	The Substance Abuse and Mental Health Data Archive (SAMHDA) provides a public web portal that allows people to review and download public-use and restricted-use data files and documentation. SAMHSA has partnered with the National Center for Health Statistics (NCHS) to host restricted-use NSDUH data at its Federal Statistical Research Data Centers (RDCs). RDCs are secure facilities that provide access to a range of restricted-use microdata for statistical purposes. SAMHSA is the most recent federal partner to work with NCHS in making NSDUH restricted-use microdata available to approved researchers at RDC sites. Eligible researchers may now apply for access to these data through SAMHSA's RDC website.	SAMHSA RDC policy	[1] https://www.datafiles.samhsa.gov/about-us/policies (Accessed: 4/20/23)
5.2	**Governance for data linkage, sharing, access, and use based on data access authorization or applicable regulations/policies (i.e., the origin of the governance)**			
5.2.1	Whether the data can be linked	Information not available/found	Information not available/found	

Dataset 2 - National Survey on Drug Use and Health (NSDUH)				
		Raw Language	**Interpretation**	**Source**
5.2.2	With what other data can it be linked or can it not be linked (scope of linkage)	Information not available/found	Information not available/found	
5.2.3	Whether data can be shared	Information not available/found	Information not available/found	
5.2.4	How data can be shared (de-identification status, disclosure review)	Information not available/found	Information not available/found	
5.2.5	How data can be accessed (access type, data use agreement, data access committee/group approval, IRB LOD, etc.)	Information not available/found	Information not available/found	
5.2.6	How data can be used (data use limitations)	RDC Data Access Agreement (DAA): These data were collected with the assurance that they will be used only for health statistical reporting and analysis.	RDC DAA specifies that the data were collected with the assurance that they will be used only for health statistical reporting and analysis	[1] https://www.samhsa.gov/data/sites/default/files/2021-11/Access-Agreement-20211101-Fillable.pdf (Accessed: 4/20/23)
5.2.7	Other (specify)	Information not available/found	Information not available/found	
6	**PII Elements**			
6.1	PII elements collected	Information not available/found	Information not available/found	
6.2	PII elements holder (i.e., party that holds the PII)	Information not available/found	Information not available/found	
6.3	Use of common data model, if any, for data collection	Information not available/found	Information not available/found	
7	**Prior Data Linkages**			
7.1	**Dataset linked with other datasets**			
7.1.1	Name of other linked dataset	Information not available/found	Information not available/found	
7.1.2	Other dataset type (clinical, EHR, survey, claims, SDOH, etc.)	Information not available/found	Information not available/found	
7.1.3	Other dataset source(s)	Information not available/found	Information not available/found	
7.1.4	Linking methodology (PPRL or non-PPRL); linkage technology	Information not available/found	Information not available/found	
7.1.5	PII elements used for the linkage	Information not available/found	Information not available/found	
7.1.6	Entity resolver (data originator or data linker or third party)	Information not available/found	Information not available/found	
7.1.7	Party performing the linkages	Information not available/found	Information not available/found	
7.1.8	Linkage quality assessment	Information not available/found	Information not available/found	
7.1.9	Linked data sharing method (linkage maps or pre-linked dataset)	Information not available/found	Information not available/found	

USE CASE 1 - GOVERNANCE INFORMATION				
Use Case 1: Effects of the COVID-19 pandemic on mental health of children. Are COVID-19 pandemic related mental health outcomes more severe for children in foster care?				
Dataset 3 - Monitoring the Future (MTF): A Continuing Study of American Youth (Restricted-Use)				
	Dataset Source	National Addiction and HIV Data Archive Program (NAHDAP)		
	Dataset Source Agency	Funded by NIDA, conducted by Survey Research Center in the Institute for Social Research at the University of Michigan		
	Dataset Type (Clinical, EHR, Survey, SDOH, etc.)	Survey (2017-2021)		
	Information Sources	U-Mich Legal meeting, NAHDAP staff meeting; MTF team email communication; website		
Dataset 3 - Monitoring the Future (MTF)				
		Raw Language	**Interpretation**	**Source**
1	**Data Collection**			
1.1	**Authorizations and Applicable Regulations/Policies**			
1.1.1	**Authorizations**		1. Assent from children 2. Consent from parents 3. MTF (U-Mich) IRB approval	
1.1.1.1	Assent	MTF is an ongoing study; MTF obtains assent from children, but also has a consent process with the parents for data collection. [1]	Assent from children authorizes data collection	[1] U-Mich Legal Meeting
1.1.1.2	Consent	Consent forms are sent to the parents of the targeted respondents 3 weeks prior to the targeted survey dates; this survey is sent to school administrators, which are then sent to the parents who elicit the consent forms. [1] Informed consent (active or passive, per school policy) is obtained from parents of students younger than 18 years and from students aged 18 years or older. About three weeks prior to the questionnaire administration date, parents of the target respondents are sent a letter by first-class mail, usually from the principal, announcing and describing the MTF study and providing parents with an opportunity to decline participation of their child if they wish. [2]	Consent from parents authorizes data collection	[1] NAHDAP Meeting [2] https://monitoringthefuture.org/wp-content/uploads/2022/12/mtf2022.pdf (Accessed: 4/18/23)
1.1.1.3	IRB/equivalent Privacy Board determination	The MTF PIs needed approval from their IRB to collect the data. [1]	MTF (U-Mich) IRB approval authorizes data collection	[1] NAHDAP Meeting
1.1.1.4	Local/state/federal law	Information not available/found	Information not available/found	
1.1.1.5	Institutional Certification	Information not available/found	Information not available/found	
1.1.1.6	Data originator agreement	Information not available/found	Information not available/found	
1.1.1.7	Repository agreements/policies	Information not available/found	Information not available/found	
1.1.1.8	Other (specify)	Information not available/found	Information not available/found	
1.1.2	**Applicable Regulations/Policies**			
1.1.2.1	Local regulations/policies	Information not available/found	Information not available/found	
1.1.2.2	Tribal regulations/policies	Information not available/found	Information not available/found	

Dataset 3 - Monitoring the Future (MTF)				
		Raw Language	**Interpretation**	**Source**
1.1.2.3	State regulations/policies	Information not available/found	Information not available/found	
1.1.2.4	Federal regulations/policies	Information not available/found	Information not available/found	
1.1.2.5	International regulations/policies	N/A - Since 1975 the MTF survey has measured drug and alcohol use and related attitudes among adolescent students nationwide. [1]	N/A	[1] https://nida.nih.gov/research-topics/trends-statistics/monitoring-future (Accessed: 4/18/23)
1.1.2.6	Contractual obligations	Information not available/found	Information not available/found	
1.1.2.7	Repository policies	Information not available/found	Information not available/found	
1.2	**Governance for data linkage, sharing, access, and use based on data collection authorization or applicable regulations/policies (i.e., the origin of the governance)**			
1.2.1	Whether the data can be linked	Does not authorize/specify - U-Mich Legal is unaware of any previous linkage of the MTF data; the consent/assent does not have any language that pertains to linkage. [1]	Does not authorize/specify	[1] U-Mich Legal Meeting
1.2.2	With what other data can it be linked or can it not be linked (scope of linkage)	Does not authorize/specify - U-Mich Legal is unaware of any previous linkage of the MTF data; the consent/assent does not have any language that pertains to linkage. [1]	Does not authorize/specify	[1] U-Mich Legal Meeting
1.2.3	Whether data can be shared	Information not available/found - The consent references the Grant of confidentiality from the U.S. Department of Justice, and may also include language about sharing de-identified data. [1]	Information not available/found	[1] U-Mich Legal Meeting
1.2.4	How data can be shared (de-identification status, disclosure review)	Information not available/found	Information not available/found	
1.2.5	How data can be accessed (access type, data use agreement, data access committee/group approval, IRB LOD, etc.)	Does not authorize/specify	Does not authorize/specify	
1.2.6	How data can be used (data use limitations)	The consent specifies that the data can be used for broad research. [1]	Consent specifies that the data can be used for broad research	[1] U-Mich Legal Meeting
1.2.7	Other (specify)	Information not available/found	Information not available/found	
2	**Data Linkage**			
2.1	**Authorizations and Applicable Regulations/Policies**			
2.1.1	**Authorizations**			
2.1.1.1	Assent	Does not authorize/specify - U-Mich Legal is unaware of any previous linkage of the MTF data; the consent/assent does not have any language that pertains to linkage. [1]	Does not authorize/specify	[1] U-Mich Legal Meeting
2.1.1.2	Consent	Does not authorize/specify - U-Mich Legal is unaware of any previous linkage of the MTF data; the consent/assent does not have any language that pertains to linkage. [1]	Does not authorize/specify	[1] U-Mich Legal Meeting
2.1.1.3	IRB/equivalent Privacy Board determination	Does not authorize/specify	Does not authorize/specify	
2.1.1.4	Local/state/federal laws	Does not authorize/specify	Does not authorize/specify	
2.1.1.5	Institutional Certification	Does not authorize/specify	Does not authorize/specify	
2.1.1.6	Data originator agreement	Does not authorize/specify	Does not authorize/specify	
2.1.1.7	Repository agreements/policies	Does not authorize/specify	Does not authorize/specify	
2.1.1.8	Other (specify)	Information not available/found	Information not available/found	
2.1.2	**Applicable Regulations/Policies**			

Dataset 3 - Monitoring the Future (MTF)				
		Raw Language	**Interpretation**	**Source**
2.1.2.1	Local regulations/policies	Information not available/found	Information not available/found	
2.1.2.2	Tribal regulations/policies	Information not available/found	Information not available/found	
2.1.2.3	State regulations/policies	Information not available/found	Information not available/found	
2.1.2.4	Federal regulations/policies	Information not available/found	Information not available/found	
2.1.2.5	International regulations/policies	N/A - Since 1975 the MTF survey has measured drug and alcohol use and related attitudes among adolescent students nationwide. [1]	N/A	[1] https://nida.nih.gov/research-topics/trends-statistics/monitoring-future (Accessed: 4/18/23)
2.1.2.6	Contractual obligations	Information not available/found	Information not available/found	
2.1.2.7	Repository policies	Information not available/found	Information not available/found	
2.2	**Governance for data linkage, sharing, access, and use based on data linkage authorization or applicable regulations/policies (i.e., the origin of the governance)**			
2.2.1	Whether the data can be linked	Does not authorize/specify	Does not authorize/specify	
2.2.2	With what other data can it be linked or can it not be linked (scope of linkage)	Does not authorize/specify	Does not authorize/specify	
2.2.3	Whether data can be shared	Does not authorize/specify	Does not authorize/specify	
2.2.4	How data can be shared (de-identification status, disclosure review)	Does not authorize/specify	Does not authorize/specify	
2.2.5	How data can be accessed (access type, data use agreement, data access committee/group approval, IRB LOD, etc.)	Does not authorize/specify	Does not authorize/specify	
2.2.6	How data can be used (data use limitations)	Does not authorize/specify	Does not authorize/specify	
2.2.7	Other (specify)	Information not available/found	Information not available/found	
3	**Data Sharing**			
3.1	**Authorizations and Applicable Regulations/Policies**			
3.1.1	**Authorizations**		1. ICPSR (U-Mich) IRB approval 2. MTF (U-Mich) IRB approval 3. MTF PI determination (for sharing/accessing data through ICPSR VDE) 4. Family Educational Rights and Privacy Act (FERPA)	
3.1.1.1	Assent	Information not available/found - The consent references the Grant of confidentiality from the U.S. Department of Justice, and may also include language about sharing de-identified data. [1]	Information not available/found	[1] U-Mich Legal Meeting
3.1.1.2	Consent	Information not available/found - The consent references the Grant of confidentiality from the U.S. Department of Justice, and may also include language about sharing de-identified data. [1]	Information not available/found	[1] U-Mich Legal Meeting

Dataset 3 - Monitoring the Future (MTF)				
		Raw Language	**Interpretation**	**Source**
3.1.1.3	IRB/equivalent Privacy Board determination	For MTF, the RUDDDA was not executed since the data originating organization is also within U-Mich, and the practice at the time MTF was deposited; the current practice is to execute a memorandum of understanding for studies within U-Mich to deposit data (however, MTF also does not have a memorandum of understanding in place). In lieu of these documents (RUDDDA and MoU), MTF did have IRB oversight of ICPSR taking over the dissemination of the restricted-use data; the IRB number of the MTF study is linked to ICPSR's reference number (for ICPSR's IRB)… U-Mich authorized NAHDAP to share the MTF data; IRB oversight was provided and authorization from the MTF PI. [1] ICPSR accepts data with identifying information under conditions consistent with the informed consent of the study participants and the relevant Institutional Review Board (IRB) approval. [2]	Two IRBs authorize data sharing: 1. ICPSR (U-Mich) IRB 2. MTF (U-Mich) IRB	[1] NAHDAP Meeting [2] https://www.icpsr.umich.edu/web/pages/datamanagement/confidentiality/ (Accessed: 4/18/23)
3.1.1.4	Local/state/federal laws	Generally, schools must have written permission from the parent or eligible student in order to release any information from a student's education record. However, FERPA allows schools to disclose those records, without consent, to the following parties or under the following conditions (34 CFR § 99.31): -School officials with legitimate educational interest; -Other schools to which a student is transferring; -Specified officials for audit or evaluation purposes; -Appropriate parties in connection with financial aid to a student; -Organizations conducting certain studies for or on behalf of the school; -Accrediting organizations; -To comply with a judicial order or lawfully issued subpoena; -Appropriate officials in cases of health and safety emergencies; and -State and local authorities, within a juvenile justice system, pursuant to specific State law. [1]	Family Educational Rights and Privacy Act (FERPA) authorize data sharing	[1] https://www2.ed.gov/policy/gen/guid/fpco/ferpa/index.html (Accessed: 4/25/23)
3.1.1.5	Institutional Certification	Information not available/found	Information not available/found	
3.1.1.6	Data originator agreement	N/A - For MTF, the RUDDDA was not executed since the data originating organization is also within U-Mich, and the practice at the time MTF was deposited; the current practice is to execute a memorandum of understanding for studies within U-Mich to deposit data (however, MTF also does not have a memorandum of understanding in place). [1]	N/A	[1] NAHDAP Meeting
3.1.1.7	Repository agreements/policies	Information not available/found	Information not available/found	

Dataset 3 - Monitoring the Future (MTF)				
		Raw Language	**Interpretation**	**Source**
3.1.1.8	Other	The determination for sharing data through the VDE is made by the NADHAP staff based on the level of disclosure risk and/or discussion with the data originator. For MTF, the PI requested that the restricted-use MTF data be shared/accessed through VDE. [1]	MTF PI authorizes data sharing (through the ICPSR VDE)	[1] NAHDAP meeting
3.1.2	**Applicable Regulations/Policies**			
3.1.2.1	Local regulations/policies	Information not available/found	Information not available/found	
3.1.2.2	Tribal regulations/policies	Information not available/found	Information not available/found	
3.1.2.3	State regulations/policies	Information not available/found	Information not available/found	
3.1.2.4	Federal regulations/policies	MTF also has a grant of confidentality from the U.S. Department of Justice (which is different from the usual NIH grant of confidentality)... In brief, the primary regulations are all IRB related, plus any/all applicable federal agency-specific policies (NIH data sharing, FERPA, etc). [1] Grant of confidentiality from the U.S. Department of Justice is different from a certificate of confidentiality because it comes from the authority to issue this is for drugs and other controlled substance research; the grant of confidentiality specifies that researchers cannot reidentify participants, and its purpose is to protect participant confidentiality. [2] Generally, schools must have written permission from the parent or eligible student in order to release any information from a student's education record. However, FERPA allows schools to disclose those records, without consent, to the following parties or under the following conditions (34 CFR § 99.31): -School officials with legitimate educational interest; -Other schools to which a student is transferring; -Specified officials for audit or evaluation purposes; -Appropriate parties in connection with financial aid to a student; -Organizations conducting certain studies for or on behalf of the school; -Accrediting organizations; -To comply with a judicial order or lawfully issued subpoena; -Appropriate officials in cases of health and safety emergencies; and -State and local authorities, within a juvenile justice system, pursuant to specific State law. [3]	1. Grant of Confidentality from the U.S. Department of Justice 2. Family Educational Rights and Privacy Act (FERPA)	[1] MTF Email Communication [2] U-Mich Legal Meeting [3] https://www2.ed.gov/policy/gen/guid/fpco/ferpa/index.html (Accessed: 4/25/23)

Dataset 3 - Monitoring the Future (MTF)				
		Raw Language	**Interpretation**	**Source**
3.1.2.5	International regulations/policies	N/A - Since 1975 the MTF survey has measured drug and alcohol use and related attitudes among adolescent students nationwide. [1]	N/A	[1] https://nida.nih.gov/research-topics/trends-statistics/monitoring-future (Accessed: 4/18/23)
3.1.2.6	Contractual obligations	Information not available/found	Information not available/found	
3.1.2.7	Repository policies	Information not available/found	Information not available/found	
3.2	**Governance for data linkage, sharing, access, and use based on data sharing authorization or applicable regulations/policies (i.e., the origin of the governance)**			
3.2.1	Whether the data can be linked	Does not authorize/specify	Does not authorize/specify	
3.2.2	With what other data can it be linked or can it not be linked (scope of linkage)	Does not authorize/specify	Does not authorize/specify	
3.2.3	Whether data can be shared	In lieu of these documents (RUDDDA and MoU), MTF did have IRB oversight of ICPSR taking over the dissemination of the restricted-use data; the IRB number of the MTF study is linked to ICPSR's reference number (for ICPSR's IRB)… U-Mich authorized NAHDAP to share the MTF data; IRB oversight was provided and authorization from the MTF PI. [1]	ICPSR (U-Mich) IRB and MTF (U-Mich) IRB approval specify that data can be shared through ICPSR/NAHDAP	[1] NAHDAP Meeting
3.2.4	How data can be shared (de-identification status, disclosure review)	The MTF data is fully de-identified survey data... For MTF restricted-use data (that contains state and zipcode), there are restrictions for the type of analysis at the state or lower geographic-level. [1] There is another document for MTF that is provided to data users which specifies the risk / disclosure review procedures for output data generated from the use of MTF data in the VDE – the output cannot be removed from the VDE until it has been reviewed by NAHDAP/ICPSR staff. [1] With the exception of deposits placed in openICPSR , our public data-sharing service, ICPSR reviews all datasets to assess disclosure risk. ICPSR trains data curators to apply specified procedures to protect respondent confidentiality in all of the data ICPSR curates, archives, and distributes. For example, ICPSR checks each study for identifiers present in the data. [2] FERPA states that if you have collected personally identifying information on a student, you are not allowed to share that information without consent unless you meet one of the FERPA conditions (one of those being, that the data is fully de-identified). [3]	FERPA and ICPSR/NAHDAP specify that data shared through ICPSR VDE: 1. Must be fully de-identified (for MTF restricted-use data, this does not include state and zipcode) 2. Must undergo disclosure review by ICPSR/NAHDAP staff prior to sharing, and disclosure review of analysis outputs is required prior to removing output data from the VDE	[1] NAHDAP Meeting [2] https://www.icpsr.umich.edu/web/pages/datamanagement/confidentiality/ (Accessed: 4/18/23) [3] U-Mich Legal Meeting

Dataset 3 - Monitoring the Future (MTF)				
		Raw Language	**Interpretation**	**Source**
3.2.5	How data can be accessed (access type, data use agreement, data access committee/group approval, IRB LOD, etc.)	The National Addiction & HIV Data Archive Program (NAHDAP), funded by the National Institute on Drug Abuse (NIDA), is hosted at the Inter-university Consortium for Political and Social Research (ICPSR). NAHDAP data in the ICPSR Virtual Data Enclave (VDE) are restricted from general dissemination to protect the confidentiality of the individuals and/or organizations represented in the data. To access data in the VDE, a Restricted Data Use Agreement (RDUA) must be established between the University of Michigan and the researcher's institution. [1] Please note that many of NAHDAP's studies contain confidential data and are not available for Web download. Instructions on how to obtain those datasets are provided at the top of the study home page under "Access Notes." You will be required to provide Institutional Review Board (IRB) approval or exemption for your project. [2] Yes, IRB approval is required to access the MTF data. [3] NADHAP staff will review all requests to access data, including review of the research proposal: - Research proposals are abstract length and include research questions, why they need access to the data, specifically, why they need access to the restricted-use data - NAHDAP staff review these requests for the ability to use the data to answer the proposed research questions - For MTF restricted-use data (that contains state and zipcode), there are restrictions for the type of analysis at the state or lower geographic-level There is no data use committee specific to MTF that reviews, but ICPSR has a data stewardship committee to assist with reviewing requests if NADHAP staff need advice. This committee does not typically review – usually NAHDAP staff will reach out to the Directors of NAHDAP (Joy/Amy), and as needed, the MTF Team to address any questions. [3]	ICPSR/NAHDAP specifies that for data access, a user: 1. Must execute NAHDAP VDE RDUA between ICPSR (U-Mich) and the researcher's institution 2. Must only access data through the ICPSR VDE (virtual enclave) 3. Must obtain IRB approval or exemption from the researcher's institution 4. Must obtain review and approval from NAHDAP on the proposed research	[1] https://www.icpsr.umich.edu/web/pages/NAHDAP/vde/index.html (Accessed: 4/18/23) [2] https://www.icpsr.umich.edu/web/pages/NAHDAP/data/index.html (Accessed: 4/18/23) [3] NAHDAP Meeting
3.2.6	How data can be used (data use limitations)	The MTF data can be used for general research which entails broad research, NAHDAP/ICPSR reviews the proposed research (but not for scientific merit). [1]	NAHDAP/ICPSR specifies that the data can be used for broad research	[1] NAHDAP Meeting
3.2.7	Other (specify)	Information not available/found	Information not available/found	
4	**Data Access**			
4.1	**Authorizations and Applicable Regulations/Policies**			

Dataset 3 - Monitoring the Future (MTF)				
		Raw Language	**Interpretation**	**Source**
4.1.1	**Authorizations**		1. MTF PI determination (for sharing/accessing data through ICPSR VDE)	
4.1.1.1	Assent	Information not available/found	Information not available/found	
4.1.1.2	Consent	Information not available/found	Information not available/found	
4.1.1.3	IRB/equivalent Privacy Board determination	Information not available/found	Information not available/found	
4.1.1.4	Local/state/federal laws	Information not available/found	Information not available/found	
4.1.1.5	Institutional Certification	Information not available/found	Information not available/found	
4.1.1.6	Data originator agreement	Information not available/found	Information not available/found	
4.1.1.7	Repository agreements/policies	There is a NAHDAP-specific RDUA that is used for MTF, which is tailored for accessing data through the VDE; this RDUA also includes data use specifications. [1]	NAHDAP Restricted Data Use Agreement for Restricted Data in the Virtual Data Enclave (NAHDAP VDE RDUA) authorizes data access	[1] NAHDAP Meeting
4.1.1.8	Other (specify)	The determination for sharing data through the VDE is made by the NADHAP staff based on the level of disclosure risk and/or discussion with the data originator. For MTF, the PI requested that the restricted-use MTF data be shared/accessed through VDE. [1]	MTF PI authorizes data access (through the ICPSR VDE)	[1] NAHDAP Meeting
4.1.2	**Applicable Regulations/Policies**			
4.1.2.1	Local regulations/policies	Information not available/found	Information not available/found	
4.1.2.2	Tribal regulations/policies	Information not available/found	Information not available/found	
4.1.2.3	State regulations/policies	Information not available/found	Information not available/found	
4.1.2.4	Federal regulations/policies	Information not available/found	Information not available/found	
4.1.2.5	International regulations/policies	N/A - Since 1975 the MTF survey has measured drug and alcohol use and related attitudes among adolescent students nationwide. [1]	N/A	[1] https://nida.nih.gov/research-topics/trends-statistics/monitoring-future (Accessed: 4/18/23)
4.1.2.6	Contractual obligations	Information not available/found	Information not available/found	

Dataset 3 - Monitoring the Future (MTF)				
		Raw Language	**Interpretation**	**Source**
4.1.2.7	Repository policies	The National Addiction & HIV Data Archive Program (NAHDAP), funded by the National Institute on Drug Abuse (NIDA), is hosted at the Inter-university Consortium for Political and Social Research (ICPSR). NAHDAP data in the ICPSR Virtual Data Enclave (VDE) are restricted from general dissemination to protect the confidentiality of the individuals and/or organizations represented in the data. To access data in the VDE, a Restricted Data Use Agreement (RDUA) must be established between the University of Michigan and the researcher's institution. [1] Please note that many of NAHDAP's studies contain confidential data and are not available for Web download. Instructions on how to obtain those datasets are provided at the top of the study home page under "Access Notes." You will be required to provide Institutional Review Board (IRB) approval or exemption for your project. [2] Yes, IRB approval is required to access the MTF data. [3] NADHAP staff will review all requests to access data, including review of the research proposal: - Research proposals are abstract length and include research questions, why they need access to the data, specifically, why they need access to the restricted-use data - NAHDAP staff review these requests for the ability to use the data to answer the proposed research questions - For MTF restricted-use data (that contains state and zipcode), there are restrictions for the type of analysis at the state or lower geographic-level There is no data use committee specific to MTF that reviews, but ICPSR has a data stewardship committee to assist with reviewing requests if NADHAP staff need advice. This committee does not typically review – usually NAHDAP staff will reach out to the Directors of NAHDAP (Joy/Amy), and as needed, the MTF Team to address any questions. [3]	NAHDAP/ICPSR VDE policy	[1] https://www.icpsr.umich.edu/web/pages/NAHDAP/vde/index.html (Accessed: 4/18/23) [2] https://www.icpsr.umich.edu/web/pages/NAHDAP/data/index.html (Accessed: 4/18/23) [3] NAHDAP Meeting
4.2	**Governance for data linkage, sharing, access, and use based on data access authorization or applicable regulations/policies (i.e., the origin of the governance)**			
4.2.1	Whether the data can be linked	Does not authorize/specify	Does not authorize/specify	
4.2.2	With what other data can it be linked or can it not be linked (scope of linkage)	Does not authorize/specify	Does not authorize/specify	
4.2.3	Whether data can be shared	Does not authorize/specify	Does not authorize/specify	
4.2.4	How data can be shared (de-identification status, disclosure review)	There is another document for MTF that is provided to data users which specifies the risk / disclosure review procedures for output data generated from the use of MTF data in the VDE – the output cannot be removed from the VDE until it has been reviewed by NAHDAP/ICPSR staff. [1]	ICPSR/NAHDAP specifies that disclosure review of analysis outputs is required prior to removing output data from the VDE.	[1] NAHDAP Meeting

Dataset 3 - Monitoring the Future (MTF)				
		Raw Language	**Interpretation**	**Source**
4.2.5	How data can be accessed (access type, data use agreement, data access committee/group approval, IRB LOD, etc.)	The National Addiction & HIV Data Archive Program (NAHDAP), funded by the National Institute on Drug Abuse (NIDA), is hosted at the Inter-university Consortium for Political and Social Research (ICPSR). NAHDAP data in the ICPSR Virtual Data Enclave (VDE) are restricted from general dissemination to protect the confidentiality of the individuals and/or organizations represented in the data. To access data in the VDE, a Restricted Data Use Agreement (RDUA) must be established between the University of Michigan and the researcher's institution. [1] Please note that many of NAHDAP's studies contain confidential data and are not available for Web download. Instructions on how to obtain those datasets are provided at the top of the study home page under "Access Notes." You will be required to provide Institutional Review Board (IRB) approval or exemption for your project. [2] Yes, IRB approval is required to access the MTF data. [3] NADHAP staff will review all requests to access data, including review of the research proposal: - Research proposals are abstract length and include research questions, why they need access to the data, specifically, why they need access to the restricted-use data - NAHDAP staff review these requests for the ability to use the data to answer the proposed research questions - For MTF restricted-use data (that contains state and zipcode), there are restrictions for the type of analysis at the state or lower geographic-level There is no data use committee specific to MTF that reviews, but ICPSR has a data stewardship committee to assist with reviewing requests if NADHAP staff need advice. This committee does not typically review – usually NAHDAP staff will reach out to the Directors of NAHDAP (Joy/Amy), and as needed, the MTF Team to address any questions. [3]	ICPSR/NAHDAP specifies that for data accessed through the ICPSR VDE, a user: 1. Must execute NAHDAP VDE RDUA between ICPSR (U-Mich) and the researcher's institution 2. Must only access data through the ICPSR VDE (virtual enclave) 3. Must obtain IRB approval or exemption from the researcher's institution 4. Must obtain review and approval from NAHDAP on the proposed research	[1] https://www.icpsr.umich.edu/web/pages/NAHDAP/vde/index.html (Accessed: 4/18/23) [2] https://www.icpsr.umich.edu/web/pages/NAHDAP/data/index.html (Accessed: 4/18/23) [3] NAHDAP Meeting
4.2.6	How data can be used (data use limitations)	The MTF data can be used for general research which entails broad research, NAHDAP/ICPSR reviews the proposed research (but not for scientific merit). [1]	NAHDAP/ICPSR specifies that the data can be used for broad research	[1] NAHDAP Meeting
4.2.7	Other (specify)	Information not available/found	Information not available/found	
5	**Data Use**			
5.1	**Authorizations and Applicable Regulations/Policies**			
5.1.1	**Authorizations**		1. Consent from adults	
5.1.1.1	Assent	Information not available/found	Information not available/found	

Dataset 3 - Monitoring the Future (MTF)				
		Raw Language	**Interpretation**	**Source**
5.1.1.2	Consent	The consent specifies that the data can be used for broad research. [1]	Consent from adults authorizes data use	[1] U-Mich Legal Meeting
5.1.1.3	IRB/equivalent Privacy Board determination	Information not available/found	Information not available/found	
5.1.1.4	Local/state/federal laws	Information not available/found	Information not available/found	
5.1.1.5	Institutional Certification		Information not available/found	
5.1.1.6	Data originator agreement	Information not available/found	Information not available/found	
5.1.1.7	Repository agreements/policies	There is a NAHDAP-specific RDUA that is used for MTF, which is tailored for accessing data through the VDE; this RDUA also includes data use specifications. [1]	NAHDAP Restricted Data Use Agreement for Restricted Data in the Virtual Data Enclave (NAHDAP VDE RDUA) authorizes data use	[1] NAHDAP Meeting
5.1.1.8	Other (specify)	Information not available/found	Information not available/found	
5.1.2	**Applicable Regulations/Policies**		Information not available/found	
5.1.2.1	Local regulations/policies	Information not available/found	Information not available/found	
5.1.2.2	Tribal regulations/policies	Information not available/found	Information not available/found	
5.1.2.3	State regulations/policies	Information not available/found	Information not available/found	
5.1.2.4	Federal regulations/policies	MTF also has a grant of confidentality from the U.S. Department of Justice (which is different from the usual NIH grant of confidentality)... In brief, the primary regulations are all IRB related, plus any/all applicable federal agency-specific policies (NIH data sharing, FERPA, etc). [1] Grant of confidentiality from the U.S. Department of Justice is different from a certificate of confidentiality because it comes from the authority to issue this is for drugs and other controlled substance research; the grant of confidentiality specifies that researchers cannot reidentify participants, and its purpose is to protect participant confidentiality. [2]	Grant of Confidentality from the U.S. Department of Justice	[1] MTF Email Communication [2] U-Mich Legal Meeting
5.1.2.5	International regulations/policies	N/A - Since 1975 the MTF survey has measured drug and alcohol use and related attitudes among adolescent students nationwide. [1]	N/A	[1] https://nida.nih.gov/research-topics/trends-statistics/monitoring-future (Accessed: 4/18/23)
5.1.2.6	Contractual obligations		N/A	
5.1.2.7	Repository policies	There is a NAHDAP-specific RDUA that is used for MTF, which is tailored for accessing data through the VDE; this RDUA also includes data use specifications. [1]	NAHDAP/ICPSR VDE policy	[1] NAHDAP Meeting
5.2	**Governance for data linkage, sharing, access, and use based on data access authorization or applicable regulations/policies (i.e., the origin of the governance)**			
5.2.1	Whether the data can be linked	Does not authorize/specify	Does not authorize/specify	
5.2.2	With what other data can it be linked or can it not be linked (scope of linkage)	Does not authorize/specify	Does not authorize/specify	
5.2.3	Whether data can be shared	Does not authorize/specify	Does not authorize/specify	
5.2.4	How data can be shared (de-identification status, disclosure review)	Does not authorize/specify	Does not authorize/specify	
5.2.5	How data can be accessed (access type, data use agreement, data access committee/group approval, IRB LOD, etc.)	Does not authorize/specify	Does not authorize/specify	

Dataset 3 - Monitoring the Future (MTF)				
		Raw Language	**Interpretation**	**Source**
5.2.6	How data can be used (data use limitations)	The consent specifies that the data can be used for broad research. [1] NAHDAP/ICPSR RDUA: That the Confidential Data will be used solely for research or statistical purposes relative to the research project identified in the online application to obtain confidential data incorporated into this Agreement, and for no other purpose whatsoever without the prior consent of NAHDAP/ICPSR. [2]	Consent and the NAHDAP/ICPSR RDUA specify that the data be used solely for research or statistical purposes	[1] U-Mich Legal Meeting [2] https://www.icpsr.umich.edu/files/NAHDAP/NAHDAPGenericVDERDUA.pdf
5.2.7	Other (specify)	Information not available/found	Information not available/found	
6	**PII Elements**			
6.1	PII elements collected	We collect name, email, address, cell phone #. [1]	Name, email, address, and cell phone number are collected from participants	[1] MTF Email Communication
6.2	PII elements holder (i.e., party that holds the PII)	We hold the PII in Ann Arbor and we are covered by the Grant of Confidentality in regards to holding the PII vested at UM. [1]	Data collector (MTF / U-Mich, Ann Arbor).	[1] MTF Email Communication
6.3	Use of common data model, if any, for data collection	N/A - MTF does not use a common data model; however, MTF adheres to a metadata schema at the study and variable level – based on the data documentation initiative (DDI), which is used commonly across social science domain repositories. [1]	N/A	[1] NAHDAP Meeting
7	**Prior Data Linkages**			
7.1	**Dataset linked with other datasets**			
7.1.1	Name of other dataset linked to this dataset	N/A - The MTF data is fully de-identified survey data, so there have not been any data users that have linked the MTF data at the individual-level; however, NAHDAP may approve research projects that propose to link the MTF with other aggregate data. [1]	N/A	[1] NAHDAP Meeting
7.1.2	Other dataset type (clinical, EHR, survey, claims, SDOH, etc.)	N/A	N/A	N/A
7.1.3	Other dataset source(s)	N/A	N/A	N/A
7.1.4	Linking methodology (PPRL or non-PPRL); linkage technology	N/A	N/A	N/A
7.1.5	PII elements used for the linkage	N/A	N/A	N/A
7.1.6	Entity resolver (data originator or data linker or third party)	N/A	N/A	N/A
7.1.7	Party performing the linkages	N/A	N/A	N/A
7.1.8	Linkage quality assessment	N/A	N/A	N/A
7.1.9	Linked data sharing method (linkage maps or pre-linked dataset)	N/A	N/A	N/A

USE CASE 1 - GOVERNANCE INFORMATION				
Use Case 1: Effects of the COVID-19 pandemic on mental health of children. Are COVID-19 pandemic related mental health outcomes more severe for children in foster care?				
Dataset 4 - AFCARS (Foster Care Files, 2017-2020)				
	Dataset Source	Adoption and Foster Care Analysis and Reporting System (AFCARS)		
	Dataset Source Agency	Administration for Children and Families (ACF)		
	Dataset Type (Clinical, EHR, Survey, SDOH, etc.	Case-Level / Report Data		
	Information Sources	Website, interview with NDACAN Team (archive for AFCARS data)		
Dataset 4 - AFCARS				
		Raw Language	**Interpretation**	**Source**
1	**Data Collection**			
1.1	**Authorizations and Applicable Regulations/Policies**			
1.1.1	**Authorizations**		1. Social Security Act (Section 479) 2. 45 CFR § 1355.40 D46	
1.1.1.1	Assent	N/A	N/A	
1.1.1.2	Consent	N/A	N/A	
1.1.1.3	IRB/equivalent Privacy Board determination	N/A - No IRB approval is necessary for the states to collect information, or to forward the data to the Children's Bureau. [1]	N/A	[1] NDACAN Team Meeting
1.1.1.4	Local/state/federal law	The Adoption and Foster Care Analysis and Reporting System (AFCARS) is a federally mandated data collection system intended to provide case specific information on all children covered by the protections of Title IV-B/E of the Social Security Act (Section 427). Under the final AFCARS' rule, states are required to collect data on all adopted children who are placed by the state's child welfare agency or by private agencies under contract with the public child welfare agency. States are encouraged to report other private adoptions not involving the public welfare agency that are finalized in the state as well. In addition, states are required to collect data on all children in foster care for whom the state child welfare agency has responsibility for placement, care, or supervision. [1] AFCARS is a data collection system for national adoption and foster care data authorized under section 479 of the Social Security Act (the Act). Section 479(c)(3)(A) of the Act requires the collection of comprehensive national information with respect to the demographic characteristics of children in foster care and those who are adopted with state involvement and their biological, foster, and adoptive parents. Section 474(f) of the Act requires HHS to impose penalties for non-compliant AFCARS data. Section 1102 of the Act instructs the Secretary to promulgate regulations necessary for the effective administration of the functions for which HHS is responsible under the Act. [2] Social Security Act (Section 479) authorizes data collection. Renewals and extensions of AFCARS data collection are published in the Federal Register, and can affect Tribal areas (e.g. [FR-2023-05427 Notice AFCARS data collection renewal for 3 years] "State and tribal title IV–E agencies are	Social Security Act (Section 479) authorizes data collection by states and Tribal agencies	[1] https://www.ndacan.acf.hhs.gov/datasets/dataset-details.cfm?ID=255 (Accessed: 4/19/23) [2] https://www.federalregister.gov/documents/2020/05/12/2020-09817/adoption-and-foster-care-analysis-and-reporting-system (Accessed: 4/19/23) [3] NDACAN Email Communication [4] https://www.acf.hhs.gov/cb/fact-sheet/about-afcars

Dataset 4 - AFCARS				
		Raw Language	**Interpretation**	**Source**
		required to report AFCARS case-level information on all children in foster care and children who have been adopted or placed in a guardianship with title IV–E agency involvement.") [3 & 4] [3] Comment on data collection at the source: "Information in case records are kept confidential, but because the child is in the placement and care of the title IV-E agency, information on the child's case must be disclosed to courts and providers under specific circumstances, to assist the child and family." [3]		
1.1.1.5	Institutional Certification	N/A	N/A	
1.1.1.6	Data originator agreement	N/A	N/A	
1.1.1.7	Repository agreements/policies	Information not available/found	Information not available/found	
1.1.1.8	Other (specify)	Information not available/found	Information not available/found	
1.1.2	**Applicable Regulations/Policies**			
1.1.2.1	Local regulations/policies	Information not available/found	Information not available/found	
1.1.2.2	Tribal regulations/policies	Information not available/found	Information not available/found	
1.1.2.3	State regulations/policies	Information not available/found	Information not available/found	
1.1.2.4	Federal regulations/policies	AFCARS data collection is done by the states in accordance with 45 CFR 1355.40, which defines the policies around data collection and provision of data to the AFCARS system. Each state forwards their data to the Children's Bureau, where the data are gathered and tested for completeness, consistency, and validity. Once a year, the data are compiled into an annual file and forwarded to NDACAN for archiving and distribution to qualified applicants. No IRB approval is necessary for the states to collect information, or to forward the data to the Children's Bureau. [1]	45 CFR 1355.40 D46	[1] NDACAN Team Meeting [2] https://www.ndacan.acf.hhs.gov/datasets/dataset-details.cfm?ID=255 (Accessed: 4/19/23)
1.1.2.5	International regulations/policies	N/A	N/A	
1.1.2.6	Contractual obligations	N/A	N/A	
1.1.2.7	Repository policies	N/A	N/A	
1.2	**Governance for data linkage, sharing, access, and use based on data collection authorization or applicable regulations/policies (i.e., the origin of the governance)**			
1.2.1	Whether the data can be linked	Does not authorize/specify - linkage (with external data) is not specified in any authorization. [1]	Does not authorize/specify	[1] NDACAN Team Meeting
1.2.2	With what other data can it be linked or can it not be linked (scope of linkage)	Does not authorize/specify - linkage (with external data) is not specified in any authorization. [1]	Does not authorize/specify	[1] NDACAN Team Meeting
1.2.3	Whether data can be shared	Information not available/found	Information not available/found	
1.2.4	How data can be shared (de-identification status, disclosure review)	Information not available/found	Information not available/found	
1.2.5	How data can be accessed (access type, data use agreement, data access committee/group approval, IRB LOD, etc.)	Information not available/found	Information not available/found	
1.2.6	How data can be used (data use limitations)	Information not available/found	Information not available/found	
1.2.7	Other (specify)	Information not available/found	Information not available/found	
2	**Data Linkage**			
2.1	**Authorizations and Applicable Regulations/Policies**			
2.1.1	**Authorizations**			
2.1.1.1	Assent	N/A	N/A	
2.1.1.2	Consent	N/A	N/A	

Dataset 4 - AFCARS				
		Raw Language	**Interpretation**	**Source**
2.1.1.3	IRB/equivalent Privacy Board determination	N/A	N/A	
2.1.1.4	Local/state/federal laws	N/A	N/A	
2.1.1.5	Institutional Certification	N/A	N/A	
2.1.1.6	Data originator agreement	N/A	N/A	
2.1.1.7	Repository agreements/policies	N/A	N/A	
2.1.1.8	Other (specify)	Does not authorize/specify - NDACAN currently only approves data linkages that are at the aggregate-level (as per NDACAN Director). If NDACAN were to receive a request to link external data with AFCARS the individual-level, the Children's Bureau would likely also need to provide input / decision on whether that linkage is an appropriate use of the data, in addition to NDACAN Director. At this stage, requests to link AFCARS data at the individual-level with external data have not been received. NDACAN does not store linkable IDs (or any other PII/PHI) per their ATO. The states collect the child welfare data and produce abstracts for NDACAN; states hold the PII/PHI and must encrypt the original participant IDs prior to sending the data to NDACAN. NDACAN receives basic demographic information from the states (DOB, Race, Ethnicity, Sex) – NDACAN does not receive: SSN, Telephone #, Address, or geographical information below the county-level [1]	Does not authorize/specify	[1] NDACAN Team Meeting Summary
2.1.2	**Applicable Regulations/Policies**			
2.1.2.1	Local regulations/policies	Information not available/found	Information not available/found	
2.1.2.2	Tribal regulations/policies	N/A - There is no information about tribes or tribal data within AFCARS, beyond ethnoracial identity of Native American. [1] Social Security Act (Section 479) authorizes data collection. Renewals and extensions of AFCARS data collection are published in the Federal Register, and can affect Tribal areas (e.g. [FR-2023-05427 Notice AFCARS data collection renewal for 3 years] "State and tribal title IV–E agencies are required to report AFCARS case-level information on all children in foster care and children who have been adopted or placed in a guardianship with title IV–E agency involvement.") [2]	N/A	[1] NDACAN Team Meeting Summary [2] NDACAN Email Communication
2.1.2.3	State regulations/policies	Information not available/found	Information not available/found	
2.1.2.4	Federal regulations/policies	Information not available/found	Information not available/found	
2.1.2.5	International regulations/policies	N/A	N/A	
2.1.2.6	Contractual obligations	N/A	N/A	
2.1.2.7	Repository policies	N/A	N/A	
2.2	**Governance for data linkage, sharing, access, and use based on data linkage authorization or applicable regulations/policies (i.e., the origin of the governance)**			
2.2.1	Whether the data can be linked	Information not available/found	Information not available/found	
2.2.2	With what other data can it be linked or can it not be linked (scope of linkage)	Information not available/found	Information not available/found	
2.2.3	Whether data can be shared	Information not available/found	Information not available/found	
2.2.4	How data can be shared (de-identification status, disclosure review)	Information not available/found	Information not available/found	
2.2.5	How data can be accessed (access type, data use agreement, data access committee/group approval, IRB LOD, etc.)	Information not available/found	Information not available/found	
2.2.6	How data can be used (data use limitations)	Information not available/found	Information not available/found	

Dataset 4 - AFCARS				
		Raw Language	**Interpretation**	**Source**
2.2.7	Other (specify)	Information not available/found	Information not available/found	
3	**Data Sharing**			
3.1	**Authorizations and Applicable Regulations/Policies**			
3.1.1	**Authorizations**		1. Contractual agreement between NDACAN and Children's Bureau	
3.1.1.1	Assent	N/A	N/A	
3.1.1.2	Consent	N/A	N/A	
3.1.1.3	IRB/equivalent Privacy Board determination	N/A - No IRB approval is necessary for the states to collect information, or to forward the data to the Children's Bureau. [1]	N/A	[1] NDACAN Team Meeting
3.1.1.4	Local/state/federal laws	Information not available/found	Information not available/found	
3.1.1.5	Institutional Certification	N/A	N/A	N/A
3.1.1.6	Data originator agreement	N/A	N/A	N/A
3.1.1.7	Repository agreements/policies	N/A - NDACAN Contributor Agreement is only required for the deposition of external data to NDACAN, and does not apply to the AFCARS data. [1]	N/A	[1] NDACAN Team Meeting
3.1.1.8	Other	The initial AFCARS program was authorized in the 1990s, NDACAN had established a memorandum of understanding (MoU) between NDACAN (archive) and the data depositor (Children's Bureau) while NDACAN was under a cooperative agreement. NDACAN is currently authorized by contract with the Children's Bureau to receive AFCARS data, so the MoU is no longer applicable. [1]	Contractual agreement between NDACAN and Children's Bureau authorizes data sharing	[1] NDACAN Team Meeting
3.1.2	**Applicable Regulations/Policies**			
3.1.2.1	Local regulations/policies	Information not available/found	Information not available/found	
3.1.2.2	Tribal regulations/policies	N/A - There is no information about tribes or tribal data within AFCARS, beyond ethnoracial identity of Native American. [1]	N/A	[1] NDACAN Team Meeting
3.1.2.3	State regulations/policies	Information not available/found	Information not available/found	
3.1.2.4	Federal regulations/policies	Information not available/found	Information not available/found	
3.1.2.5	International regulations/policies	N/A	N/A	
3.1.2.6	Contractual obligations	Information not available/found	Information not available/found	
3.1.2.7	Repository policies	Information not available/found	Information not available/found	
3.2	**Governance for data linkage, sharing, access, and use based on data sharing authorization or applicable regulations/policies (i.e., the origin of the governance)**			
3.2.1	Whether the data can be linked	Does not authorize/specify - linkage (with external data) is not specified in any authorization. [1]	Does not authorize/specify	[1] NDACAN Team Meeting
3.2.2	With what other data can it be linked or can it not be linked (scope of linkage)	Does not authorize/specify - linkage (with external data) is not specified in any authorization. [1]	Does not authorize/specify	[1] NDACAN Team Meeting
3.2.3	Whether data can be shared	The initial AFCARS program was authorized in the 1990s, NDACAN had established a memorandum of understanding (MoU) between NDACAN (archive) and the data depositor (Children's Bureau) while NDACAN was under a cooperative agreement. NDACAN is currently authorized by contract with the Children's Bureau to receive AFCARS data, so the MoU is no longer applicable. [1]	Contractual agreement between NDACAN and Children's Bureau specifies that AFCARS data can be shared through NDACAN.	[1] NDACAN Team Meeting

Dataset 4 - AFCARS				
		Raw Language	**Interpretation**	**Source**
3.2.4	How data can be shared (de-identification status, disclosure review)	AFCARS comes to NDACAN with no PII whatsoever, with none of the other pieces of data listed above – except maybe dates. Dates are submitted with AFCARS (such as birth date or date of removal) but these are either masked or permuted so that original dates are not included in the distributed files. In short, nothing on the list above (except the counties) is included in the data users' receive [1] NDACAN does not store linkable IDs (or any other PII/PHI) per their ATO; linkage would not be possible using AFCARS data stored in NDACAN. [1] Before distributing the AFCARS data, NDACAN makes certain manipulations to the foster care data to protect the privacy of the children in foster care. - The county FIPS code for the children from counties with fewer than 1,000 records in the annual database are recoded to indicate not provided for reasons of confidentiality. - The child's day of birth (DOB) is recoded to the 15th of the month. NOTE: All derived age variables are based on the actual DOB, so may not agree with an age computed from the recoded DOB. - All other dates in the file are adjusted to the recoded date of birth so that the span of time between any two dates is preserved. As a result, all dates in the file are recoded, but all time spans are accurate. [2] Confidentiality Protections in the NDACAN Administrative Datasets. [3]	Contractual agreement between NDACAN and Children's Bureau specifies that AFCARS data shared through NDACAN: 1. Must be de-identified of all 18 HIPAA identifiers 2. Must undergo disclosure review by NDACAN staff will conduct disclosure review of data prior to sharing (removing county FIPS code with >1,000 records, recode DoB to the 15th and adjust all other dates accordingly)	[1] NDACAN Team Meeting [2] https://www.ndacan.acf.hhs.gov/datasets/pdfs_user_guides/afcars-userguide-2000-present.pdf (Accessed: 4/19/23) [3] Confidentiality Protections document from NDACAN Team
3.2.5	How data can be accessed (access type, data use agreement, data access committee/group approval, IRB LOD, etc.)	Information not available/found	Information not available/found	
3.2.6	How data can be used (data use limitations)	Does not authorize/specify	Does not authorize/specify	
3.2.7	Other (specify)	Information not available/found	Information not available/found	
4	**Data Access**			
4.1	**Authorizations and Applicable Regulations/Policies**			
4.1.1	**Authorizations**		1. Contractual agreement between Children's Bureau and NDACAN 2. NDACAN Terms of Use Agreement	
4.1.1.1	Assent	N/A	N/A	
4.1.1.2	Consent	N/A	N/A	
4.1.1.3	IRB/equivalent Privacy Board determination	N/A - No (IRB approval is not required to access AFCARS data), but researchers who work at a college or university are required to report their research to their own IRB. [1]	N/A	[1] NDACAN Team Meeting
4.1.1.4	Local/state/federal laws	Information not available/found	Information not available/found	
4.1.1.5	Institutional Certification	N/A	N/A	
4.1.1.6	Data originator agreement	N/A	N/A	

Dataset 4 - AFCARS				
		Raw Language	**Interpretation**	**Source**
4.1.1.7	Repository agreements/policies	NDACAN is authorized by the Children's Bureau to provide access to AFCARS to individual researchers who provide their contact information and submit a Terms of Use Agreement. [1] The Children's Bureau authorizes NDACAN to share the data with researchers (under contractual agreement). The Terms of Use Agreement specifies how the data can be used by researchers - the origin of this authorization / terms of use agreement may need to be traced back to when NDACAN was established. [1]	1. Contractual agreement between Children's Bureau and NDACAN 2. NDACAN Terms of Use Agreement	[1] NDACAN Team Meeting
4.1.1.8	Other (specify)	NDACAN is authorized by the Children's Bureau to provide access to AFCARS to individual researchers who provide their contact information and submit a Terms of Use Agreement. [1] The Children's Bureau authorizes NDACAN to share the data with researchers (under contractual agreement). The Terms of Use Agreement specifies how the data can be used by researchers - the origin of this authorization / terms of use agreement may need to be traced back to when NDACAN was established. [1]	Two repository agreements authorize data access: 1. Contractual agreement between Children's Bureau and NDACAN 2. NDACAN Terms of Use Agreement	[1] NDACAN Team Meeting
4.1.2	**Applicable Regulations/Policies**			
4.1.2.1	Local regulations/policies	Information not available/found	Information not available/found	
4.1.2.2	Tribal regulations/policies	N/A - There is no information about tribes or tribal data within AFCARS, beyond ethnoracial identity of Native American. [1]	N/A	[1] NDACAN Team Meeting
4.1.2.3	State regulations/policies	Information not available/found	Information not available/found	
4.1.2.4	Federal regulations/policies	Information not available/found	Information not available/found	
4.1.2.5	International regulations/policies	N/A	N/A	
4.1.2.6	Contractual obligations	Information not available/found	Information not available/found	

Dataset 4 - AFCARS				
		Raw Language	**Interpretation**	**Source**
4.1.2.7	Repository policies	NDACAN provides individuals with authorization to use the AFCARS datasets (as formatted and provided by NDACAN) if they submit a Terms of Use Agreement. NDACAN reviews and approves this document, which contains a description of their research purpose and their affirmation to limit their use of the data to research and not re-sale, attempts to identify individuals, or other activities defined as misuse in the Agreement and by U.S. laws. Please refer to the Terms of Use Agreement PDF on this web page: https://www.ndacan.acf.hhs.gov/datasets/request-dataset.cfm [1] The Children's Bureau authorizes NDACAN to share the data with researchers (under contractual agreement). The Terms of Use Agreement specifies how the data can be used by researchers - the origin of this authorization / terms of use agreement may need to be traced back to when NDACAN was established. [1]	NDACAN policy	[1] NDACAN Team Meeting
4.2	**Governance for data linkage, sharing, access, and use based on data access authorization or applicable regulations/policies (i.e., the origin of the governance)**			
4.2.1	Whether the data can be linked	Does not authorize/specify - linkage (with external data) is not specified in any authorization. [1]	Does not authorize/specify	[1] NDACAN Team Meeting
4.2.2	With what other data can it be linked or can it not be linked (scope of linkage)	Does not authorize/specify - linkage (with external data) is not specified in any authorization. [1]	Does not authorize/specify	[1] NDACAN Team Meeting

Dataset 4 - AFCARS				
		Raw Language	**Interpretation**	**Source**
4.2.3	Whether data can be shared	The initial AFCARS program was authorized in the 1990s, NDACAN had established a memorandum of understanding (MoU) between NDACAN (archive) and the data depositor (Children's Bureau) while NDACAN was under a cooperative agreement. NDACAN is currently authorized by contract with the Children's Bureau to receive AFCARS data, so the MoU is no longer applicable. [1]	Contractual agreement between NDACAN and Children's Bureau specifies that AFCARS data can be shared through NDACAN	[1] NDACAN Team Meeting
4.2.4	How data can be shared (de-identification status, disclosure review)	AFCARS comes to NDACAN with no PII whatsoever, with none of the other pieces of data listed above – except maybe dates. Dates are submitted with AFCARS (such as birth date or date of removal) but these are either masked or permuted so that original dates are not included in the distributed files. In short, nothing on the list above (except the counties) is included in the data users' receive. [1] NDACAN does not store linkable IDs (or any other PII/PHI) per their ATO; linkage would not be possible using AFCARS data stored in NDACAN. [1] Before distributing the AFCARS data, NDACAN makes certain manipulations to the foster care data to protect the privacy of the children in foster care. - The county FIPS code for the children from counties with fewer than 1,000 records in the annual database are recoded to indicate not provided for reasons of confidentiality. - The child's day of birth (DOB) is recoded to the 15th of the month. NOTE: All derived age variables are based on the actual DOB, so may not agree with an age computed from the recoded DOB. - All other dates in the file are adjusted to the recoded date of birth so that the span of time between any two dates is preserved. As a result, all dates in the file are recoded, but all time spans are accurate. [2] Confidentiality Protections in the NDACAN Administrative Datasets. [3]	Contractual agreement between NDACAN and Children's Bureau specifies that AFCARS data shared through NDACAN: 1. Data must be de-identified of all 18 HIPAA identifiers 2. Data must undergo disclosure review of data prior to sharing (removing county FIPS code with >1,000 records, recode DoB to the 15th and adjust all other dates accordingly)	[1] NDACAN Team Meeting [2] https://www.ndacan.acf.hhs.gov/datasets/pdfs_user_guides/afcars-userguide-2000-present.pdf (Accessed: 4/19/23) [3] Confidentiality Protections document from NDACAN Team

Dataset 4 - AFCARS				
		Raw Language	**Interpretation**	**Source**
4.2.5	How data can be accessed (access type, data use agreement, data access committee/group approval, IRB LOD, etc.)	NDACAN may request additional information or deny a request for a dataset if the research purpose written on their Terms of Use Agreement is unclear, indicates a situation where additional researchers on a team need to submit their own Terms of Use Agreements, or indicates an inappropriate use of the data. Such instances are forwarded to the NDACAN Director for review. [1] NDACAN staff review all requests for AFCARS data; occasionally, the Children's Bureau CORs will review special requests, but they are typically not involved. [1] No technical controls, except for a 10-day download window. Individuals who request datasets from NDACAN supply a research purpose which is checked by the NDACAN Archiving Assistant to prevent non-approved usage. [1]	Contractural agreement between NDACAN and Children's Bureau specifies that for data accessed: 1. User must execute of the NDACAN Terms of Use Agreement 2. User must obtain review and approval from NDACAN staff on the proposed research	[1] NDACAN Team Meeting
4.2.6	How data can be used (data use limitations)	The User agrees to use the Research Data only for purposes that support the User's task defined in Section I.1 above. The User also agrees to ensure the integrity, security, and confidentiality of the Research Data by complying with the terms of this Agreement and applicable law, including the Privacy Act of 1974 (5 U.S.C. 552a). [1]	NDACAN Terms of Use Agreement specifies that the data can be used by researchers in accordance with their approved research described in Section I.1 of the Terms of Use Agreement	[1] https://www.ndacan.acf.hhs.gov/datasets/order_forms/termsofuseagreement.pdf (Accessed: 4/19/23)
4.2.7	Other (specify)	Information not available/found	Information not available/found	
5	**Data Use**			
5.1	**Authorizations and Applicable Regulations/Policies**			
5.1.1	**Authorizations**			
5.1.1.1	Assent	N/A	N/A	
5.1.1.2	Consent	N/A	N/A	
5.1.1.3	IRB/equivalent Privacy Board determination	Information not available/found - the authorizations for data use may need to be traced back to when NDACAN was established. [1]	N/A	[1] NDACAN Team Meeting
5.1.1.4	Local/state/federal laws	Information not available/found	Information not available/found	
5.1.1.5	Institutional Certification	Information not available/found	N/A	
5.1.1.6	Data originator agreement	Information not available/found	N/A	

Dataset 4 - AFCARS				
		Raw Language	**Interpretation**	**Source**
5.1.1.7	Repository agreements/policies	The Children's Bureau authorizes NDACAN to share the data with researchers (under contractual agreement). The Terms of Use Agreement specifies how the data can be used by researchers - the origin of this authorization / terms of use agreement may need to be traced back to when NDACAN was established. [1]	NDACAN Terms of Use Agreement authorizes data use	[1] NDACAN Team Meeting
5.1.1.8	Other (specify)	Information not available/found	Information not available/found	
5.1.2	**Applicable Regulations/Policies**			
5.1.2.1	Local regulations/policies	Information not available/found	Information not available/found	
5.1.2.2	Tribal regulations/policies	N/A - There is no information about tribes or tribal data within AFCARS, beyond ethnoracial identity of Native American. [1]	N/A	[1] NDACAN Team Meeting
5.1.2.3	State regulations/policies	Information not available/found	Information not available/found	
5.1.2.4	Federal regulations/policies	Information not available/found	Information not available/found	
5.1.2.5	International regulations/policies	N/A	N/A	
5.1.2.6	Contractual obligations	Information not available/found	Information not available/found	
5.1.2.7	Repository policies	NDACAN provides individuals with authorization to use the AFCARS datasets (as formatted and provided by NDACAN) if they submit a Terms of Use Agreement. NDACAN reviews and approves this document, which contains a description of their research purpose and their affirmation to limit their use of the data to research and not re-sale, attempts to identify individuals, or other activities defined as misuse in the Agreement and by U.S. laws. Please refer to the Terms of Use Agreement PDF on this web page: https://www.ndacan.acf.hhs.gov/datasets/request-dataset.cfm [1] The Children's Bureau authorizes NDACAN to share the data with researchers (under contractual agreement). The Terms of Use Agreement specifies how the data can be used by researchers - the origin of this authorization / terms of use agreement may need to be traced back to when NDACAN was established. [1]	NDACAN policy	[1] NDACAN Team Meeting
5.2	**Governance for data linkage, sharing, access, and use based on data access authorization or applicable regulations/policies (i.e., the origin of the governance)**			

Dataset 4 - AFCARS				
		Raw Language	**Interpretation**	**Source**
5.2.1	Whether the data can be linked	Does not authorize/specify - linkage (with external data) is not specified in any authorization. [1]	Does not authorize/specify	[1] NDACAN Team Meeting
5.2.2	With what other data can it be linked or can it not be linked (scope of linkage)	Does not authorize/specify - linkage (with external data) is not specified in any authorization. [1]	Does not authorize/specify	[1] NDACAN Team Meeting
5.2.3	Whether data can be shared	Information not available/found	Information not available/found	
5.2.4	How data can be shared (de-identification status, disclosure review)	Information not available/found	Information not available/found	
5.2.5	How data can be accessed (access type, data use agreement, data access committee/group approval, IRB LOD, etc.)	Information not available/found	Information not available/found	
5.2.6	How data can be used (data use limitations)	The User agrees to use the Research Data only for purposes that support the User's task defined in Section I.1 above. The User also agrees to ensure the integrity, security, and confidentiality of the Research Data by complying with the terms of this Agreement and applicable law, including the Privacy Act of 1974 (5 U.S.C. 552a). [1]	NDACAN Terms of Use Agreement specifies that the data can be used by researchers in accordance with their approved research described in Section I.1 of the Terms of Use Agreement	[1] https://www.ndacan.acf.hhs.gov/datasets/order_forms/termsofuseagreement.pdf (Accessed: 4/19/23)
5.2.7	Other (specify)	Information not available/found	Information not available/found	
6	**PII Elements**			

Dataset 4 - AFCARS				
		Raw Language	**Interpretation**	**Source**
6.1	PII elements collected	NDACAN receives basic demographic information from the states (DOB, Race, Ethnicity, Sex) – NDACAN does not receive original participant IDs, SSN, Telephone #, Address, or geographical information (below the county-level), which are PII held by the states. [1] The AFCARS datasets made available to researchers by the Children's Bureau, via NDACAN, contain no PII. [2]	Participant ID, SSN, telephone number, address, and geographical location are collected by states but not held as part of the AFCARS data	[1] NDACAN Team Meeting [2] AFCARS Email Communication
6.2	PII elements holder (i.e., party that holds the PII)	Individual states collect child welfare data and produce abstracts for NDACAN; the states hold the PII/PHI and must encrypt the original participant IDs prior to sending the data to NDACAN. [1]	Data originator (individual states)	[1] NDACAN Team Meeting
6.3	Use of common data model, if any, for data collection	N/A - Not that staff is aware of. [1]	N/A	[1] NDACAN Team Meeting
7	**Prior Data Linkages**			
7.1	**Dataset linked with other datasets**			

Dataset 4 - AFCARS				
		Raw Language	**Interpretation**	**Source**
7.1.1	Name of other linked dataset	As commented, AFCARS can only be linked by individual IDs with NCANDS and NYTD (both housed in NDACAN) and Family First Prevention Services Act data (not housed at NDACAN) – which are all systems under the Children's Bureau and share a common "AFCARS" ID to allow for linkage. So if those count as outside sources then yes. Otherwise AFCARS can't be linked at an individual level with any other data. [1] *FPSA removed from list of previously linked datasets by NDACAN staff in email communication The National Youth in Transition Database (NYTD) collects information on youth in foster care, including sex, race, ethnicity, date of birth, and foster care status. It also collects information about the outcomes of those youth who have aged out of foster care. States began collecting data in 2010, and the first data set was submitted in May 2011. [2]	AFCARS data has been linked with the following datasets within NDACAN: 1. National Child Abuse and Neglect Data System (NCANDS) 2. National Youth in Transition Database (NYTD)	[1] NDACAN Team Meeting [2] https://www.acf.hhs.gov/cb/research-data-technology/reporting-systems/nytd (Accessed: 4/19/23)
7.1.2	Other dataset type (clinical, EHR, survey, claims, SDOH, etc.)	Every year, NCANDS data are submitted voluntarily by the 50 states, the District of Columbia, and the Commonwealth of Puerto Rico. The NCANDS reporting year is based on the FFY calendar which spans October 1 to September 30. States submit case-level data, called a Child File, by constructing an electronic file of child-specific records for each report of alleged child abuse and neglect that received a CPS response in the form of an investigation or alternative response. Case-level data include information about the characteristics of the reports of abuse and neglect, the children involved, the types of maltreatment, the CPS findings, the risk factors of the child and the caregivers, the services provided, and the perpetrators. [1] To meet the law's mandate, ACF published a proposed rule in the Federal Register on July 14, 2006, and a final rule on February 26, 2008. The regulation establishes the National Youth in Transition Database (NYTD) and requires that States engage in two data collection activities. First, States are to collect information on each youth who receives independent living services paid for or provided by the State agency that administers the Chafee Program. Second, States are to collect demographic and outcome information on certain youth in foster care whom the State will follow over time to collect additional outcome information. This information will allow ACF to track which independent living services States provide and assess the collective outcomes of youth. [2] As jurisdictions consider how to provide Title IV-E prevention services to children and families who are served in the community and are otherwise unknown to the Title IV-E agency (also known as a Community Pathway),	Case-level / report data	[1] https://www.acf.hhs.gov/cb/fact-sheet/about-ncands (Accessed: 4/25/23) [2] https://www.acf.hhs.gov/cb/fact-sheet/about-nytd (Accessed: 4/25/23) [3] https://campaigns.pcgus.com/human-services/family-first/documents/Family%20First%20Prevention%20Services%20Act_Community%20Pathways%20and%20Data%20Reporting.pdf (Accessed: 4/25/23)

Dataset 4 - AFCARS				
		Raw Language	**Interpretation**	**Source**
		they must develop processes to collect and report data that is required by the federal Administration for Children and Families (ACF). The Family First Prevention Services Act (FFPSA) data elements to be reported to ACF are defined in Revised Technical Bulletin #1, and one of these elements is a unique child identifier (current or future AFCARS record number). [3] *FPSA removed from list of previously linked datasets by NDACAN staff in email communication		
7.1.3	Other dataset source(s)	As commented, AFCARS can only be linked by individual IDs with NCANDS and NYTD (both housed in NDACAN) and Family First Prevention Services Act (FFPSA) data (not housed at NDACAN) – which are all systems under the Children's Bureau and share a common "AFCARS" ID to allow for linkage. So if those count as outside sources then yes. Otherwise AFCARS can't be linked at an individual level with any other data. [1] *FPSA removed from list of previously linked datasets by NDACAN staff in email communication	Sources of data for AFCARS linkage: NCANDS and NYTD which are housed in NDACAN	
7.1.4	Linking methodology (PPRL or non-PPRL); linkage technology	Information not available/found	Information not available/found	
7.1.5	PII elements used for the linkage	Information not available/found	Information not available/found	
7.1.6	Entity resolver (data originator or data linker or third party)	Information not available/found	Information not available/found	
7.1.7	Party performing the linkages	Information not available/found	Information not available/found	
7.1.8	Linkage quality assessment	Information not available/found	Information not available/found	
7.1.9	Linked data sharing method (linkage maps or pre-linked dataset)	Information not available/found	Information not available/found	

TABLE 1: USE CASE 2 - DATASET GOVERNANCE

[Legend: Blank cell in Table 1 = information not available/found; N/A = information confirmed to not exist]

	Governance	Dataset 1 - NCCR					Dataset 2 - CDC COVID					Dataset 3 - T-MSIS				
		Data Collection	Data Linkage	Data Sharing	Data Access	Data Use	Data Collection	Data Linkage	Data Sharing	Data Access	Data Use	Data Collection	Data Linkage	Data Sharing	Data Access	Data Use
1	**Authorization/s**															
1.1	Assent					N/A	N/A					N/A	N/A	N/A	N/A	N/A
1.2	Consent	Consent authorizes data collection for some data providers who originally obtained the data under a research consent		Consent authorizes data sharing for some data providers		N/A	N/A					N/A	N/A	N/A	N/A	N/A
1.3	IRB/equivalent Privacy Board determination	IRB approval authorizes data collection for certain states (e.g., New Jersey)	N/A	IRB approval authorizes data sharing for some data providers		IRB approval from the state registry authorizes data use (as needed/if applicable)	N/A				RIDURA specifies that for data access, a user must obtain IRB LOD from researcher's institution (as needed / applicable)	N/A	CMS Privacy Board approval of a RIF application specifying linkage authorizes data linkage			
1.4	Local/state/federal law	State regulations/policies authorizes data collection by central cancer registries as part of public health surveillance	N/A	Sections 405, 410, 412 and 413 of the Public Health Service Act (42 U.S.C. §§ 284, 285, 285a-1 and 285a-2 authorize data sharing			State laws/regulations authorizes data collection					Three Federal laws authorize data collection: 1. Section 4753 of the Balanced Budget Act of 1997 2. Section 6504 of the Patient Protection and Affordable Care Act 3. Medicaid and CHIP Managed Care Final Rule			CMS Privacy Board authorizes data access	
1.5	Institutional Certification	Institutional Certification authorizes data collection for some state or regional cancer registries	N/A				N/A					N/A	N/A			

TABLE 1: USE CASE 2 - DATASET GOVERNANCE

[Legend: Blank cell in Table 1 = information not available/found; N/A = information confirmed to not exist]

	Governance	Dataset 1 - NCCR					Dataset 2 - CDC COVID					Dataset 3 - T-MSIS				
		Data Collection	Data Linkage	Data Sharing	Data Access	Data Use	Data Collection	Data Linkage	Data Sharing	Data Access	Data Use	Data Collection	Data Linkage	Data Sharing	Data Access	Data Use
1.6	Data originator agreement	N/A	Three agreements authorize data linkage (as needed / if applicable): 1. Protocol for linkage approved by the data provider authorizes data linkage 2. Data Sharing (or Use) Agreement or a hosting agreement (ISA) or BAA executed between IMS and the data provider authorize data linkage 3. MOU or MOA executed between NCI and the data provider authorizes data linkage	Protocol for linkage approved by the data provider authorizes data sharing			N/A					N/A	N/A			
1.7	Repository agreements/policies	N/A			NCCR Data Use Agreement authorizes data access	NCCR Data Use Agreement authorizes data use	N/A			Registration Information and Data Use Restrictions Agreement (RIDURA) authorizes data access	Registration Information and Data Use Restrictions Agreement (RIDURA) authorizes data use	N/A	Information Exchange Agreement (IEA) between CMS and the Participating Agency (on top of the Data Use Agreement) authorizes data linkage for federal entities performing linkage with non-standard TAFs containing PII	Information Exchange Agreement (IEA) between CMS and the Participating Agency (on top of the Data Use Agreement) authorizes data sharing of linked data by federal agencies performing linkage with non-standard TAFs	Data Use Agreement authorizes data access	Data Use Agreement (DUA) authorizes data use
1.8	Other (specify)			Two contracts authorize data sharing: 1. Contract between SEER registries and NCI 2. Subcontract between NPCR registries and NAACCR and grant between NPCR registries and NAACCR			N/A									
2	**Applicable Regulations/Policies**															
2.1	Local regulations/policies	N/A	N/A													T-MSIS policy

TABLE 1: USE CASE 2 - DATASET GOVERNANCE

[Legend: Blank cell in Table 1 = information not available/found; N/A = information confirmed to not exist]

	Governance	Dataset 1 - NCCR					Dataset 2 - CDC COVID					Dataset 3 - T-MSIS				
		Data Collection	Data Linkage	Data Sharing	Data Access	Data Use	Data Collection	Data Linkage	Data Sharing	Data Access	Data Use	Data Collection	Data Linkage	Data Sharing	Data Access	Data Use
2.2	Tribal regulations/policies	N/A	N/A	N/A	N/A	N/A										
2.3	State regulations/policies	State regulations/policies the participating NCCR registries: Arkansas, California, Connecticut, Florida, Georgia, Hawaii, Idaho, Illinois, Iowa, Kentucky, Louisiana, Massachusetts, Michigan, New Jersey, New Mexico, New York, Ohio, Pennsylvania, Seattle (Puget Sound), Tennessee, Texas, Utah, and Wisconsin	N/A													
2.4	Federal regulations/policies	N/A	N/A									1. Privacy Act 2. HIPAA Privacy Rule		1. Privacy Act 2. HIPAA Privacy Rule	1. Privacy Act 2. HIPAA Privacy Rule 3. Common Rule	1. Privacy Act 2. HIPAA Privacy Rule
2.5	International regulations/policies	N/A	N/A	N/A	N/A		N/A	N/A	N/A							
2.6	Contractual Obligations	1. Contractual obligations for registries operated by the state health department to the NCI 2. Obligations for registries participating through a grant under NAACCR	1. Interconnection Services Agreement (as needed / if applicable) 2. Business Associate Agreement (BAA) (as needed / if applicable)													
2.7	Repository Policies	N/A	N/A		NCCR policy	NCCR policy			CDC COVID-19 case surveillance restricted access data policy	Registration Information and Data Use Restrictions Agreement (RIDURA) authorizes data access	CDC COVID-19 case surveillance restricted access data policy		CMS Research Data Center (ResDAC) policy		CMS Research Data Center (ResDAC) policy	CMS Research Data Center (ResDAC) policy
3	Data Linking/Sharing/Access/Use Governance Based on Authorizations and Applicable Regulations/Policies															

TABLE 1: USE CASE 2 - DATASET GOVERNANCE

[Legend: Blank cell in Table 1 = information not available/found; N/A = information confirmed to not exist]

	Governance	Dataset 1 - NCCR					Dataset 2 - CDC COVID					Dataset 3 - T-MSIS				
		Data Collection	**Data Linkage**	**Data Sharing**	**Data Access**	**Data Use**	**Data Collection**	**Data Linkage**	**Data Sharing**	**Data Access**	**Data Use**	**Data Collection**	**Data Linkage**	**Data Sharing**	**Data Access**	**Data Use**
3.1	Whether the data can be linked	Certain state laws/regulations specify that the data can be linked (e.g., Lousiana)	Data originator agreements specify that the data can be linked	Data originator agreements specify that the data can be linked	DUA specifies that users are not authorized to link data across SEER databases							Does not authorize/specify	1. CMS Privacy Board authorizes data linkage for researchers for research purposes via an approved RIF Application that specifies the scope of linkage 2. CMS Privacy Board, Chief Data Officer approval of a letter of justification for linkage, and Information Exchange Agreement (IEA) authorizes data linkage for federal entities performing linkage with non-standard TAFs containing PII		CMS Privacy Board authorizes data linkage for researchers for research purposes via an approved RIF Application that specifies the scope of linkage	Does not authorize/specify
3.2	With what other data can it be linked or can it not be linked (scope of linkage)	Certain state laws/regulations specify the scope of linkage (e.g., Idaho)		Data originator agreements specify that the data can be linked according to the protocol for linkage approved by the data provider	DUA specifies that users are not authorized to link data across SEER databases							Does not authorize/specify	1. CMS policy and Data Use Agreement specify that data can only be linked to external data in accordance with the RIF Application approved by the CMS Privacy Board which specifies the scope of linkage 2. Information Exchange Agreement (IEA) between CMS and Participating Agency specifies the scope of linkage for federal entities performing linkage with non-standard TAFs containing PII		CMS policy and Data Use Agreement specify that data can only be linked to external data in accordance with the RIF Application approved by the CMS Privacy Board which specifies the scope of linkage	Does not authorize/specify

TABLE 1: USE CASE 2 - DATASET GOVERNANCE

[Legend: Blank cell in Table 1 = information not available/found; N/A = information confirmed to not exist]

	Governance	Dataset 1 - NCCR					Dataset 2 - CDC COVID					Dataset 3 - T-MSIS				
		Data Collection	Data Linkage	Data Sharing	Data Access	Data Use	Data Collection	Data Linkage	Data Sharing	Data Access	Data Use	Data Collection	Data Linkage	Data Sharing	Data Access	Data Use
3.3	Whether data can be shared	Does not authorize/specify		1. Consent and IRB approval specify that data can be shared (when applicable if consent and IRB were used by data provider) 2. Sections 405, 410, 412 and 413 of the Public Health Service Act (42 U.S.C. §§ 284, 285, 285a-1 and 285a-2) specify that NCI can share data 3. Contracts between SEER registries and NCI, subcontracts between NPCR registries and NAACCR and grants between NPCR registries and NAACCR specify that data can be shared								HIPAA Privacy Rule that CMS is subject to under the Privacy Act specifies that de-identified data can be shared.	CMS policy, Data Use Agreement, and Information Exchange Agreement (IEA) authorize data sharing of linked data by federal entities but prohibit re-sharing of un-linked CMS data	CMS policy, Data Use Agreement, and Information Exchange Agreement (IEA) authorize data sharing of linked data by federal entities but prohibit re-sharing of un-linked CMS data	Does not authorize/specify	Does not authorize/specify
3.4	How data can be shared (de-identification status, disclosure review)	Does not authorize/specify	Data originator agreements specify that no geographic identifiers (except quintiles), exact dates, or exact ages will be shared	1. Contract between SEER registries and NCI, and Subcontract between NPCR registries and NAACCR, specifies that data must be de-identified for sharing through NCCR (including removal of PII, dates, and geographic information) 2. Data originator agreements specify that no geographic identifiers (except quintiles), exact dates, or exact ages will be shared					RIDURA specifies that COVID-19 data shared from various jurisdictions: 1. Must be de-identified of all direct identifiers prior to sharing 2. Must undergo disclosure review to suppress data fields with low frequency (<5) prior to sharing			HIPAA Privacy Rule that CMS is subject to under the Privacy Act specifies that T-MSIS data be de-identified of all 18 HIPAA identifiers prior to sharing.	1. CMS policy specifies that federal entities performing data linkage using non-standard TAFs containing PII must take the data into their SORN under the Privacy Act 2. CMS policy specifies that federal entities performing data linkage must agree to treat secondarily shared data as if the entity is a HIPAA Covered Entity and follow a process similar to CMS for releasing data including entering into a DUA with the researcher	1. HIPAA specifies that de-identified data can be shared 2. CMS policy specifies that federal entities performing data linkage using non-standard TAFs containing PII must take the data into their SORN under the Privacy Act 3. CMS policy specifies that federal entities performing data linkage must agree to treat secondarily shared data as if the entity is a HIPAA Covered Entity and follow a process similar to CMS for releasing data including entering into a DUA with the researcher	CMS policy specifies that data shared through the VRDC: must undergo the VRDC Review Process which is disclosure review by CCW VRDC staff of analysis outputs prior to removing output data from the VRDC	Does not authorize/specify

TABLE 1: USE CASE 2 - DATASET GOVERNANCE

[Legend: Blank cell in Table 1 = information not available/found; N/A = information confirmed to not exist]

	Governance	Dataset 1 - NCCR					Dataset 2 - CDC COVID					Dataset 3 - T-MSIS				
		Data Collection	**Data Linkage**	**Data Sharing**	**Data Access**	**Data Use**	**Data Collection**	**Data Linkage**	**Data Sharing**	**Data Access**	**Data Use**	**Data Collection**	**Data Linkage**	**Data Sharing**	**Data Access**	**Data Use**
3.5	How data can be accessed (access type, data use agreement, data access committee/group approval, IRB LOD, etc.)	Does not authorize/specify			NCCR specifies that for data access, a user: 1. Must submit a Data Analysis Plan 2. Must execute the NCCR DUA 3. Must obtain review and approval by the NCI Office of Data Sharing and the Surveillance Research Program's Data Release group on the proposed research 4. Must obtain IRB LOD from the researcher's institution, or from NCI central IRB (BRANY) if the researchers institution does not have an IRB (as needed /applicable) 5. Must obtain IRB approval from the state registry 6. Must submit using an institutional account (known as eRA Commons) and obtain verification of Signing Official by the NCI Office of Data Sharing and the Surveillance Research Program's Data Release group 7. Must access and use the data within SEER*Stat (client server software) or NCCR Data Platform	NCCR specifies that for data access, a user: 1. Must submit a Data Analysis Plan 2. Must execute the NCCR DUA 3. Must obtain review and approval by the NCI Office of Data Sharing and the Surveillance Research Program's Data Release group on the proposed research 4. Must obtain IRB LOD from the researcher's institution, or from NCI central IRB (BRANY) if the researchers institution does not have an IRB (as needed / applicable) 5. Must obtain IRB approval from the state registry 6. Must submit using an insititutional account (known as eRA Commons) and obtain verification of Signing Official by the NCI Office of Data Sharing and the Surveillance Research Program's Data Release group 7. Must access and use the data within SEER*Stat (client server software) or NCCR Data Platform				CDC COVID-19 case surveillance restricted access data policy specifies that users review the following documents to determine interest in accessing the COVID-19 Case Surveillance Restricted Access Detailed Data file: 1. CDC COVID-19 Case Surveillance Restricted Access Detailed Data: Summary, Guidance, Limitations Information, and Restricted Access Data Use Agreement Information 2. Data Dictionary for the COVID-19 Case Surveillance Restricted Access Detailed Data 3. After reviewing the above information, the user must complete the RIDURA, which specifies that for data access the user: (a). Obtain IRB LOD from researcher's institution (as needed /applicable) (b). Must access the data through GitHub private repository	RIDURA specifies that for data access, a user: 1. Must sign and complete the RIDURA 2. Obtain IRB LOD from researcher's institution (as needed / applicable) 3. Must access the data through GitHub private repository	Does not authorize/specify	Does not authorize/specify		CMS specifies that for data access, a user: 1. Must submit RIF Data Use Agreement, Attachment A: RIF Application, RIF Application Key Personnel Supplement, RIF Specifications Worksheet, and Data Management Plan Self-Attestation Questionnaire (DMP SAQ) 2. Must obtain review and approval from ResDAC team on the proposed research 3. Must obtain review and approval from CMS' Data Privacy Safeguard Program (DPSP) on the Data Management Plan Self-Attestation Questionnaire (DMP SAQ) 4. Must obtain IRB LOD from the requesting institution 5. Must obtain review and approval from CMS Privacy Board on the proposed research 6. Must access data through the VRDC or through encrypted shipped disks	Does not authorize/specify

TABLE 1: USE CASE 2 - DATASET GOVERNANCE

[Legend: Blank cell in Table 1 = information not available/found; N/A = information confirmed to not exist]

	Governance	Dataset 1 - NCCR					Dataset 2 - CDC COVID					Dataset 3 - T-MSIS				
		Data Collection	**Data Linkage**	**Data Sharing**	**Data Access**	**Data Use**	**Data Collection**	**Data Linkage**	**Data Sharing**	**Data Access**	**Data Use**	**Data Collection**	**Data Linkage**	**Data Sharing**	**Data Access**	**Data Use**
3.6	How data can be used (including data use limitations)	Does not authorize/specify			NCCR Data Use Agreement specifies that Authorized User will use or disclose the Data only for the purposes for approved research	NCCR Data Use Agreement specifies that Authorized User will use or disclose the Data only for the purposes for approved research				RIDURA specifies that the COVID-19 Case Surveillance Restricted Data Access can be used for broad research (must be of public health significance)	RIDURA specifies that the COVID-19 Case Surveillance Restricted Data Access can be used for broad research (must be of public health significance)	Does not authorize/specify	CMS Privacy Board specifies that TAFs can be used for resesarch purposes		Does not authorize/specify	1. T-MSIS policy specifies that there must be a research use to justify the initial disclosure and findings from the research must be publicly available. 2. Data Use Agreement (DUA) specifies that data will be used solely for the study described in detail in the RIF Request Application
3.7	Other (specify)	Does not authorize/specify										Does not authorize/specify	Does not authorize/specify		Does not authorize/specify	Does not authorize/specify

TABLE 2: USE CASE 2 - DATASET LINKAGE DETERMINATION

Definitions: Limitations are restrictions on data linkage and use (e.g., dataset must only be linked with other disease-relevant data, dataset must be used in a physical enclave, etc.). **Controls** are processes established to ensure compliance with governance for data sharing, access, and use (e.g., user must access data in a physical enclave, user must sign data use agreement, user must receive data access committee approval, etc.). **Authorization gaps** exist when there is no explicit authorization or the authorization is not available or found in the information collected by the project team for various data life cycle stages (data collection, linking, sharing, access and use).

		Can the datasets be linked?	What limitations do the linked datasets inherit?	What controls do the linked dataset require?	What authorization gaps exist?
1	NCCR (Dataset 1) and CDC COVID (Dataset 2) linkage	**Yes, NCCR and CDC COVID can be linked provided:** **A. SEER staff:** 1. Share NCCR data fully de-identified of all direct identifiers (including exact ages and dates and geographic information, except quintiles) [Control 1a] **B. CDC COVID staff:** 2.De-identify data of all direct identifiers [Control 2a] 3.Perform disclosure review to suppress data fields with low frequency (<5) prior to sharing CDC COVID data [Control 2b] **C. The researcher/user:** 1. Obtains authorizations for data linkage and sharing for CDC COVID [Authorization gap 1] - Assumption 2. Links NCCR data only based on applicable state laws which specify the scope of linkage (e.g., Idaho) and in accordance with protocol for linkage approved by the data provider/s (i.e., state registry) [Limitation 1a] 3. Uses or disclose linked NCCR and CDC COVID data only for the purposes for approved research which must be of public health significance [Limitations 1b, 2a] 4. Submits NCCR Data Analysis Plan, NCCR DUA, and CDC COVID Restricted Access Data Use Agreement (RIDURA) [Control 1b, 1c, 2b] 5. Obtains approval from NCI Office of Data Sharing and the Surveillance Research Program's Data Release group on the proposed linkage [Control 1d] 6. Obtains IRB LOD from the researcher's institution for NCCR and CDC COVID data (or NCI BRANY as needed for NCCR) and IRB approval from the NCCR state registry [Control 1e, 1f, 2c] 7. Uses an institutional account (known as eRA Commons) and obtains verification of Signing Official by the NCI Office of Data Sharing and the Surveillance Research Program's Data Release group [Control 1g] 8. Works with CDC COVID staff to upload data obtained from CDC GitHub private repository into SEER*Stat (client server software), where NCCR data can be accessed [Controls 1h, 2e] - Assumption for NCCR and CDC COVID	**Researchers/users:** 1a. Must only link NCCR data based on applicable state laws which specify the scope of linkage (e.g., Idaho) and in accordance with protocol for linkage approved by the data providers (i.e., state registry) [NCCR] 1b. Must use or disclose NCCR data only for the purposes for approved research [NCCR] 2a. Must use CDC COVID data for broad research (must be of public health significance) [CDC COVID]	**For sharing NCCR data, SEER staff must:** 1a. Share NCCR data fully de-identified of all direct identifiers (including exact ages and dates and geographic information, except quintiles) [NCCR] **For sharing CDC COVID data, staff must:** 2a. De-identify data of all direct identifiers [CDC COVID] 2b. Perform disclosure review to suppress data fields with low frequency (<5) prior to sharing CDC COVID data [CDC COVID] **For accessing NCCR data, researchers/users must:** 1b. Submit Data Analysis Plan [NCCR] 1c. Execute the NCCR DUA [NCCR] 1d. Obtain review by and approval from the NCI Office of Data Sharing and the Surveillance Research Program's Data Release group on the proposed research [NCCR] 1e. Obtain IRB LOD from the researcher's institution, or from NCI central IRB (BRANY) if the researchers institution does not have an IRB (as needed/applicable) [NCCR] 1f. Obtain IRB approval from the state registry [NCCR] 1g. Use an institutional account (known as eRA Commons) and obtain verification of Signing Official by the NCI Office of Data Sharing and the Surveillance Research Program's Data Release group [NCCR] 1h. Access the data from SEER*Stat (client server software) [NCCR] **For accessing CDC COVID data, researchers/users must:** 2b. Execute the Restricted Access Data Use Agreement (RIDURA) [CDC COVID] 2c. Obtain IRB LOD from researcher's institution (as needed/applicable) [CDC COVID] 2e. Access data from CDC GitHub private repository [CDC COVID]	2a. For CDC COVID-19, information on authorizations for linkage and sharing is not available/found.

TABLE 2: USE CASE 2 - DATASET LINKAGE DETERMINATION

Definitions: Limitations are restrictions on data linkage and use (e.g., dataset must only be linked with other disease-relevant data, dataset must be used in a physical enclave, etc.). **Controls** are processes established to ensure compliance with governance for data sharing, access, and use (e.g., user must access data in a physical enclave, user must sign data use agreement, user must receive data access committee approval, etc.). **Authorization gaps** exist when there is no explicit authorization or the authorization is not available or found in the information collected by the project team for various data life cycle stages (data collection, linking, sharing, access and use).

		Can the datasets be linked?	What limitations do the linked datasets inherit?	What controls do the linked dataset require?	What authorization gaps exist?
2	NCCR (Dataset 1) and T-MSIS (Dataset 3) linkage	**Yes, NCCR and T-MSIS can be linked provided:** **A. SEER staff:** 1. Share NCCR data fully de-identified of all direct identifiers (including exact ages and dates and geographic information, except quintiles) [Control 1a] **B. Federal entities who create linked T-MSIS data and share the linked data:** 2. De-identify T-MSIS data of all 18 HIPAA identifiers as per HIPAA [Control 3a] 3. Take the data into their SORN under the Privacy Act when performing data linkage using non-standard TAFs containing PII [Control 3b] 4. Treat secondarily shared data as if they are a HIPAA Covered Entity and follow a process similar to CMS for releasing data including entering into a DUA with the researcher [Control 3c] 5. Links NCCR data only based on applicable state laws for participating NCCR registries, in accordance with protocol for linkage approved by the NCCR data provider/s, the RIF Application approved by ResDAC team and CMS Privacy Board, and the Information Exchange Agreement (IEA) between CMS and Participating Agency (for federal entities performing linkage with non-standard TAFs containing PII)- all off	**Researchers/users:** 1a. Must only link NCCR data based on applicable state laws which specify the scope of linkage (e.g., Idaho) and in accordance with protocol for linkage approved by the data providers (i.e., state registry) [NCCR] 1b. Must use or disclose NCCR data only for the purposes for approved research [NCCR] **Federal entities who create linked T-MSIS data**: 3a. Must only link T-MSIS data in accordance with the RIF Application approved by the CMS Privacy Board which specifies the scope of linkage and in accordance with the Information Exchange Agreement (IEA) between CMS and Participating Agency which specifies the scope of linkage for federal entities performing linkage with non-standard TAFs containing PII [T-MSIS] **Researchers/users:** 3b. Must use T-MSIS data for research use that justifies the initial disclosure and solely for the study described in detail in the RIF Request Application, and must ensure findings are publicly available [T-MSIS]	**For sharing NCCR data, SEER staff must:** 1a. Share NCCR data fully de-identified of all direct identifiers (including exact ages and dates and geographic information, except quintiles) [NCCR] **For sharing T-MSIS data:** **A. Federal entities who create linked T-MSIS data and share the linked data must:** 3a. De-identify T-MSIS data of all 18 HIPAA identifiers as per HIPAA [T-MSIS] 3b. Take the data into their SORN under the Privacy Act when performing data linkage using non-standard TAFs containing PII [T-MSIS] 3c. Treat secondarily shared data as if they are a HIPAA Covered Entity and follow a process similar to CMS for releasing data including entering into a DUA with the researcher [T-MSIS] **B. CCW VRDC staff must:** 3d. Perform disclosure review of analysis outputs prior to sharing output data from the VRDC [T-MSIS] **For accessing NCCR data, researchers/users must:** 1b. Submit Data Analysis Plan [NCCR] 1c. Execute the NCCR DUA [NCCR] 1d. Obtain review by and approval from the NCI Office of Data Sharing and the Surveillance Research	None

TABLE 2: USE CASE 2 - DATASET LINKAGE DETERMINATION

Definitions: Limitations are restrictions on data linkage and use (e.g., dataset must only be linked with other disease-relevant data, dataset must be used in a physical enclave, etc.). **Controls** are processes established to ensure compliance with governance for data sharing, access, and use (e.g., user must access data in a physical enclave, user must sign data use agreement, user must receive data access committee approval, etc.). **Authorization gaps** exist when there is no explicit authorization or the authorization is not available or found in the information collected by the project team for various data life cycle stages (data collection, linking, sharing, access and use).

		Can the datasets be linked?	What limitations do the linked datasets inherit?	What controls do the linked dataset require?	What authorization gaps exist?
		which specify the scope of linkage [Limitations 1a, 3a] **C. CCW VRDC staff:** 6. Perform disclosure review of analysis outputs prior to sharing output data from the VRDC [Control 3d] **D. The researcher/user:** 1. Uses or discloses linked NCCR and T-MSIS data only for the purposes for approved research for NCCR and research use that justifies the initial disclosure and solely for the study described in the T-MSIS RIF Request Application, and ensures findings are publicly available [Limitations 1b, 3b] 2. Submits NCCR Data Analysis Plan, NCCR DUA, T-MSIS RIF Data Use Agreement, Attachment A: RIF Application, RIF Application Key Personnel Supplement, RIF Specifications Worksheet, and Data Management Plan Self-Attestation Questionnaire (DMP SAQ) [Controls 1b, 1c, 3e] 3. Obtains approval from NCI Office of Data Sharing and the Surveillance Research Program's Data Release group, ResDAC team, and CMS Privacy Board on the proposed linkage [Control 1d, 3f, 3g] 4. Obtains IRB LOD from the researcher's institution for NCCR and T-MSIS (or NCI BRANY as needed for NCCR) and IRB approval from the NCCR state registry [Controls 1e, 1f, 3h] 5. Uses an institutional account (known as eRA Commons) and obtains verification of Signing Official by the NCI Office of Data Sharing and the Surveillance Research Program's Data Release group for NCCR [Control 1g] 6. Works with T-MSIS and NCCR to establish federal agency authorization for T-MSIS linkage with NCCR and for use within SEER*Stat [Controls 1h, 3i] - Assumption for NCCR and T-MSIS		Program's Data Release group on the proposed research [NCCR] 1e. Obtain IRB LOD from the researcher's institution, or from NCI central IRB (BRANY) if the researchers institution does not have an IRB (as needed / applicable) [NCCR] 1f. Obtain IRB approval from the state registry [NCCR] 1g. Use an institutional account (known as eRA Commons) and obtain verification of Signing Official by the NCI Office of Data Sharing and the Surveillance Research Program's Data Release group [NCCR] 1h. Access the data from SEER*Stat (client server software) [NCCR] **For accessing T-MSIS data, researchers/users must:** 3e. Submit RIF Data Use Agreement, Attachment A: RIF Application, RIF Application Key Personnel Supplement, RIF Specifications Worksheet, and Data Management Plan Self-Attestation Questionnaire (DMP SAQ) [T-MSIS] 3f. Obtain review and approval from ResDAC team on the proposed research [T-MSIS] 3g. Obtain review and approval from CMS' Data Privacy Safeguard Program (DPSP) on the Data Management Plan Self-Attestation Questionnaire (DMP SAQ) [T-MSIS] 3h. Obtain IRB LOD from the requesting institution [T-MSIS] 3i. Access data from the VRDC or through encrypted shipped disks [T-MSIS]	

TABLE 2: USE CASE 2 - DATASET LINKAGE DETERMINATION

Definitions: Limitations are restrictions on data linkage and use (e.g., dataset must only be linked with other disease-relevant data, dataset must be used in a physical enclave, etc.). **Controls** are processes established to ensure compliance with governance for data sharing, access, and use (e.g., user must access data in a physical enclave, user must sign data use agreement, user must receive data access committee approval, etc.). **Authorization gaps** exist when there is no explicit authorization or the authorization is not available or found in the information collected by the project team for various data life cycle stages (data collection, linking, sharing, access and use).

		Can the datasets be linked?	What limitations do the linked datasets inherit?	What controls do the linked dataset require?	What authorization gaps exist?
3	CDC COVID (Dataset 2) and T-MSIS (Dataset 3) linkage	**Yes, CDC COVID and T-MSIS can be linked provided:** **A. CDC COVID staff:** 1.De-identify data of all direct identifiers [Control 2a] 2.Perform disclosure review to suppress data fields with low frequency (<5) prior to sharing CDC COVID data [Control 2b] **B. Federal entities who create linked T-MSIS data and share the linked data:** 3. De-identify T-MSIS data of all 18 HIPAA identifiers as per HIPAA [Control 3a] 4. Take the data into their SORN under the Privacy Act when performing data linkage using non-standard TAFs containing PII [Control 3b] 5. Treat secondarily shared data as if they are a HIPAA Covered Entity and follow a process similar to CMS for releasing data including entering into a DUA with the researcher [Control 3c] 6. Link T-MSIS data only in accordance with the RIF Application approved by the CMS Privacy Board which specifies the scope of linkage and in accordance with the Information Exchange Agreement (IEA) between CMS and Participating Agency which specifies the scope of linkage for federal entities performing linkage with non-standard TAFs containing PII [Limitation 3a] **C. CCW VRDC staff:**	**Researchers/users:** 2a. Must use CDC COVID data for broad research (must be of public health significance) [CDC COVID] **Federal entities who create linked T-MSIS data:** 3a. Must only link T-MSIS data in accordance with the RIF Application approved by the CMS Privacy Board which specifies the scope of linkage and in accordance with the Information Exchange Agreement (IEA) between CMS and Participating Agency which specifies the scope of linkage for federal entities performing linkage with non-standard TAFs containing PII [T-MSIS] **Researchers/users:** 3b. Must use T-MSIS data for research use that justifies the initial disclosure and solely for the study described in detail in the RIF Request Application, and must ensure findings are publicly available [T-MSIS]	**For sharing CDC COVID data, staff must:** 2a. De-identify data of all direct identifiers [CDC COVID] 2b. Perform disclosure review to suppress data fields with low frequency (<5) prior to sharing CDC COVID data [CDC COVID] **For sharing T-MSIS data:** **A. Federal entities who create linked T-MSIS data and share the linked data must:** 3a. De-identify T-MSIS data of all 18 HIPAA identifiers as per HIPAA [T-MSIS] 3b. Take the data into their SORN under the Privacy Act when performing data linkage using non-standard TAFs containing PII [T-MSIS] 3c. Treat secondarily shared data as if they are a HIPAA Covered Entity and follow a process similar to CMS for releasing data including entering into a DUA with the researcher [T-MSIS] **B. CCW VRDC staff must:** 3d. Perform disclosure review of analysis outputs prior to sharing output data from the VRDC [T-MSIS] **For accessing CDC COVID data, researchers/users must:** 2b. Execute the Restricted Access Data Use Agreement (RIDURA) [CDC COVID]	2a. For CDC COVID-19, information on authorizations for linkage and sharing is not available/found.

TABLE 2: USE CASE 2 - DATASET LINKAGE DETERMINATION

Definitions: Limitations are restrictions on data linkage and use (e.g., dataset must only be linked with other disease-relevant data, dataset must be used in a physical enclave, etc.). **Controls** are processes established to ensure compliance with governance for data sharing, access, and use (e.g., user must access data in a physical enclave, user must sign data use agreement, user must receive data access committee approval, etc.). **Authorization gaps** exist when there is no explicit authorization or the authorization is not available or found in the information collected by the project team for various data life cycle stages (data collection, linking, sharing, access and use).

		Can the datasets be linked?	What limitations do the linked datasets inherit?	What controls do the linked dataset require?	What authorization gaps exist?
		7. Perform disclosure review of analysis outputs prior to sharing output data from the VRDC [Control 3d] **D. The researcher/user:** 1. Obtains authorizations for data linkage and sharing for CDC COVID [Authorization gap 1] - **Assumption** 2. Uses the linked CDC COVID data and T-MSIS data for broad research use (must be of public health significance) that justifies the initial disclosure and solely for the study described in the T-MSIS RIF Request Application, and ensures findings are publicly available [Limitations 2a, 3b] 3. Submits CDC COVID Restricted Access Data Use Agreement (RIDURA), T-MSIS RIF Data Use Agreement, Attachment A: RIF Application, RIF Application Key Personnel Supplement, RIF Specifications Worksheet, and Data Management Plan Self-Attestation Questionnaire (DMP SAQ) [Controls 2b, 3e] 4. Obtains approval from ResDAC team on the proposed linkage [Control 3f] 5. Obtains approval from CMS' Data Privacy Safeguard Program (DPSP) on the Data Management Plan Self-Attestation Questionnaire (DMP SAQ) [Control 3g] 6. Obtains IRB LOD from the researcher's institutions [Controls 2c, 3h] 7. Works with T-MSIS and CDC COVID to establish federal agency authorization for T-MSIS linkage with CDC COVID and either (1) upload data obtained from CDC GitHub private repository into VRDC, where T-MSIS data can be accessed - or - (2) obtain T-MSIS data in encrypted shipped disks and upload to CDC GitHub private repository [Controls 2e, 3i] - ***Assumption for CDC COVID and T-MSIS***		2c. Obtain IRB LOD from researcher's institution (as needed/applicable) [CDC COVID] 2e. Access data from CDC GitHub private repository [CDC COVID] **For accessing T-MSIS data, researchers/users must:** 3e. Submit RIF Data Use Agreement, Attachment A: RIF Application, RIF Application Key Personnel Supplement, RIF Specifications Worksheet, and Data Management Plan Self-Attestation Questionnaire (DMP SAQ) [T-MSIS] 3f. Obtain review and approval from ResDAC team on the proposed research [T-MSIS] 3g. Obtain review and approval from CMS' Data Privacy Safeguard Program (DPSP) on the Data Management Plan Self-Attestation Questionnaire (DMP SAQ) [T-MSIS] 3h. Obtain IRB LOD from the requesting institution [T-MSIS] 3i. Access data from the VRDC or through encrypted shipped disks [T-MSIS]	

TABLE 2: USE CASE 2 - DATASET LINKAGE DETERMINATION

Definitions: Limitations are restrictions on data linkage and use (e.g., dataset must only be linked with other disease-relevant data, dataset must be used in a physical enclave, etc.). **Controls** are processes established to ensure compliance with governance for data sharing, access, and use (e.g., user must access data in a physical enclave, user must sign data use agreement, user must receive data access committee approval, etc.). **Authorization gaps** exist when there is no explicit authorization or the authorization is not available or found in the information collected by the project team for various data life cycle stages (data collection, linking, sharing, access and use).

		Can the datasets be linked?	What limitations do the linked datasets inherit?	What controls do the linked dataset require?	What authorization gaps exist?
4	All datasets	**Yes, NCCR, CDC COVID, and T-MSIS can be linked provided:** **A. SEER/CDC COVID/CCW VRDS staff**: 1. Share NCCR, CDC COVID, and T-MSIS data fully de-identified of all direct identifiers [Controls 1a, 2a, 3a] 2. Perform disclosure review -for CDC COVID data, to suppress data fields with low frequency (<5) prior to sharing -for T-MSIS data, to review analysis outputs prior to sharing output data from the VRDC [Controls 2b, 3d] **B. Federal entities who create linked T-MSIS data and share the linked data**: 3. Take the data into their SORN under the Privacy Act when performing data linkage using non-standard TAFs containing PII [Control 3b] 4. Treat secondarily shared data as if they are a HIPAA Covered Entity and follow a process similar to CMS for releasing data including entering into a DUA with the researcher [Control 3c] 5. Link NCCR data only based on applicable state laws for participating NCCR registries, in accordance with protocol for linkage approved by the NCCR data providers, and the RIF Application approved by ResDAC team and CMS Privacy Board and in accordance with the Information Exchange	**Researchers/users:** 1a. Must only link NCCR data based on applicable state laws which specify the scope of linkage (e.g., Idaho) and in accordance with protocol for linkage approved by the data providers (i.e., state registry) [NCCR] 1b. Must use or disclose NCCR data only for the purposes for approved research [NCCR] 2a. Must use CDC COVID data for broad research (must be of public health significance) [CDC COVID] 3b. Must use T-MSIS data for research use that justifies the initial disclosure and solely for the study described in detail in the RIF Request Application, and must ensure findings are publicly available [T-MSIS] **Federal entities who create linked T-MSIS data:** 3a. Must only link T-MSIS data in accordance with the RIF Application approved by the CMS Privacy Board which specifies the scope of linkage and in accordance with the Information Exchange Agreement (IEA) between CMS and Participating Agency which specifies the scope of linkage for federal entities performing linkage with non-standard TAFs containing PII [T-MSIS]	**For sharing NCCR data, SEER staff must:** 1a. Share NCCR data fully de-identified of all direct identifiers (including exact ages and dates and geographic information, except quintiles) [NCCR] **For sharing CDC COVID data, staff must:** 2a. De-identify data of all direct identifiers [CDC COVID] 2b. Perform disclosure review to suppress data fields with low frequency (<5) prior to sharing CDC COVID data [CDC COVID] **For sharing T-MSIS data:** **A. Federal entities who create linked T-MSIS data and share the linked data must:** 3a. De-identify T-MSIS data of all 18 HIPAA identifiers as per HIPAA [T-MSIS] 3b. Take the data into their SORN under the Privacy Act when performing data linkage using non-standard TAFs containing PII [T-MSIS] 3c. Treat secondarily shared data as if they are a HIPAA Covered Entity and follow a process similar to CMS for releasing data including entering into a DUA with the researcher [T-MSIS] **B. CCW VRDC staff must:** 3d. Perform disclosure review of analysis outputs prior to sharing output data from the VRDC [T-MSIS]	2a. For CDC COVID-19, information on authorizations for linkage and sharing is not available/found.

TABLE 2: USE CASE 2 - DATASET LINKAGE DETERMINATION

Definitions: Limitations are restrictions on data linkage and use (e.g., dataset must only be linked with other disease-relevant data, dataset must be used in a physical enclave, etc.). **Controls** are processes established to ensure compliance with governance for data sharing, access, and use (e.g., user must access data in a physical enclave, user must sign data use agreement, user must receive data access committee approval, etc.). **Authorization gaps** exist when there is no explicit authorization or the authorization is not available or found in the information collected by the project team for various data life cycle stages (data collection, linking, sharing, access and use).

		Can the datasets be linked?	What limitations do the linked datasets inherit?	What controls do the linked dataset require?	What authorization gaps exist?
		Agreement (IEA) between CMS and Participating Agency - all of which specify the scope of linkage [Limitations 1a, 3a] **C. The researcher/user:** 1. Obtains authorizations for data linkage and sharing for CDC COVID [Authorization gap 1] - ***Assumption*** 2. Uses the linked NCCR, CDC COVID and T-MSIS data for broad research use (must be of public health significance) that justifies the initial disclosure and solely for the study described in the T-MSIS RIF Request Application, and ensures findings are publicly available [Limitations 1b, 2a, 3b] 3. Submits NCCR Data Analysis Plan, NCCR DUA, and CDC COVID Restricted Access Data Use Agreement (RIDURA), T-MSIS RIF Data Use Agreement, Attachment A: RIF Application, RIF Application Key Personnel Supplement, RIF Specifications Worksheet, and Data Management Plan Self-Attestation Questionnaire (DMP SAQ) [Controls 1b, 1c, 2b, 3e] 4. Obtains approval from NCI Office of Data Sharing and the Surveillance Research Program's Data Release group and ResDAC team on the proposed linkage [Controls 1d, 3f] 5. Obtains IRB LOD from the researcher's institution for NCCR, CDC COVID and T-MSIS data (or NCI BRANY as needed for NCCR) and IRB approval from the NCCR state registry [Controls 1e, 1f, 2c, 3h] 6. Uses an institutional account (known as eRA Commons) and obtains verification of Signing Official by the NCI Office of Data Sharing and the Surveillance Research Program's Data Release group for NCCR [Control 1g] 7. Works with T-MSIS, NCCR, and CDC COVID staff to establish federal agency authorization for T-MSIS linkage with NCCR and CDC COVID and to upload CDC COVID data obtained from CDC GitHub private repository and T-MSIS data obtained via encrypted shipped disks to SEER*Stat (client server software), where NCCR data can be accessed [Controls 1h, 2e, 3i] - ***Assumption for NCCR, CDC COVID and T-MSIS***		**For accessing NCCR data, researchers/users must:** 1b. Submit Data Analysis Plan [NCCR] 1c. Execute the NCCR DUA [NCCR] 1d. Obtain review by and approval from the NCI Office of Data Sharing and the Surveillance Research Program's Data Release group on the proposed research [NCCR] 1e. Obtain IRB LOD from the researcher's institution, or from NCI central IRB (BRANY) if the researchers institution does not have an IRB (as needed/applicable) [NCCR] 1f. Obtain IRB approval from the state registry [NCCR] 1g. Use an institutional account (known as eRA Commons) and obtain verification of Signing Official by the NCI Office of Data Sharing and the Surveillance Research Program's Data Release group [NCCR] 1h. Access the data from SEER*Stat (client server software) [NCCR] **For accessing CDC COVID data, researchers/users must:** 2b. Execute the Restricted Access Data Use Agreement (RIDURA) [CDC COVID] 2c. Obtain IRB LOD from researcher's institution (as needed/applicable) [CDC COVID] 2e. Access data from CDC GitHub private repository [CDC COVID] **For accessing T-MSIS data, researchers/users must:** 3e. Submit RIF Data Use Agreement, Attachment A: RIF Application, RIF Application Key Personnel Supplement, RIF Specifications Worksheet, and Data Management Plan Self-Attestation Questionnaire (DMP SAQ) [T-MSIS] 3f. Obtain review and approval from ResDAC team on the proposed research [T-MSIS] 3g. Obtain review and approval from CMS' Data Privacy Safeguard Program (DPSP) on the Data Management Plan Self-Attestation Questionnaire (DMP SAQ) [T-MSIS] 3h. Obtain IRB LOD from the requesting institution [T-MSIS] 3i. Access data from the VRDC or through encrypted shipped disks [T-MSIS]	

USE CASE 2 - GOVERNANCE INFORMATION				
Use Case 2: What is the impact of COVID-19 infection on pediatric cancer survivors? Or what is the impact of COVID-19 infection on future pediatric cancer outcomes?				
Dataset 5 - NCCR Data				
	Dataset Source	National Childhood Cancer Registry (NCCR)		
	Dataset Source Agency	NIH/NCI		
	Dataset Type (Clinical, EHR, Survey, SDOH, etc.)	Clinical/Demographics		
	Information Sources	Website, interview with NCCR staff member		
Dataset 5 - NCCR Data				
		Raw Language	**Interpretation**	**Source**
1	**Data Collection**			
1.1	**Authorizations and Applicable Regulations/Policies**			
1.1.1	**Authorizations**		1. Content (if applicable) 2. IRB approval for some states (e.g., New Jersey) 3. State Laws/Regulations (if applicable) 4. Institutional Certification (if applicable)	
1.1.1.1	Assent	Patients are not asked to provide consent for information to be reported to the central cancer registry. [1] We work with some data providers that obtained data originally under a research consent (PPCR, COG). In those cases, local IRBs have been part of reviewing if the consent covers broad data sharing or not and the process to reconsent individiuals. This may also involve obtaining additional organizational approvals. We may also work with BRANY as the central IRB for NCI to determine if the NCCR is a data repository that is exempt from human subjects research as it is all secondary research. We are also requiring investigators to obtain IRB approval in the first year of use to monitor how IRBs view the NCCR. [2]	Information not available/found	[1] NCCR Meeting [2] NCCR Email Communication
1.1.1.2	Consent	Patients are not asked to provide consent for information to be reported to the central cancer registry. [1] We work with some data providers that obtained data originally under a research consent (PPCR, COG). In those cases, local IRBs have been part of reviewing if the consent covers broad data sharing or not and the process to reconsent individiuals. This may also involve obtaining additional organizational approvals. We may also work with BRANY as the central IRB for NCI to determine if the NCCR is a data repository that is exempt from human subjects research as it is all secondary research. We are also requiring investigators to obtain IRB approval in the first year of use to monitor how IRBs view the NCCR. [2]	Consent authorizes data collection for some data providers who originally obtained the data under a research consent	[1] NCCR Meeting [2] NCCR Email Communication
1.1.1.3	IRB/equivalent Privacy Board determination	New Jersey has a Data Repository IRB approved protocol the Rutgers uses to oversee the activities of the registry in data collection and its eventual use for research. The New Jersey repository IRB covers all data collection including SEER sponsored data linkages include the GHI linkage, SEER-Medicare, SEER-Medicaid, and will also cover linkages with CVS and Walgreens. The data repository also covers non-traditional data sources for incidence, treatment and follow-up such as Meaningful Use, claims, NDI, etc. [1]	IRB approval authorizes data collection for certain states (e.g., New Jersey)	[1] NCCR Meeting

Dataset 5 - NCCR Data				
		Raw Language	**Interpretation**	**Source**
1.1.1.4	Local/state/federal laws	Central cancer registries collect information about cancer diagnosis and treatment according to state laws. As a reportable illness, healthcare providers disclose information about cancer patients to the central cancer registry and sharing this information is exempt from HIPAA https://www.hhs.gov/hipaa/for-professionals/privacy/guidance/disclosures-public-health-activities/index.html . Therefore, patients are not asked to provide consent for information to be reported to the central cancer registry. In some cases, the state health department directly manages the activities of the central cancer registries but in other cases this is contracted to another entity, usually a university (as is the case in Florida, Georgia, Kentucky, etc.). Additionally, some state laws and regulations about reporting information about cancer patients ranges from outdated to regularly updated by state legislatures or health departments. Some registries are bound by law to be incident-only and do not have authority to hold longitudinal, post-cancer information (Michigan). It is important to read state law and regulations for each state when there are questions about what kind of data are reported by providers to the central cancer registry (see attached examples). Subsequently, SEER registries have a contractual relationship with the NCI to license de-identified data to the NCI. With NPCR registries, the direct relationship is with NAACCR and Information Management Services, Inc. (IMS) (serves as the honest broker) to hold data. Additionally, IMS executes Interconnection Services Agreements with each registry in a few cases Business Associate Agreements to act as an honest broker and IT vendor for each registry's isolated environment in the IMS private data center. For some projects, central cancer registries do seek IRB approval but thus far this does not apply to NCCR data submissions. [1] All cancer registries are authorized by the state regulations health care providers are responsible for reporting cancer diagnosis to the central cancer registry (in some cases, the registry is the state health department, or subcontractor to the state health department e.g., for California, UCSF operates regional registries for the state health department)'. Regulations vary by state (NCCR staff sent list of sample state legislature that are tracked) – for example, Michigan registries are not permitted to hold longitudinal data, on incidence data. State legislature typically defines 1) what is required to be reported, 2) who is required to report, 3) where the reporting is done, and 4) in what format (typically states will say whatever is in the NAACCR format should be reported – but not very detailed, which is why NCCR obtains other data like pharmacy data). The state laws are not backed up by any federal laws. Providers are exempted from having to tell or ask patients about the disclosure of PII to the central cancer registry. [1] State regulations/policies authorize original data collection by central cancer registries as part of public health surveillance. Data are then de-identified and licensed to the NCI under contract. Most data flow through this process, even if the data provider is centrally coordinated by NCI (such as Walgreens, Rite Aid, and CVS providing pharmacy claims). Some data are obtained through PPRL and cannot be provided to registries, but only held by the NCI. The NCI is part of the National Institutes of Health (NIH) and the NIH is an entity of the Public Health Service (Sections 202 and 401 of the Public Health Service Act, 42 U.S.C. §§ 203, 281). Pursuant to Section 402 of the Public Health Service Act (42 U.S.C. § 282), the Director of the NIH has the authority to conduct certain functions in furtherance of the public health purposes of Section 301 of the Public Health Service Act. The NCI is authorized to conduct the cancer research, control and information dissemination activities described herein pursuant to Sections 405, 410, 412 and 413 of the Public Health Service Act (42 U.S.C. §§ 284, 285, 285a-1 and 285a-2). [2]	State regulations/policies authorizes data collection by central cancer registries as part of public health surveillance	[1] NCCR Meeting [2] NCCR Email Communication
1.1.1.5	Institutional Certification	State or regional central cancer registries may obtain institutional certification to collect and retain data about cancer patients. [1]	Institutional Certification authorizes data collection for some state or regional cancer registries	[1] NCCR Email Communication
1.1.1.6	Data originator agreement	N/A	N/A	
1.1.1.7	Repository agreements/policies	N/A	N/A	
1.1.1.8	Other (specify)	Information not available/found	Information not available/found	
1.1.2	**Applicable Regulations/Policies**			
1.1.2.1	Local regulations/policies	N/A	N/A	
1.1.2.2	Tribal regulations/policies	N/A - No tribal data [1]	N/A	[1] NCCR Meeting

Dataset 5 - NCCR Data				
		Raw Language	**Interpretation**	**Source**
1.1.2.3	State regulations/policies	Data from Cancer in North America (CiNA) (North American Association of Central Cancer Registries (NAACCR) 1995-2018 and the NCI's Surveillance, Epidemiology and End Results (SEER) Registries), submitted December 2020). Registries include: California, Connecticut, Florida, Georgia, Hawaii, Idaho, Illinois, Iowa, Kentucky, Louisiana, Massachusetts, New Jersey, New Mexico, New York, Ohio, Pennsylvania, Seattle (Puget Sound), Tennessee, Texas, Utah, Wisconsin. These 23 NCCR registries represent 66% of all U.S. children, adolescents, and young adults ages 0-39 based on 2018 U.S. Populations.	State regulations/policies the participating NCCR registries: Arkansas, California, Connecticut, Florida, Georgia, Hawaii, Idaho, Illinois, Iowa, Kentucky, Louisiana, Massachusetts, Michigan, New Jersey, New Mexico, New York, Ohio, Pennsylvania, Seattle (Puget Sound), Tennessee, Texas, Utah, and Wisconsin	[1] https://datacatalog.ccdi.cancer.gov/dataset/NCCR-NCCRExplorer (Accessed: 4/28/23) [2] NCCR Email Communication
1.1.2.4	Federal regulations/policies	N/A - the state laws are not backed up by any federal laws. [1]	N/A	[1] NCCR Meeting
1.1.2.5	International regulations/policies	N/A	N/A	
1.1.2.6	Contractual obligations	All cancer registries are authorized by the state regulations health care providers are responsible for reporting cancer diagnosis to the central cancer registry (in some cases, the registry is the state health department, or subcontractor to the state health department e.g., for California, UCSF operates regional registries for the state health department)'. [1] Contractual obligations for registries operated by the state health department to the NCI. Some registries participate through a grant under NAACCR. [2]	1. Contractual obligations for registries operated by the state health department to the NCI 2. Obligations for registries participating through a grant under NAACCR	[1] NCCR Meeting [2] NCCR Email Communication
1.1.2.7	Repository policies	N/A	N/A	
1.2	**Governance for data linkage, sharing, access, and use based on data collection authorization or applicable regulations/policies (i.e., the origin of the governance)**			
1.2.1	Whether the data can be linked	External Linkages. Louisiana Tumor Registry (LTR) data may be linked with external databases in order to improve the accuracy and completeness of follow-up data or for research. Health care provider shall mean every licensed health care facility and licensed health care provider, as defined in R.S. 40:1231.1(A), in the state of Louisiana. [1] State public health surveillance laws vary in restrictions on use of linked data but generally matching patients from different data sources to the cancer registry data, adding to traditional registry data, and then de-identifying data for submission to the NCI is permitted. [2]	Certain state laws/regulations specify that the data can be linked (e.g., Lousiana)	[1] SEER Legal Authority for Detailed Pharmacy Treatment Data Collection from NCCR Staff [2] NCCR Email Communication
1.2.2	With what other data can it be linked or can it not be linked (scope of linkage)	Idaho Code Title 57, Chapter 17, 57-1703 (6) defines the scope of data collection: "Population-based" refers to all cancers and reportable benign tumors diagnosed and/or treated within the state of Idaho by hospitals or other facilities providing screening, diagnostic or therapeutic services to patients with respect to cancer, and from physicians, surgeons, and all other health care providers diagnosing or providing treatment for cancer patients.'... Linkages to gather treatment and other information are specifically allowed in 57-1805(5): 'Nothing in this chapter shall prevent the department or authorized contractor from identifying and reporting cases using data linkages with death records, statewide cancer registries, and other potential sources.' [1] State public health surveillance laws vary in restrictions on use of linked data but generally matching patients from different data sources to the cancer registry data, adding to traditional registry data, and then de-identifying data for submission to the NCI is permitted. [2]	Certain state laws/regulations specify the scope of linkage (e.g., Idaho)	[1] SEER Legal Authority for Detailed Pharmacy Treatment Data Collection from NCCR Staff [2] NCCR Email Communication
1.2.3	Whether data can be shared	Does not authorize/specify	Does not authorize/specify	[1] SEER Legal Authority for Detailed Pharmacy Treatment Data Collection from NCCR Staff
1.2.4	How data can be shared (de-identification status, disclosure review)	Does not authorize/specify	Does not authorize/specify	[1] SEER Legal Authority for Detailed Pharmacy Treatment Data Collection from NCCR Staff
1.2.5	How data can be accessed (access type, data use agreement, data access committee/group approval, IRB LOD, etc.)	Does not authorize/specify	Does not authorize/specify	[1] SEER Legal Authority for Detailed Pharmacy Treatment Data Collection from NCCR Staff
1.2.6	How data can be used (data use limitations)	Does not authorize/specify	Does not authorize/specify	[1] SEER Legal Authority for Detailed Pharmacy Treatment Data Collection from NCCR Staff
1.2.7	Other (specify)	Does not authorize/specify	Does not authorize/specify	[1] SEER Legal Authority for Detailed Pharmacy Treatment Data Collection from NCCR Staff
2	**Data Linkage**			
2.1	**Authorizations and Applicable Regulations/Policies**			
2.1.1	**Authorizations**			
2.1.1.1	Assent	Information not available/found	Information not available/found	
2.1.1.2	Consent	Information not available/found	Information not available/found	

Dataset 5 - NCCR Data				
		Raw Language	**Interpretation**	**Source**
2.1.1.3	IRB/equivalent Privacy Board determination	N/A	N/A	
2.1.1.4	Local/state/federal laws	N/A	N/A	
2.1.1.5	Institutional Certification	N/A	N/A	
2.1.1.6	Data originator agreement	The linked data will come to NCI/NCCR de-identified when NCCR is ready to receive data, NCCR will post the requirements for linkage SEER data (https://seer.cancer.gov/tools/submission.html), so NAACCR (subcontractor) will submit under Call for Data. [1] For each linkage, registries review and approve a protocol specific to that linkage. Approvals are granted by senior personnel at the registry and are often reviewed by legal advisors within the state health departments who manage the registry or the registry sub-contractor. In some cases, data providers and registries sign additional Data Sharing (or Use) Agreements. This usually happens between the registry and IMS. In some cases, a MOU or MOA is signed between a data provider and the NCI. [1] Below is a table of the VPR-only registries that are participating in linkage with the NCCR. These nine registries have a Data Sharing Agreement with IMS allowing IMS to link the data and agreeing for data on prior/subsequent cancers to be incorporated into the NCCR. The table shows whether the registry would like the full record incorporated or just a list of basic data items (which now included stage and treatment). It also identifies whether the registry would like IMS to extract the data for incorporation into the NCCR. [1] Data Sharing (or Use) Agreement or a hosting agreement (ISA) or BAA executed between IMS and the data provider (as needed / if applicable) authorizes data storage, linkage, management, and destruction of original data [2]	Three agreements authorize data linkage (as needed / if applicable): 1. Protocol for linkage approved by the data provider authorizes data linkage 2. Data Sharing (or Use) Agreement or a hosting agreement (ISA) or BAA executed between IMS and the data provider authorize data linkage 3. MOU or MOA executed between NCI and the data provider authorizes data linkage	[1] NCCR Meeting [2] NCCR Email Communication
2.1.1.7	Repository agreements/policies	Information not available/found	Information not available/found	
2.1.1.8	Other (specify)	Information not available/found	Information not available/found	
2.1.2	**Applicable Regulations/Policies**			
2.1.2.1	Local regulations/policies	N/A	N/A	
2.1.2.2	Tribal regulations/policies	N/A - No tribal data. [1]	N/A	[1] NCCR Meeting
2.1.2.3	State regulations/policies	N/A	N/A	
2.1.2.4	Federal regulations/policies	N/A	N/A	
2.1.2.5	International regulations/policies	N/A	N/A	
2.1.2.6	Contractual obligations	So far, NCCR has not released any linked data (still setting up the mechanisms to do that). The way it usually works depending on if its NPCR registry (under NAACCR subcontract) or SEER registry (direct contract with NCI), they sign an Interconnection Services Agreement, and in some cases also a BAA with IMS if required by the registry. The purpose of these agreements are for IMS to act as the honest broker to hold the PII and perform the linkage using Match Pro there is an addendum to some of these contracts specifically for NCCR, since it was novel that the PII be held in one file across all registries because care is received/given across state lines across the participant's lifetime. [1]	1. Interconnection Services Agreement (as needed / if applicable) 2. Business Associate Agreement (BAA) (as needed / if applicable)	[1] NCCR Meeting
2.1.2.7	Repository policies	N/A	N/A	
2.2	**Governance for data linkage, sharing, access, and use based on data linkage authorization or applicable regulations/policies (i.e., the origin of the governance)**			

Dataset 5 - NCCR Data				
		Raw Language	**Interpretation**	**Source**
2.2.1	Whether the data can be linked	The linked data will come to NCI/NCCR de-identified when NCCR is ready to receive data, NCCR will post the requirements for linkage SEER data (https://seer.cancer.gov/tools/submission.html), so NAACCR (subcontractor) will submit under Call for Data. [1] For each linkage, registries review and approve a protocol specific to that linkage. Approvals are granted by senior personnel at the registry and are often reviewed by legal advisors within the state health departments who manage the registry or the registry sub-contractor. In some cases, data providers and registries sign additional Data Sharing (or Use) Agreements. This usually happens between the registry and IMS. In some cases, a MOU or MOA is signed between a data provider and the NCI. [1] Below is a table of the VPR-only registries that are participating in linkage with the NCCR. These nine registries have a Data Sharing Agreement with IMS allowing IMS to link the data and agreeing for data on prior/subsequent cancers to be incorporated into the NCCR. The table shows whether the registry would like the full record incorporated or just a list of basic data items (which now included stage and treatment). It also identifies whether the registry would like IMS to extract the data for incorporation into the NCCR. [1] State public health surveillance laws vary in restrictions on use of linked data but generally matching patients from different data sources to the cancer registry data, adding to traditional registry data, and then de-identifying data for submission to the NCI is permitted. [2]	Data originator agreements specify that the data can be linked	[1] NCCR Meeting [2] NCCR Email Communication
2.2.2	With what other data can it be linked or can it not be linked (scope of linkage)	The linked data will come to NCI/NCCR de-identified when NCCR is ready to receive data, NCCR will post the requirements for linkage SEER data (https://seer.cancer.gov/tools/submission.html), so NAACCR (subcontractor) will submit under Call for Data. [1] For each linkage, registries review and approve a protocol specific to that linkage. Approvals are granted by senior personnel at the registry and are often reviewed by legal advisors within the state health departments who manage the registry or the registry sub-contractor. In some cases, data providers and registries sign additional Data Sharing (or Use) Agreements. This usually happens between the registry and IMS. In some cases, a MOU or MOA is signed between a data provider and the NCI. [1] Below is a table of the VPR-only registries that are participating in linkage with the NCCR. These nine registries have a Data Sharing Agreement with IMS allowing IMS to link the data and agreeing for data on prior/subsequent cancers to be incorporated into the NCCR. The table shows whether the registry would like the full record incorporated or just a list of basic data items (which now included stage and treatment). It also identifies whether the registry would like IMS to extract the data for incorporation into the NCCR. [1] State public health surveillance laws vary in restrictions on use of linked data but generally matching patients from different data sources to the cancer registry data, adding to traditional registry data, and then de-identifying data for submission to the NCI is permitted. [2]	Data originator agreements specify that the data can be linked according to the protocol for linkage approved by the data provider	[1] NCCR Meeting [2] NCCR Email Communication
2.2.3	Whether data can be shared	Information not available/found	Information not available/found	
2.2.4	How data can be shared (de-identification status, disclosure review)	Generally, we have many policies to follow per original agreements with data providers. We do not share any geographic identifiers, although we do provide geographically-calculated fields in the form of quintiles (e.g., Yost Index at the time of cancer diagnoses calculated at the census tract level using original address provided to the registries but stripped of census tract and address by the time it is received by NCI for release). We also only provide calculated time between dates (i.e., no dates or month/year) or age in years or categorical groupings. We also have a data access committee that will review all requests for individual-level data and will require IRB review of those data requests by investigators prior to release. [1]	Data originator agreements specify that no geographic identifiers (except quintiles), exact dates, or exact ages will be shared	[1] NCCR Email Communication
2.2.5	How data can be accessed (access type, data use agreement, data access committee/group approval, IRB LOD, etc.)	Information not available/found	Information not available/found	
2.2.6	How data can be used (data use limitations)	Information not available/found	Information not available/found	

Dataset 5 - NCCR Data				
		Raw Language	**Interpretation**	**Source**
2.2.7	Other (specify)	Information not available/found	Information not available/found	
3	**Data Sharing**			
3.1	**Authorizations and Applicable Regulations/Policies**			
3.1.1	**Authorizations**			
3.1.1.1	Assent	Information not available/found	Information not available/found	[1] NCCR Meeting
3.1.1.2	Consent	Patients are not asked to provide consent for information to be reported to the central cancer registry. [1] We work with some data providers that obtained data originally under a research consent (PPCR, COG). In those cases, local IRBs have been part of reviewing if the consent covers broad data sharing or not and the process to reconsent individiuals. This may also involve obtaining additional organizational approvals. We may also work with BRANY as the central IRB for NCI to determine if the NCCR is a data repository that is exempt from human subjects research as it is all secondary research. We are also requiring investigators to obtain IRB approval in the first year of use to monitor how IRBs view the NCCR. [2]	Consent authorizes data sharing for some data providers	[1] NCCR Meeting [2] NCCR Email Communication
3.1.1.3	IRB/equivalent Privacy Board determination	For some projects, central cancer registries do seek IRB approval, but thus far this does not apply to NCCR data submissions. [1] IRB approval authorizes data collection for certain states (e.g., Wisconsin) [2] We work with some data providers that obtained data originally under a research consent (PPCR, COG). In those cases, local IRBs have been part of reviewing if the consent covers broad data sharing or not and the process to reconsent individiuals. This may also involve obtaining additional organizational approvals. We may also work with BRANY as the central IRB for NCI to determine if the NCCR is a data repository that is exempt from human subjects research as it is all secondary research. We are also requiring investigators to obtain IRB approval in the first year of use to monitor how IRBs view the NCCR. [2]	IRB approval authorizes data sharing for some data providers	[1] NCCR Meeting [2] NCCR Email Communication
3.1.1.4	Local/state/federal laws	The NCI is part of the National Institutes of Health (NIH) and the NIH is an entity of the Public Health Service (Sections 202 and 401 of the Public Health Service Act, 42 U.S.C. §§ 203, 281). Pursuant to Section 402 of the Public Health Service Act (42 U.S.C. § 282), the Director of the NIH has the authority to conduct certain functions in furtherance of the public health purposes of Section 301 of the Public Health Service Act. The NCI is authorized to conduct the cancer research, control and information dissemination activities described herein pursuant to Sections 405, 410, 412 and 413 of the Public Health Service Act (42 U.S.C. §§ 284, 285, 285a-1 and 285a-2). [1]	Sections 405, 410, 412 and 413 of the Public Health Service Act (42 U.S.C. §§ 284, 285, 285a-1 and 285a-2 authorize data sharing	[1] NCCR Email Communication
3.1.1.5	Institutional Certification	Information not available/found	Information not available/found	
3.1.1.6	Data originator agreement	NPCR-NCCR registries have a DAA (Data Assurances Agreement) with NCI to allow NCI to release data. SEER-NCCR registries agree to NCI's release of data under contract. Both types of registries agree to data release when they approve protocols for data linkages against their own registry data. [1] The details of how NCI will release data are generally described in original protocols approved by registries and data providers. SEER-NCCR registries must also complete NIH DMSP as part of their contracts with NCI. Other agreements with data providers may generally describe data sharing and data destruction policies. [1]	Protocol for linkage approved by the data provider authorizes data sharing	[1] NCCR Email Communication
3.1.1.7	Repository agreements/policies	Information not available/found	Information not available/found	

Dataset 5 - NCCR Data				
		Raw Language	**Interpretation**	**Source**
3.1.1.8	Other	SEER registries have a contractual relationship with the NCI to license de-identified data to the NCI. With NPCR registries, the direct relationship is with NAACCR and IMS to hold data. [1] Not all SEER registries participate in NCCR – SEER registries have a contractual relationship with NCI and license the data to the NCI to be used in data releases in the de-identified form: -IMS is the honest broker and they hold the PII on behalf of the registries that own the data -For the NPCR registries that participate, they have a subcontract through NAACCR to submit the data to IMS to be de-identified and submitted to NCCR so, NPCR has a direct relationship with NAACCR, not NCI. [1] Currently, the 5 NPCR registries that participate do so through a sub award from NAACCR's NCCR contract with NCI. Beginning in 2023, additional registries covering Oregon, Arkansas, Colorado, Missouri, etc. will join via the NAACCR sub award or, if they are already a SEER research support registry, through a task order with the NCI. The main implication is that when NCI has a direct relationship with the registries there is better access to defining submission standards and using the technology infrastructure that makes data submission, cleaning, linkages, etc. easier. [1] Contractual obligations for registries operated by the state health department to the NCI. Some registries participate through a grant under NAACCR. [2] NPCR-NCCR registries have a DAA (Data Assurances Agreement) with NCI to allow NCI to release data. SEER-NCCR registries agree to NCI's release of data under contract. Both types of registries agree to data release when they approve protocols for data linkages against their own registry data. [2]	Two contracts authorize data sharing: 1. Contract between SEER registries and NCI 2. Subcontract between NPCR registries and NAACCR and grant between NPCR registries and NAACCR	[1] NCCR Meeting [2] NCCR Email Communication
3.1.2	**Applicable Regulations/Policies**			
3.1.2.1	Local regulations/policies	Information not available/found	Information not available/found	
3.1.2.2	Tribal regulations/policies	N/A - No tribal data. [1]	N/A	[1] NCCR Meeting
3.1.2.3	State regulations/policies	Information not available/found	Information not available/found	
3.1.2.4	Federal regulations/policies	Information not available/found	Information not available/found	
3.1.2.5	International regulations/policies	N/A	N/A	
3.1.2.6	Contractual obligations	Information not available/found	Information not available/found	[1] NCCR Email Communication
3.1.2.7	Repository policies	Information not available/found	Information not available/found	
3.2	**Governance for data linkage, sharing, access, and use based on data sharing authorization or applicable regulations/policies (i.e., the origin of the governance)**			
3.2.1	Whether the data can be linked	NPCR-NCCR registries have a DAA (Data Assurances Agreement) with NCI to allow NCI to release data. SEER-NCCR registries agree to NCI's release of data under contract. Both types of registries agree to data release when they approve protocols for data linkages against their own registry data. [1]	Data originator agreements specify that the data can be linked	[1] NCCR Email Communication
3.2.2	With what other data can it be linked or can it not be linked (scope of linkage)	NPCR-NCCR registries have a DAA (Data Assurances Agreement) with NCI to allow NCI to release data. SEER-NCCR registries agree to NCI's release of data under contract. Both types of registries agree to data release when they approve protocols for data linkages against their own registry data. [1]	Data originator agreements specify that the data can be linked according to the protocol for linkage approved by the data provider	[1] NCCR Email Communication

Dataset 5 - NCCR Data				
		Raw Language	**Interpretation**	**Source**
3.2.3	Whether data can be shared	We work with some data providers that obtained data originally under a research consent (PPCR, COG). In those cases, local IRBs have been part of reviewing if the consent covers broad data sharing or not and the process to reconsent individuals. This may also involve obtaining additional organizational approvals. We may also work with BRANY as the central IRB for NCI to determine if the NCCR is a data repository that is exempt from human subjects research as it is all secondary research. We are also requiring investigators to obtain IRB approval in the first year of use to monitor how IRBs view the NCCR. [1] The NCI is part of the National Institutes of Health (NIH) and the NIH is an entity of the Public Health Service (Sections 202 and 401 of the Public Health Service Act, 42 U.S.C. §§ 203, 281). Pursuant to Section 402 of the Public Health Service Act (42 U.S.C. § 282), the Director of the NIH has the authority to conduct certain functions in furtherance of the public health purposes of Section 301 of the Public Health Service Act. The NCI is authorized to conduct the cancer research, control and information dissemination activities described herein pursuant to Sections 405, 410, 412 and 413 of the Public Health Service Act (42 U.S.C. §§ 284, 285, 285a-1 and 285a-2). [1] SEER registries have a contractual relationship with the NCI to license de-identified data to the NCI. With NPCR registries, the direct relationship is with NAACCR and IMS to hold data. [1] Not all SEER registries participate in NCCR – SEER registries have a contractual relationship with NCI and license the data to the NCI to be used in data releases in the de-identified form: -IMS is the honest broker and they hold the PII on behalf of the registries that own the data -For the NPCR registries that participate, they have a subcontract through NAACCR to submit the data to IMS to be de-identified and submitted to NCCR so, NPCR has a direct relationship with NAACCR, not NCI. [1] Currently, the 5 NPCR registries that participate do so through a sub award from NAACCR's NCCR contract with NCI. Beginning in 2023, additional registries covering Oregon, Arkansas, Colorado, Missouri, etc. will join via the NAACCR sub award or, if they are already a SEER research support registry, through a task order with the NCI. The main implication is that when NCI has a direct relationship with the registries there is better access to defining submission standards and using the technology infrastructure that makes data submission, cleaning, linkages, etc. easier. [1] Contractual obligations for registries operated by the state health department to the NCI. Some registries participate through a grant under NAACCR. [2] NPCR-NCCR registries have a DAA (Data Assurances Agreement) with NCI to allow NCI to release data. SEER-NCCR registries agree to NCI's release of data under contract. Both types of registries agree to data release when they approve protocols for data linkages against their own registry data. [2]	1. Consent and IRB approval specify that data can be shared (when applicable if consent and IRB were used by data provider) 2. Sections 405, 410, 412 and 413 of the Public Health Service Act (42 U.S.C. §§ 284, 285, 285a-1 and 285a-2) specify that NCI can share data 3. Contracts between SEER registries and NCI, subcontracts between NPCR registries and NAACCR and grants between NPCR registries and NAACCR specify that data can be shared	[1] NCCR Email Communication [2] NCCR Email Communication
3.2.4	How data can be shared (de-identification status, disclosure review)	Not all SEER registries participate in NCCR – SEER registries have a contractual relationship with NCI and license the data to the NCI to be used in data releases in the de-identified form: -IMS is the honest broker and they hold the PII on behalf of the registries that own the data -For the NPCR registries that participate, they have a subcontract through NAACCR to submit the data to IMS to be de-identified and submitted to NCCR so, NPCR has a direct relationship with NAACCR, not NCI. [1] SEER registries have a contractual relationship with the NCI to license de-identified data to the NCI. With NPCR registries, the direct relationship is with NAACCR and IMS to hold data. [1] Data are de-identified prior to sharing. Month/Year dates are not provided (interval only); no geographic identifiers in the NCCR data platform or SEER Stat file. [1] Contract between SEER registries and NCI, and Subcontract between NCI and NAACCR who then awards grants to NPCR, specifies that data must be de-identified for sharing through NCCR (including removal of PII, dates, and geographic information). Protocols and other agreements with original data providers cover NCI's data release practices. [2]	1. Contract between SEER registries and NCI, and Subcontract between NPCR registries and NAACCR, specifies that data must be de-identified for sharing through NCCR (including removal of PII, dates, and geographic information) 2. Data originator agreements specify that no geographic identifiers (except quintiles), exact dates, or exact ages will be shared	[1] NCCR Meeting [2] NCCR Email Communication
3.2.5	How data can be accessed (access type, data use agreement, data access committee/group approval, IRB LOD, etc.)	Information not available/found	Information not available/found	
3.2.6	How data can be used (data use limitations)	Information not available/found	Information not available/found	

Dataset 5 - NCCR Data				
		Raw Language	**Interpretation**	**Source**
3.2.7	Other (specify)	Information not available/found	Information not available/found	
4	**Data Access**			
4.1	**Authorizations and Applicable Regulations/Policies**			
4.1.1	**Authorizations**			
4.1.1.1	Assent	Information not available/found	Information not available/found	[1] NCCR Meeting
4.1.1.2	Consent	Information not available/found	Information not available/found	[1] NCCR Meeting
4.1.1.3	IRB/equivalent Privacy Board determination	Information not available/found	Information not available/found	
4.1.1.4	Local/state/federal laws	Information not available/found	Information not available/found	
4.1.1.5	Institutional Certification	Information not available/found	Information not available/found	
4.1.1.6	Data originator agreement	Information not available/found	Information not available/found	
4.1.1.7	Repository agreements/policies	There will be a DUA for NCCR data approved by the registries and NCI as part of the SEER data access request process. NCI is in the process of planning the tiered access process for researchers and others to use NCCR data that will be similar to the existing SEER processes (see 1/17/2023 ppt attached and https://seer.cancer.gov/data/access.html). [1] This National Childhood Cancer Registry (NCCR) Research Data Use Agreement (DUA) (the "Agreement") outlines the terms and conditions for access to data in the National Cancer Institute (NCI) NCCR Research, Research Plus, Specialized, and Customized Databases (collectively, the "Databases"). [2] NCCR Data Use Agreement authorizes data access to researchers. NCCR data access committee will review all requests for individual-level data. NCI is evaluating the possibilities of having the NCCR Data Platform deemed exempt from human subjects research. All data requestors will be required to have IRB approval in the first year of the NCCR Data Platform. [3]	NCCR Data Use Agreement authorizes data access	[1] NCCR Meeting [2] https://seer.cancer.gov/data-software/documentation/seerstat/nov2022/nccr-dua-nov2022.html (Accessed: 4/28/23) [3] NCCR Email Communication
4.1.1.8	Other (specify)	Information not available/found	Information not available/found	
4.1.2	**Applicable Regulations/Policies**			
4.1.2.1	Local regulations/policies	Information not available/found	Information not available/found	
4.1.2.2	Tribal regulations/policies	N/A - No tribal data. [1]	N/A	[1] NCCR Meeting
4.1.2.3	State regulations/policies	Information not available/found	Information not available/found	
4.1.2.4	Federal regulations/policies	Information not available/found	Information not available/found	
4.1.2.5	International regulations/policies	N/A	N/A	
4.1.2.6	Contractual obligations	Information not available/found	Information not available/found	
4.1.2.7	Repository policies	There will be a DUA for NCCR data approved by the registries and NCI as part of the SEER data access request process. NCI is in the process of planning the tiered access process for researchers and others to use NCCR data that will be similar to the existing SEER processes (see 1/17/2023 ppt attached and https://seer.cancer.gov/data/access.html). [1] This National Childhood Cancer Registry (NCCR) Research Data Use Agreement (DUA) (the "Agreement") outlines the terms and conditions for access to data in the National Cancer Institute (NCI) NCCR Research, Research Plus, Specialized, and Customized Databases (collectively, the "Databases"). [2]	NCCR policy	[1] NCCR Meeting [2] https://seer.cancer.gov/data-software/documentation/seerstat/nov2022/nccr-dua-nov2022.html (Accessed: 4/28/23)
4.2	**Governance for data linkage, sharing, access, and use based on data access authorization or applicable regulations/policies (i.e., the origin of the governance)**			
4.2.1	Whether the data can be linked	NCCR DUA does not allow Data Linkage across SEER Databases [1]	DUA specifies that users are not authorized to link data across SEER databases	[1] https://seer.cancer.gov/data-software/documentation/seerstat/nov2022/nccr-dua-nov2022.html (Accessed: 4/28/23)
4.2.2	With what other data can it be linked or can it not be linked (scope of linkage)	NCCR DUA does not allow Data Linkage across SEER Databases	DUA specifies that users are not authorized to link data across SEER databases	[1] https://seer.cancer.gov/data-software/documentation/seerstat/nov2022/nccr-dua-nov2022.html (Accessed: 4/28/23)+A1
4.2.3	Whether data can be shared	Information not available/found	Information not available/found	
4.2.4	How data can be shared (de-identification status, disclosure review)	Information not available/found	Information not available/found	

Dataset 5 - NCCR Data				
		Raw Language	**Interpretation**	**Source**
4.2.5	How data can be accessed (access type, data use agreement, data access committee/group approval, IRB LOD, etc.)	There will be a DUA for NCCR data approved by the registries and NCI as part of the SEER data access request process. NCI is in the process of planning the tiered access process for researchers and others to use NCCR data that will be similar to the existing SEER processes (see 1/17/2023 ppt attached and https://seer.cancer.gov/data/access.html). Researchers seeking individual-level information and combinations of datasets and fields that increase the risk of re-identifiability (census-tract information or geographically-associated derived variables like the Yost index) will be required to submit a data analysis plan and possibly IRB-approval. NCI has a central IRB (BRANY) for this purpose if investigators do not use their own IRB. [1] If researcher puts in a request that requires IRB approval, they can go through the central IRB for NCCR. Each project has its own DUA (if necessary) and NCCR tracks limitations or restrictions moving forward for researchers looking to obtain data, there will be a NCCR DUA (which will be similar to the SEER DUA). [1] Currently, the process is planned as researchers submit an institutional account (known by eRA Commons) and have a Signing Official verified by NCI's Office of Data Sharing and the Surveillance Research Program for access to SEER*Stat or the NCCR Data Platform. If they require individual-level data through the Data Platform there will be an additional process to provide a data analysis plan and more information to approve release of a custom dataset. For additional information about what data will be in the Data Platform in the first year, see the sample codebook. [1] An NCCR dataset will be available in SEER*Stat that cannot be downloaded but researchers can compute various statistics in that software. Data providers (claims, PPCR, etc.) and registries sign various data sharing agreements and approve protocols to share data... SEER Stat will construct a standard registry dataset for all the registry data coming in, which will be updated annually and available via SEER*Stat to approved researchers (but download of the data won't be permitted – SEER*Stat is not an enclave, but a client server software. [1] Central IRB Review (slide 6): Likely no IRB review required from users of Tier 1-3 for SEER Data Products -Awaiting review and decision of the BRANY cIRB -Will have similar review for NCCR Data Products to determine need for IRB review IRB might be required by the user's home institution -NCI SRP retained the services of Biomedical Research Alliance of New York (BRANY) -Offers the services of a central IRB. [2] Must access and use the data within SEER*Stat (client server software) or NCCR Data Platform. [3]	NCCR specifies that for data access, a user: 1. Must submit a Data Analysis Plan 2. Must execute the NCCR DUA 3. Must obtain review and approval by the NCI Office of Data Sharing and the Surveillance Research Program's Data Release group on the proposed research 4. Must obtain IRB LOD from the researcher's institution, or from NCI central IRB (BRANY) if the researchers institution does not have an IRB (as needed / applicable) 5. Must obtain IRB approval from the state registry 6. Must submit using an insititutional account (known as eRA Commons) and obtain verification of Signing Official by the NCI Office of Data Sharing and the Surveillance Research Program's Data Release group 7. Must access and use the data within SEER*Stat (client server software) or NCCR Data Platform	[1] NCCR Meeting [2] Updated Data Release Process PPT from NCCR [3] NCCR Email Communication
4.2.6	How data can be used (data use limitations)	There will be a DUA for NCCR data approved by the registries and NCI as part of the SEER data access request process. NCI is in the process of planning the tiered access process for researchers and others to use NCCR data that will be similar to the existing SEER processes (see 1/17/2023 ppt attached and https://seer.cancer.gov/data/access.html). [1] National Childhood Cancer Registry (NCCR) Research Data Use Agreement: Use and Disclosure Restricted: Use and Disclosure Restricted: Authorized User will use or disclose the Data only for the purposes for which they were supplied. Requests for data with increased risk to re-identify individuals will require additional application materials and Authorized User will only use data for the purposes approved by the NCI. [2]	NCCR Data Use Agreement specifies that Authorized User will use or disclose the Data only for the purposes for approved research	[1] NCCR Meeting [2] https://seer.cancer.gov/data-software/documentation/seerstat/nov2022/nccr-dua-nov2022.html (Accessed: 4/28/23)
4.2.7	Other (specify)	Information not available/found	Information not available/found	
5	**Data Use**			
5.1	**Authorizations and Applicable Regulations/Policies**			
5.1.1	**Authorizations**			
5.1.1.1	Assent	N/A - Consent does not apply [1]	N/A	[1] NCCR Meeting
5.1.1.2	Consent	N/A - Consent does not apply [1]	N/A	[1] NCCR Meeting
5.1.1.3	IRB/equivalent Privacy Board determination	Some state registries have stated that researchers seeking to use their data at the individual level would need to go through the state registry's IRB and it is not enough that NCI has a process for reviewing requests (NCCR is working through this). [1]	IRB approval from the state registry authorizes data use (as needed/if applicable)	[1] NCCR Meeting [2] Updated Data Release Process PPT from NCCR
5.1.1.4	Local/state/federal laws	Information not available/found	Information not available/found	
5.1.1.5	Institutional Certification	Information not available/found	Information not available/found	
5.1.1.6	Data originator agreement	Information not available/found	Information not available/found	

Dataset 5 - NCCR Data				
		Raw Language	**Interpretation**	**Source**
5.1.1.7	Repository agreements/policies	There will be a DUA for NCCR data approved by the registries and NCI as part of the SEER data access request process. NCI is in the process of planning the tiered access process for researchers and others to use NCCR data that will be similar to the existing SEER processes (see 1/17/2023 ppt attached and https://seer.cancer.gov/data/access.html). [1] When requesting access to the SEER Research Plus Data, you must agree to the NCCR Research Data Use Agreement as well. The NCCR data will be made available through SEER*Stat along with the SEER Research Plus Data. [2]	NCCR Data Use Agreement authorizes data use	[1] NCCR Meeting [2] https://seer.cancer.gov/data-software/documentation/seerstat/nov2022/nccr-dua-nov2022.html (Accessed: 4/28/23)
5.1.1.8	Other (specify)	Information not available/found	Information not available/found	
5.1.2	**Applicable Regulations/Policies**			
5.1.2.1	Local regulations/policies	Information not available/found	Information not available/found	
5.1.2.2	Tribal regulations/policies	N/A - No tribal data. [1]	N/A	[1] NCCR Meeting
5.1.2.3	State regulations/policies	Information not available/found	Information not available/found	
5.1.2.4	Federal regulations/policies	Information not available/found	Information not available/found	
5.1.2.5	International regulations/policies	Information not available/found	Information not available/found	
5.1.2.6	Contractual obligations	Information not available/found	Information not available/found	
5.1.2.7	Repository policies	There will be a DUA for NCCR data approved by the registries and NCI as part of the SEER data access request process. NCI is in the process of planning the tiered access process for researchers and others to use NCCR data that will be similar to the existing SEER processes (see 1/17/2023 ppt attached and https://seer.cancer.gov/data/access.html). [1] When requesting access to the SEER Research Plus Data, you must agree to the NCCR Research Data Use Agreement as well. The NCCR data will be made available through SEER*Stat along with the SEER Research Plus Data. [2]	NCCR policy	[1] NCCR Meeting [2] https://seer.cancer.gov/data-software/documentation/seerstat/nov2022/nccr-dua-nov2022.html (Accessed: 4/28/23)
5.2	**Governance for data linkage, sharing, access, and use based on data access authorization or applicable regulations/policies (i.e., the origin of the governance)**			
5.2.1	Whether the data can be linked	Information not available/found	Information not available/found	
5.2.2	With what other data can it be linked or can it not be linked (scope of linkage)	Information not available/found	Information not available/found	
5.2.3	Whether data can be shared	Information not available/found	Information not available/found	
5.2.4	How data can be shared (de-identification status, disclosure review)	Information not available/found	Information not available/found	
5.2.5	How data can be accessed (access type, data use agreement, data access committee/group approval, IRB LOD, etc.)	There will be a DUA for NCCR data approved by the registries and NCI as part of the SEER data access request process. NCI is in the process of planning the tiered access process for researchers and others to use NCCR data that will be similar to the existing SEER processes (see 1/17/2023 ppt attached and https://seer.cancer.gov/data/access.html). Researchers seeking individual-level information and combinations of datasets and fields that increase the risk of re-identifiability (census-tract information or geographically-associated derived variables like the Yost index) will be required to submit a data analysis plan and possibly IRB-approval. NCI has a central IRB (BRANY) for this purpose if investigators do not use their own IRB. [1] If researcher puts in a request that requires IRB approval, they can go through the central IRB for NCCR. Each project has its own DUA (if necessary) and NCCR tracks limitations or restrictions moving forward for researchers looking to obtain data, there will be a NCCR DUA (which will be similar to the SEER DUA). [1] Currently, the process is planned as researchers submit an institutional account (known by eRA Commons) and have a Signing Official verified by NCI's Office of Data Sharing and the Surveillance Research Program for access to SEER*Stat or the NCCR Data Platform. If they require individual-level data through the Data Platform there will be an additional process to provide a data analysis plan and more information to approve release of a custom dataset. For additional information about what data will be in the Data Platform in the first year, see the sample codebook. [1] An NCCR dataset will be available in SEER*Stat that cannot be downloaded but researchers can compute various statistics in that software. Data providers (claims, PPCR, etc.) and registries sign various data sharing agreements and approve protocols to share data... SEER Stat will construct a standard registry dataset for all the registry data coming in, which will be updated annually and available via SEER*Stat to approved researchers (but download of the data won't be permitted – SEER*Stat is not an enclave, but a client server software. [1]	NCCR specifies that for data access, a user: 1. Must submit a Data Analysis Plan 2. Must execute the NCCR DUA 3. Must obtain review and approval by the NCI Office of Data Sharing and the Surveillance Research Program's Data Release group on the proposed research 4. Must obtain IRB LOD from the researcher's institution, or from NCI central IRB (BRANY) if the researchers institution does not have an IRB (as needed / applicable) 5. Must obtain IRB approval from the state registry 6. Must submit using an insititutional account (known as eRA Commons) and obtain verification of Signing Official by the NCI Office of Data Sharing and the Surveillance Research Program's Data Release group 7. Must access and use the data within SEER*Stat (client server software) or NCCR Data Platform	[1] NCCR Meeting [2] Updated Data Release Process PPT from NCCR [3] NCCR Email Communication

Dataset 5 - NCCR Data				
		Raw Language	**Interpretation**	**Source**
		Central IRB Review (slide 6): Likely no IRB review required from users of Tier 1-3 for SEER Data Products -Awaiting review and decision of the BRANY cIRB -Will have similar review for NCCR Data Products to determine need for IRB review IRB might be required by the user's home institution -NCI SRP retained the services of Biomedical Research Alliance of New York (BRANY) -Offers the services of a central IRB. [2] Must access and use the data within SEER*Stat (client server software) or NCCR Data Platform. [3]		
5.2.6	How data can be used (data use limitations)	National Childhood Cancer Registry (NCCR) Research Data Use Agreement: Use and Disclosure Restricted: Use and Disclosure Restricted: Authorized User will use or disclose the Data only for the purposes for which they were supplied. Requests for data with increased risk to re-identify individuals will require additional application materials and Authorized User will only use data for the purposes approved by the NCI. [1]	NCCR Data Use Agreement specifies that Authorized User will use or disclose the Data only for the purposes for approved research	[1] https://seer.cancer.gov/data-software/documentation/seerstat/nov2022/nccr-dua-nov2022.html (Accessed: 4/28/23)
5.2.7	Other (specify)	Information not available/found	Information not available/found	
6	**PII Elements**			
6.1	PII elements collected	In general, linkages are performed using patient name, address, SSN, phone number, cancer diagnosis, date of diagnosis. PII is also not captured in the NCCR. PII is captured by registries as part of normal state-mandated public health surveillance. PII is stripped from submissions by the registries to the NCI for the NCCR. Other data providers (e.g., claims, PPCR, etc.) also provide PII to perform patient matching and IMS performs matching in isolated parts of the data center owned by registries. IMS does maintain a PII-based file for systematic approaches to national-level patient-matching for linkages where patients are likely to move around and need to be de-deuplicated across state lines. [1]	Patient name, address, SSN, phone number, cancer diagnosis, and date of diagnosis are collected but not held in NCCR	[1] NCCR Meeting
6.2	PII elements holder (i.e., party that holds the PII)	Please also confirm our understanding that the PII resides with the individual registries (for both SEER and NCPR registries), and that the honest broker (IMS) performs the linkage with ISA, DUA or DSA, and BAAs in place. Correct.	Data originator (NPCR registries, SEER registries, and other providers)	[1] NCCR Meeting
6.3	Use of common data model, if any, for data collection	Registry data conforms to the NAACCR layout. IMS transforms data to match common SEER data elements, like the SEER recode of cancer diagnoses. When a data provider holds data with a standard then we work them to use that standard, for example, radiotherapy in the ASTRO minimal dataset. Some cancer centers (https://cancercontrol.cancer.gov/research-emphasis/supplement/childhood-cancer-registry plus MSK and Children's Hospital of Atlanta) have provided data in OMOP and we have an ongoing data harmonization effort to map all cancer center data received to OMOP for release in the data platform. When OMOP does not have concepts from historical data in clinical and genomic information systems we will use caDSR, PCDC data dictionaries, etc. ExtractEHR does not align to a common data model at this time. NCCR will consider using OMOP for SEER submissions. [1]	OMOP is used by some cancer centers provding data	[1] NCCR Meeting
7	**Prior Data Linkages**			
7.1	**Dataset linked with other datasets**			
7.1.1	Name of other linked dataset	N/A	N/A	
7.1.2	Other dataset type (clinical, EHR, survey, claims, SDOH, etc.)	NCCR (and IMS) uses PII to perform data linkage using Match Pro (originally developed for NCCR; which will be soon used for P3RL in the next few years). In addition, the NCCR data platform did use HealthVerity's P3RL mechanism to link some registry data to Medicaid claims (this data will also be available through the NCCR data platform), but because the way HealthVerity obtains the data, the data cannot go back into the registry's isolated enclaves/data centers in an identified format, rather, it the data can only be used de-identified (due to the Medicaid data – so that data will be de-identified in the NCCR data platform). [1] Registry, Claims, EHR, SDOH [2]	Registry, Claims, EHR, SDOH	[1] NCCR Meeting [2] NCCR Email Communication

Dataset 5 - NCCR Data				
		Raw Language	**Interpretation**	**Source**
7.1.3	Other dataset source(s)	Below is a table of the VPR-only registries that are participating in linkage with the NCCR. These nine registries have a Data Sharing Agreement with IMS allowing IMS to link the data and agreeing for data on prior/subsequent cancers to be incorporated into the NCCR. The table shows whether the registry would like the full record incorporated or just a list of basic data items (which now included stage and treatment). It also identifies whether the registry would like IMS to extract the data for incorporation into the NCCR. [1] COG, PPCR, Pharmacy providers, Medical insurance providers, Data aggregators using PPRL [2]	Participating NCCR registries that are authorized to link data, COG, PPCR, Pharmacy providers, Medical insurance providers, Data aggregators using PPRL	[1] NCCR Meeting [2] NCCR Email Communication
7.1.4	Linking methodology (PPRL or non-PPRL); linkage technology	NCCR (and IMS) uses PII to perform data linkage using Match Pro (originally developed for NCCR; which will be soon used for P3RL in the next few years). In addition, the NCCR data platform did use HealthVerity's P3RL mechanism to link some registry data to Medicaid claims (this data will also be available through the NCCR data platform), but because the way HealthVerity obtains the data, the data cannot go back into the registry's isolated enclaves/data centers in an identified format, rather, it the data can only be used de-identified (due to the Medicaid data – so that data will be de-identified in the NCCR data platform). [1] Match*Pro is used for data linkage. PPRL methods when required. Match*Pro will have PPRL technology for matches in the future. [2]	MatchPro or PPRL (when required)	[1] NCCR Meeting [2] NCCR Email Communication
7.1.5	PII elements used for the linkage	In general, linkages are performed using patient name, address, SSN, phone number, cancer diagnosis, date of diagnosis. PII is also not captured in the NCCR. PII is captured by registries as part of normal state-mandated public health surveillance. PII is stripped from submissions by the registries to the NCI for the NCCR. Other data providers (e.g., claims, PPCR, etc.) also provide PII to perform patient matching and IMS performs matching in isolated parts of the data center owned by registries. IMS does maintain a PII-based file for systematic approaches to national-level patient-matching for linkages where patients are likely to move around and need to be de-deuplicated across state lines. [1] Patient name, address, SSN, phone number, cancer diagnosis, and date of diagnosis are used for linkage when data are held in the registry enclaves. [2]	Patient name, address, SSN, phone number, cancer diagnosis, and date of diagnosis are used for linkage when data are held in the registry enclaves	[1] NCCR Meeting [2] NCCR Email Communication
7.1.6	Entity resolver (data originator or data linker or third party)	Information Management Services, Inc (IMS) uses PII to perform data linkage using Match Pro. [1] We have also worked with HealthVerity. [2]	IMS (Honest Broker) or HealthVerity	[1] NCCR Meeting [2] NCCR Email Communication
7.1.7	Party performing the linkages	SEER registries have a contractual relationship with the NCI to license de-identified data to the NCI. With NPCR registries, the direct relationship is with NAACCR and Information Management Services, Inc (IMS) (serves as the honest broker) to hold data. [1]	IMS (Honest Broker)	[1] NCCR Meeting
7.1.8	Linkage quality assessment	We have routine practices to evaluate patient matching. Sometimes manual review is done by IMS or registries.	Manual review or other linkage quality assessment methods	[2] NCCR Email Communication
7.1.9	Linked data sharing method (linkage maps or pre-linked dataset)	Information not available/found	Information not available/found	

USE CASE 2 - GOVERNANCE INFORMATION				
Use Case 2: What is the impact of COVID-19 infection on pediatric cancer survivors? Or what is the impact of COVID-19 infection on future pediatric cancer outcomes?				
Dataset 6 - COVID-19 Case Surveillance Restricted Access Data				
	Dataset Source	CDC COVID Data Tracker		
	Dataset Source Agency	CDC		
	Dataset Type (Clinical, EHR, Survey, SDOH, etc.)	Clinical/Case surveillance data		
	Information Sources	Website; CDC staff		
Dataset 6 - COVID-19				
		Raw Language	**Interpretation**	**Source**
1	**Data Collection**			
1.1	**Authorizations and Applicable Regulations/Policies**			
1.1.1	**Authorizations**			
1.1.1.1	Assent	N/A - COVID-19 case surveillance data are collected by jurisdictions and are shared voluntarily with CDC. [1]	N/A	[1] https://data.cdc.gov/Case-Surveillance/COVID-19-Case-Surveillance-Restricted-Access-Detai/mbd7-r32t (Accessed: 4/26/23)
1.1.1.2	Consent	N/A - COVID-19 case surveillance data are collected by jurisdictions and are shared voluntarily with CDC. [1]	N/A	[1] https://data.cdc.gov/Case-Surveillance/COVID-19-Case-Surveillance-Restricted-Access-Detai/mbd7-r32t (Accessed: 4/26/23)
1.1.1.3	IRB/equivalent Privacy Board determination	N/A - COVID-19 case surveillance data are collected by jurisdictions and are shared voluntarily with CDC. [1]	N/A	[1] https://data.cdc.gov/Case-Surveillance/COVID-19-Case-Surveillance-Restricted-Access-Detai/mbd7-r32t (Accessed: 4/26/23)
1.1.1.4	Local/state/federal laws	COVID-19 is a mandatory reportable condition in all U.S. state health departments, several territorial health departments, and two local health departments (New York City and District of Columbia). These state, territorial, and local health departments determine what information laboratories and health care providers in their areas are asked to collect. The state, territorial and local health departments confirm cases of COVID-19 based on national standardized criteria and may gather additional information on the cases reported. The data elements can be found on the Human Infection with 2019 Novel Coronavirus Case Report Form (CRF). [1]	State laws/regulations authorizes data collection	[1] https://data.cdc.gov/api/views/mbd7-r32t/files/9aad836e-5aa5-4047-aa5c-15996becc87c?download=true&filename=summary_guidance_and_limitations_information_and_restricted_access_data_use_agreement_information_updated.pdf (Accessed: 4/26/23)
1.1.1.5	Institutional Certification	N/A - COVID-19 case surveillance data are collected by jurisdictions and are shared voluntarily with CDC. [1]	N/A	[1] https://data.cdc.gov/Case-Surveillance/COVID-19-Case-Surveillance-Restricted-Access-Detai/mbd7-r32t (Accessed: 4/26/23)
1.1.1.6	Data originator agreement	N/A - COVID-19 case surveillance data are collected by jurisdictions and are shared voluntarily with CDC. [1]	N/A	[1] https://data.cdc.gov/Case-Surveillance/COVID-19-Case-Surveillance-Restricted-Access-Detai/mbd7-r32t (Accessed: 4/26/23)
1.1.1.7	Repository agreements/policies	N/A - COVID-19 case surveillance data are collected by jurisdictions and are shared voluntarily with CDC. [1]	N/A	[1] https://data.cdc.gov/Case-Surveillance/COVID-19-Case-Surveillance-Restricted-Access-Detai/mbd7-r32t (Accessed: 4/26/23)
1.1.1.8	Other (specify)	N/A - COVID-19 case surveillance data are collected by jurisdictions and are shared voluntarily with CDC. [1]	N/A	[1] https://data.cdc.gov/Case-Surveillance/COVID-19-Case-Surveillance-Restricted-Access-Detai/mbd7-r32t (Accessed: 4/26/23)
1.1.2	**Applicable Regulations/Policies**			
1.1.2.1	Local regulations/policies	Information not available/found	Information not available/found	
1.1.2.2	Tribal regulations/policies	Information not available/found	Information not available/found	
1.1.2.3	State regulations/policies	Information not available/found	Information not available/found	
1.1.2.4	Federal regulations/policies	Information not available/found	Information not available/found	
1.1.2.5	International regulations/policies	N/A - COVID-19 is a nationally notifiable disease. [1]	N/A	[1] https://www.cdc.gov/coronavirus/2019-ncov/php/reporting-pui.html (Accessed: 4/28/23)
1.1.2.6	Contractual obligations	Information not available/found	Information not available/found	

Dataset 6 - COVID-19				
		Raw Language	Interpretation	Source
1.1.2.7	Repository policies	Information not available/found	Information not available/found	
1.2	**Governance for data linkage, sharing, access, and use based on data collection authorization or applicable regulations/policies (i.e., the origin of the governance)**			
1.2.1	Whether the data can be linked	Information not available/found	Information not available/found	
1.2.2	With what other data can it be linked or can it not be linked (scope of linkage)	Information not available/found	Information not available/found	
1.2.3	Whether data can be shared	Information not available/found	Information not available/found	
1.2.4	How data can be shared (de-identification status, disclosure review)	Information not available/found	Information not available/found	
1.2.5	How data can be accessed (access type, data use agreement, data access committee/group approval, IRB LOD, etc.)	Information not available/found	Information not available/found	
1.2.6	How data can be used (data use limitations)	Information not available/found	Information not available/found	
1.2.7	Other (specify)	Information not available/found	Information not available/found	
2	**Data Linkage**			
2.1	**Authorizations and Applicable Regulations/Policies**			
2.1.1	**Authorizations**			
2.1.1.1	Assent	Information not available/found	Information not available/found	
2.1.1.2	Consent	Information not available/found	Information not available/found	
2.1.1.3	IRB/equivalent Privacy Board determination	Information not available/found	Information not available/found	
2.1.1.4	Local/state/federal laws	Information not available/found	Information not available/found	
2.1.1.5	Institutional Certification	Information not available/found	Information not available/found	
2.1.1.6	Data originator agreement	Information not available/found	Information not available/found	
2.1.1.7	Repository agreements/policies	Information not available/found	Information not available/found	
2.1.1.8	Other (specify)	Information not available/found	Information not available/found	
2.1.2	**Applicable Regulations/Policies**			
2.1.2.1	Local regulations/policies	Information not available/found	Information not available/found	
2.1.2.2	Tribal regulations/policies	Information not available/found	Information not available/found	
2.1.2.3	State regulations/policies	Information not available/found	Information not available/found	
2.1.2.4	Federal regulations/policies	Information not available/found	Information not available/found	
2.1.2.5	International regulations/policies	N/A - COVID-19 is a nationally notifiable disease. [1]	N/A	[1] https://www.cdc.gov/coronavirus/2019-ncov/php/reporting-pui.html (Accessed: 4/28/23)
2.1.2.6	Contractual obligations	Information not available/found	Information not available/found	
2.1.2.7	Repository policies	Information not available/found	Information not available/found	
2.2	**Governance for data linkage, sharing, access, and use based on data linkage authorization or applicable regulations/policies (i.e., the origin of the governance)**			
2.2.1	Whether the data can be linked	Information not available/found	Information not available/found	
2.2.2	With what other data can it be linked or can it not be linked (scope of linkage)	Information not available/found	Information not available/found	
2.2.3	Whether data can be shared	Information not available/found	Information not available/found	
2.2.4	How data can be shared (de-identification status, disclosure review)	Information not available/found	Information not available/found	
2.2.5	How data can be accessed (access type, data use agreement, data access committee/group approval, IRB LOD, etc.)	Information not available/found	Information not available/found	
2.2.6	How data can be used (data use limitations)	Information not available/found	Information not available/found	
2.2.7	Other (specify)	Information not available/found	Information not available/found	
3	**Data Sharing**			
3.1	**Authorizations and Applicable Regulations/Policies**			
3.1.1	**Authorizations**			
3.1.1.1	Assent	Information not available/found	Information not available/found	
3.1.1.2	Consent	Information not available/found	Information not available/found	
3.1.1.3	IRB/equivalent Privacy Board determination	Information not available/found	Information not available/found	

Dataset 6 - COVID-19				
		Raw Language	**Interpretation**	**Source**
3.1.1.4	Local/state/federal laws	Information not available/found	Information not available/found	
3.1.1.5	Institutional Certification	Information not available/found	Information not available/found	
3.1.1.6	Data originator agreement	Information not available/found	Information not available/found	
3.1.1.7	Repository agreements/policies	Information not available/found	Information not available/found	
3.1.1.8	Other	Information not available/found	Information not available/found	
3.1.2	**Applicable Regulations/Policies**			
3.1.2.1	Local regulations/policies	Information not available/found	Information not available/found	
3.1.2.2	Tribal regulations/policies	Information not available/found	Information not available/found	
3.1.2.3	State regulations/policies	Information not available/found	Information not available/found	
3.1.2.4	Federal regulations/policies	Information not available/found	Information not available/found	
3.1.2.5	International regulations/policies	N/A - COVID-19 is a nationally notifiable disease. [1]	N/A	[1] https://www.cdc.gov/coronavirus/2019-ncov/php/reporting-pui.html (Accessed: 4/28/23)
3.1.2.6	Contractual obligations	Information not available/found	Information not available/found	
3.1.2.7	Repository policies	This case surveillance publicly available dataset has 33 elements for all COVID-19 cases shared with CDC and includes demographics, geography (county and state of residence), any exposure history, disease severity indicators and outcomes, and presence of any underlying medical conditions and risk behaviors. [1]	CDC COVID-19 case surveillance restricted access data policy	[1] https://data.cdc.gov/api/views/mbd7-r32t/files/9aad836e-5aa5-4047-aa5c-15996becc87c?download=true&filename=summary_guidance_and_limitations_information_and_restricted_access_data_use_agreement_information_updated.pdf
3.2	**Governance for data linkage, sharing, access, and use based on data sharing authorization or applicable regulations/policies (i.e., the origin of the governance)**			
3.2.1	Whether the data can be linked	Information not available/found	Information not available/found	
3.2.2	With what other data can it be linked or can it not be linked (scope of linkage)	Information not available/found	Information not available/found	
3.2.3	Whether data can be shared	Information not available/found	Information not available/found	
3.2.4	How data can be shared (de-identification status, disclosure review)	This case surveillance publicly available dataset has 33 elements for all COVID-19 cases shared with CDC and includes demographics, geography (county and state of residence), any exposure history, disease severity indicators and outcomes, and presence of any underlying medical conditions and risk behaviors. [1] To prevent the release of data that could be used to identify persons, data cells are suppressed for low frequency (< 5) records. Records are never removed from the dataset, but individual field values are suppressed for geographic areas with low reporting counts (in the restricted access dataset) or rare combinations of demographic characteristics (sex, age group, race, ethnicity) (in both the restricted access and public use datasets). Suppressed values are re-coded to the NA answer option. [1] No direct identifiers or characteristics that might lead to identification have been included in the data provided. [2]	RIDURA specifies that COVID-19 data shared from various jurisdictions: 1. Must be de-identified of all direct identifiers prior to sharing 2. Must undergo disclosure review to suppress data fields with low frequency (<5) prior to sharing	[1] https://data.cdc.gov/Case-Surveillance/COVID-19-Case-Surveillance-Restricted-Access-Detai/mbd7-r32t [2] https://data.cdc.gov/api/views/mbd7-r32t/files/9aad836e-5aa5-4047-aa5c-15996becc87c?download=true&filename=summary_guidance_and_limitations_information_and_restricted_access_data_use_agreement_information_updated.pdf
3.2.5	How data can be accessed (access type, data use agreement, data access committee/group approval, IRB LOD, etc.)	Information not available/found	Information not available/found	
3.2.6	How data can be used (data use limitations)	Information not available/found	Information not available/found	
3.2.7	Other (specify)	Information not available/found	Information not available/found	
4	**Data Access**			
4.1	**Authorizations and Applicable Regulations/Policies**			

Dataset 6 - COVID-19				
		Raw Language	**Interpretation**	**Source**
4.1.1	**Authorizations**			
4.1.1.1	Assent	Information not available/found	Information not available/found	
4.1.1.2	Consent	Information not available/found	Information not available/found	
4.1.1.3	IRB/equivalent Privacy Board determination	Information not available/found	Information not available/found	
4.1.1.4	Local/state/federal laws	Information not available/found	Information not available/found	
4.1.1.5	Institutional Certification	Information not available/found	Information not available/found	
4.1.1.6	Data originator agreement	Information not available/found	Information not available/found	
4.1.1.7	Repository agreements/policies	Please review the following documents to determine your interest in accessing the COVID-19 Case Surveillance Restricted Access Detailed Data file: 1) CDC COVID-19 Case Surveillance Restricted Access Detailed Data: Summary, Guidance, Limitations Information, and Restricted Access Data Use Agreement Information 2) Data Dictionary for the COVID-19 Case Surveillance Restricted Access Detailed Data The next step is to complete the Registration Information and Data Use Restrictions Agreement (RIDURA). Once complete, CDC will review your agreement. After access is granted, Ask SRRG (eocevent394@cdc.gov) will email you information about how to access the data through GitHub. If you have questions about obtaining access, email eocevent394@cdc.gov. [1]	Registration Information and Data Use Restrictions Agreement (RIDURA) authorizes data access	[1] https://data.cdc.gov/Case-Surveillance/COVID-19-Case-Surveillance-Restricted-Access-Detai/mbd7-r32t [2] https://data.cdc.gov/api/views/mbd7-r32t/files/9aad836e-5aa5-4047-aa5c-15996becc87c?download=true&filename=summary_guidance_and_limitations_information_and_restricted_access_data_use_agreement_information_updated.pdf
4.1.1.8	Other (specify)	Information not available/found	Information not available/found	
4.1.2	**Applicable Regulations/Policies**			
4.1.2.1	Local regulations/policies	Information not available/found	Information not available/found	
4.1.2.2	Tribal regulations/policies	Information not available/found	Information not available/found	
4.1.2.3	State regulations/policies	Information not available/found	Information not available/found	
4.1.2.4	Federal regulations/policies	Information not available/found	Information not available/found	
4.1.2.5	International regulations/policies	Information not available/found	Information not available/found	
4.1.2.6	Contractual obligations	Information not available/found	Information not available/found	
4.1.2.7	Repository policies	Please review the following documents to determine your interest in accessing the COVID-19 Case Surveillance Restricted Access Detailed Data file: 1) CDC COVID-19 Case Surveillance Restricted Access Detailed Data: Summary, Guidance, Limitations Information, and Restricted Access Data Use Agreement Information 2) Data Dictionary for the COVID-19 Case Surveillance Restricted Access Detailed Data The next step is to complete the Registration Information and Data Use Restrictions Agreement (RIDURA). Once complete, CDC will review your agreement. After access is granted, Ask SRRG (eocevent394@cdc.gov) will email you information about how to access the data through GitHub. If you have questions about obtaining access, email eocevent394@cdc.gov. [1]	Registration Information and Data Use Restrictions Agreement (RIDURA) authorizes data access	[1] https://data.cdc.gov/Case-Surveillance/COVID-19-Case-Surveillance-Restricted-Access-Detai/mbd7-r32t [2] https://data.cdc.gov/api/views/mbd7-r32t/files/9aad836e-5aa5-4047-aa5c-15996becc87c?download=true&filename=summary_guidance_and_limitations_information_and_restricted_access_data_use_agreement_information_updated.pdf
4.2	**Governance for data linkage, sharing, access, and use based on data access authorization or applicable regulations/policies (i.e., the origin of the governance)**			
4.2.1	Whether the data can be linked	Information not available/found	Information not available/found	
4.2.2	With what other data can it be linked or can it not be linked (scope of linkage)	Information not available/found	Information not available/found	
4.2.3	Whether data can be shared	Information not available/found	Information not available/found	

Dataset 6 - COVID-19				
		Raw Language	**Interpretation**	**Source**
4.2.4	How data can be shared (de-identification status, disclosure review)	Information not available/found	Information not available/found	
4.2.5	How data can be accessed (access type, data use agreement, data access committee/group approval, IRB LOD, etc.)	Please review the following documents to determine your interest in accessing the COVID-19 Case Surveillance Restricted Access Detailed Data file: 1) CDC COVID-19 Case Surveillance Restricted Access Detailed Data: Summary, Guidance, Limitations Information, and Restricted Access Data Use Agreement Information 2) Data Dictionary for the COVID-19 Case Surveillance Restricted Access Detailed Data The next step is to complete the Registration Information and Data Use Restrictions Agreement (RIDURA). Once complete, CDC will review your agreement. After access is granted, Ask SRRG (eocevent394@cdc.gov) will email you information about how to access the data through GitHub. If you have questions about obtaining access, email eocevent394@cdc.gov. [1] RIDURA: 2. I am responsible for obtaining Institutional Review Board review of projects when appropriate. 3. Access and use of the data and/or information does not grant me permission to use any trade names trademarks, services marks, product names, or logos of CDC or the Department of Health and Human Services, except as may be required for reasonable and customary use in describing the CDC or the data and/or information. I will obtain express written approval from CDC prior to any use of the aforementioned. Though I agree to identify CDC as the source of the data provided, I further agree to not imply or state in	CDC COVID-19 case surveillance restricted access data policy specifies that users review the following documents to determine interest in accessing the COVID-19 Case Surveillance Restricted Access Detailed Data file: 1. CDC COVID-19 Case Surveillance Restricted Access Detailed Data: Summary, Guidance, Limitations Information, and Restricted Access Data Use Agreement Information 2. Data Dictionary for the COVID-19 Case Surveillance Restricted Access Detailed Data 3. After reviewing the above information, the user must complete the RIDURA, which specifies that for data access the user: (a). Obtain IRB LOD from researcher's institution (as needed /applicable) (b). Must access the data through GitHub private repository	[1] https://data.cdc.gov/api/views/mbd7-r32t/files/9aad836e-5aa5-4047-aa5c-15996becc87c?download=true&filename=summary_guidance_and_limitations_information_and_restricted_access_data_use_agreement_information_updated.pdf (Accessed: 4/28/23)

Dataset 6 - COVID-19				
		Raw Language	**Interpretation**	**Source**
		any written form, that use of or any interpretation based on the data are those of the original data sources or of CDC. 4. I understand that use of these data does not imply endorsement by CDC. I will not attribute any analysis conducted using these data to CDC. 5. I agree that while matching cases for public health purposes is acceptable, I will not deliberately participate in or support the combination of case surveillance data sets with other data sets for the specific purpose of matching records to identify individuals. 6. I understand that CDC has taken all reasonable steps for privacy protections to ensure the identity of data subjects cannot be disclosed. No direct identifiers or characteristics that might lead to identification have been included in the data provided. As such, I will not use the data to re-identify or attempt to re-identify any individual included in the data and will not use, publish or release the data in any personally identifiable form. Should I inadvertently re-identify an individual, I will notify CDC of such re-identification within three (3) days of any such discovery. [2] Your GitHub ID will be granted access to a private repository containing data that we use to make it easier to share data with you. If you do not have an ID, you can create one for free at GitHub.com. After review, you will receive an email invitation from a CDC staff member. [2]		
4.2.6	How data can be used (data use limitations)	Proposed use of the data: -Title of Analysis -Brief description of proposed analysis -Purpose of analysis / Public health significance -Describe the intended products from this analysis [1]	RIDURA specifies that the COVID-19 Case Surveillance Restricted Data Access can be used for broad research (must be of public health significance)	[1] https://data.cdc.gov/api/views/mbd7-r32t/files/9aad836e-5aa5-4047-aa5c-15996becc87c?download=true&filename=summary_guidance_and_limitations_information_and_restricted_access_data_use_agreement_information_updated.pdf (Accessed: 4/26/23)
4.2.7	Other (specify)	Information not available/found	Information not available/found	
5	**Data Use**			
5.1	**Authorizations and Applicable Regulations/Policies**			
5.1.1	**Authorizations**			
5.1.1.1	Assent	Information not available/found	Information not available/found	
5.1.1.2	Consent	Information not available/found	Information not available/found	
5.1.1.3	IRB/equivalent Privacy Board determination	Information not available/found	RIDURA specifies that for data access, a user must obtain IRB LOD from researcher's institution (as needed / applicable)	[1] https://data.cdc.gov/api/views/mbd7-r32t/files/9aad836e-5aa5-4047-aa5c-15996becc87c?download=true&filename=summary_guidance_and_limitations_information_and_restricted_access_data_use_agreement_information_updated.pdf (Accessed: 4/26/23)
5.1.1.4	Local/state/federal laws	Information not available/found	Information not available/found	
5.1.1.5	Institutional Certification	Information not available/found	Information not available/found	
5.1.1.6	Data originator agreement	Information not available/found	Information not available/found	

Dataset 6 - COVID-19				
		Raw Language	**Interpretation**	**Source**
5.1.1.7	Repository agreements/policies	Please review the following documents to determine your interest in accessing the COVID-19 Case Surveillance Restricted Access Detailed Data file: 1) CDC COVID-19 Case Surveillance Restricted Access Detailed Data: Summary, Guidance, Limitations Information, and Restricted Access Data Use Agreement Information 2) Data Dictionary for the COVID-19 Case Surveillance Restricted Access Detailed Data The next step is to complete the Registration Information and Data Use Restrictions Agreement (RIDURA). Once complete, CDC will review your agreement. After access is granted, Ask SRRG (eocevent394@cdc.gov) will email you information about how to access the data through GitHub. If you have questions about obtaining access, email eocevent394@cdc.gov. [1]	Registration Information and Data Use Restrictions Agreement (RIDURA) authorizes data use	[1] https://data.cdc.gov/Case-Surveillance/COVID-19-Case-Surveillance-Restricted-Access-Detai/mbd7-r32t (Accessed: 4/28/23)
5.1.1.8	Other (specify)	Information not available/found	Information not available/found	
5.1.2	**Applicable Regulations/Policies**			
5.1.2.1	Local regulations/policies	Information not available/found	Information not available/found	
5.1.2.2	Tribal regulations/policies	Information not available/found	Information not available/found	
5.1.2.3	State regulations/policies	Information not available/found	Information not available/found	
5.1.2.4	Federal regulations/policies	Information not available/found	Information not available/found	
5.1.2.5	International regulations/policies	Information not available/found	Information not available/found	
5.1.2.6	Contractual obligations	Information not available/found	Information not available/found	
5.1.2.7	Repository policies	Please review the following documents to determine your interest in accessing the COVID-19 Case Surveillance Restricted Access Detailed Data file: 1) CDC COVID-19 Case Surveillance Restricted Access Detailed Data: Summary, Guidance, Limitations Information, and Restricted Access Data Use Agreement Information 2) Data Dictionary for the COVID-19 Case Surveillance Restricted Access Detailed Data The next step is to complete the Registration Information and Data Use Restrictions Agreement (RIDURA). Once complete, CDC will review your agreement. After access is granted, Ask SRRG (eocevent394@cdc.gov) will email you information about how to access the data through GitHub. If you have questions about obtaining access, email eocevent394@cdc.gov. [1]	CDC COVID-19 case surveillance restricted access data policy	[1] https://data.cdc.gov/Case-Surveillance/COVID-19-Case-Surveillance-Restricted-Access-Detai/mbd7-r32t (Accessed: 4/28/23)
5.2	**Governance for data linkage, sharing, access, and use based on data access authorization or applicable regulations/policies (i.e., the origin of the governance)**			
5.2.1	Whether the data can be linked	Information not available/found	Information not available/found	
5.2.2	With what other data can it be linked or can it not be linked (scope of linkage)	Information not available/found	Information not available/found	
5.2.3	Whether data can be shared	Information not available/found	Information not available/found	
5.2.4	How data can be shared (de-identification status, disclosure review)	Information not available/found	Information not available/found	

Dataset 6 - COVID-19				
		Raw Language	**Interpretation**	**Source**
5.2.5	How data can be accessed (access type, data use agreement, data access committee/group approval, IRB LOD, etc.)	2. I am responsible for obtaining Institutional Review Board review of projects when appropriate. 3. Access and use of the data and/or information does not grant me permission to use any trade names trademarks, services marks, product names, or logos of CDC or the Department of Health and Human Services, except as may be required for reasonable and customary use in describing the CDC or the data and/or information. I will obtain express written approval from CDC prior to any use of the aforementioned. Though I agree to identify CDC as the source of the data provided, I further agree to not imply or state in any written form, that use of or any interpretation based on the data are those of the original data sources or of CDC. 4. I understand that use of these data does not imply endorsement by CDC. I will not attribute any analysis conducted using these data to CDC. 5. I agree that while matching cases for public health purposes is acceptable, I will not deliberately participate in or support the combination of case surveillance data sets with other data sets for the specific purpose of matching records to identify individuals. 6. I understand that CDC has taken all reasonable steps for privacy protections to ensure the identity of data subjects cannot be disclosed. No direct identifiers or characteristics that might lead to identification have been included in the data provided. As such, I will not use the data to re-identify or attempt to re-identify any individual included in the data and will not use, publish or release the data in any personally identifiable form. Should I inadvertently re-identify an individual, I will notify CDC of such re-identification within three (3) days of any such discovery. [1]	RIDURA specifies that for data access, a user: 1. Must sign and complete the RIDURA 2. Obtain IRB LOD from researcher's institution (as needed / applicable) 3. Must access the data through GitHub private repository	[1] https://data.cdc.gov/Case-Surveillance/COVID-19-Case-Surveillance-Restricted-Access-Detai/mbd7-r32t
5.2.6	How data can be used (data use limitations)	Proposed use of the data: -Title of Analysis -Brief description of proposed analysis -Purpose of analysis / Public health significance -Describe the intended products from this analysis [1]	RIDURA specifies that the COVID-19 Case Surveillance Restricted Data Access can be used for broad research (must be of public health significance)	[1] https://data.cdc.gov/api/views/mbd7-r32t/files/9aad836e-5aa5-4047-aa5c-15996becc87c?download=true&filename=summary_guidance_and_limitations_information_and_restricted_access_data_use_agreement_information_updated.pdf (Accessed: 4/26/23)
5.2.7	Other (specify)	Information not available/found	Information not available/found	
6	**PII Elements**			
6.1	PII elements collected	Human Infection with 2019 Novel Coronavirus Case Report Form (CRF) [1]	First name, last name, date of birth, sex, age (yr/mo/day), state and county of residence, race, ethnicity, and tribal name are collected from participants	[1] https://www.cdc.gov/coronavirus/2019-ncov/downloads/pui-form.pdf [2] SSRG response email
6.2	PII elements holder (i.e., party that holds the PII)	Patient identifier information is not transmitted to the CDC, and reside within the state, territorial, and/or local health departments that report these cases to the CDC. [1]	State, territorial, and/or local health departments	[1] https://www.cdc.gov/coronavirus/2019-ncov/downloads/pui-form.pdf (Accessed: 4/26/23)
6.3	Use of common data model, if any, for data collection	Information not available/found	Information not available/found	
7	**Prior Data Linkages**			
7.1	**Dataset linked with other datasets**			
7.1.1	Name of other linked dataset	Information not available/found	Information not available/found	

Dataset 6 - COVID-19				
		Raw Language	**Interpretation**	**Source**
7.1.2	Other dataset type (clinical, EHR, survey, claims, SDOH, etc.)	Information not available/found	Information not available/found	
7.1.3	Other dataset source(s)	Information not available/found	Information not available/found	
7.1.4	Linking methodology (PPRL or non-PPRL); linkage technology	Information not available/found	Information not available/found	
7.1.5	PII elements used for the linkage	Information not available/found	Information not available/found	
7.1.6	Entity resolver (data originator or data linker or third party)	Information not available/found	Information not available/found	
7.1.7	Party performing the linkages	Information not available/found	Information not available/found	
7.1.8	Linkage quality assessment	Information not available/found	Information not available/found	
7.1.9	Linked data sharing method (linkage maps or pre-linked dataset)	Information not available/found	Information not available/found	

USE CASE 2 - GOVERNANCE INFORMATION				
Use Case 2: What is the impact of COVID-19 infection on pediatric cancer survivors? Or what is the impact of COVID-19 infection on future pediatric cancer outcomes?				
Dataset 7 - T-MSIS Analytic Files (TAF)				
	Dataset Source	Transformed Medicaid Statistical Information System (T-MSIS)		
	Dataset Source Agency	CMS		
	Dataset Type (Clinical, EHR, Survey, SDOH, etc.)	Administrative/claims		
	Information Sources	Webiste, CMS staff meeting		
Dataset 7 - T-MSIS Analytic Files (TAF)				
		Raw Language	**Interpretation**	**Source**
1	**Data Collection**			
1.1	**Authorizations and Applicable Regulations/Policies**			
1.1.1	**Authorizations**		1. Section 4753 of the Balanced Budget Act of 1997 2. Section 6504 of the Patient Protection and Affordable Care Act	
1.1.1.1	Assent	N/A	N/A	
1.1.1.2	Consent	N/A	N/A	
1.1.1.3	IRB/equivalent Privacy Board determination	N/A	N/A	
1.1.1.4	Local/state/federal law	Section 4753 of the Balanced Budget Act of 1997, P.L. 105-33, amended section 1903(r) of the Act to include a statutory requirement for states to submit claims data, enrollee encounter data, and supporting information. Section 6504 of the Patient Protection and Affordable Care Act, P.L. 111-148, as amended by the Health Care and Education Reconciliation Act, P.L. 111-152 (collectively, the Affordable Care Act) strengthened this provision by requiring states to include data elements the Secretary of Health and Human Services determines necessary for program integrity, program oversight, and administration. The Medicaid managed care regulation published in May 2016 further describes the requirements for the submission of encounter data (see 42 CFR 438.242, 438.604 and 438.818). As part of encounter data reporting, CMS expects states to report all actual payment-related fields stipulated in the T-MSIS documentation and referenced in the Medicaid managed care regulations. [1]	Three Federal laws authorize data collection: 1. Section 4753 of the Balanced Budget Act of 1997 2. Section 6504 of the Patient Protection and Affordable Care Act 3. Medicaid and CHIP Managed Care Final Rule	[1] https://www.medicaid.gov/federal-policy-guidance/downloads/SHO18008.pdf (Accessed: 4/25/23) [2] CMS Staff Meeting [3] https://www.federalregister.gov/documents/2020/11/13/2020-24758/medicaid-program-medicaid-and-childrens-health-insurance-program-chip-managed-care (Accessed: 4/25/23)
1.1.1.5	Institutional Certification	N/A	N/A	
1.1.1.6	Data originator agreement	N/A	N/A	
1.1.1.7	Repository agreements/policies	N/A	N/A	
1.1.1.8	Other (specify)	Information not available/found	Information not available/found	
1.1.2	**Applicable Regulations/Policies**			
1.1.2.1	Local regulations/policies	Information not available/found	Information not available/found	
1.1.2.2	Tribal regulations/policies	Information not available/found	Information not available/found	
1.1.2.3	State regulations/policies	Information not available/found	Information not available/found	

Dataset 7 - T-MSIS Analytic Files (TAF)				
		Raw Language	**Interpretation**	**Source**
1.1.2.4	Federal regulations/policies	CMS is a covered entity under HIPAA and is subject to the Privacy Act when collecting, using, and disclosing data. (https://www.federalregister.gov/documents/2019/02/06/2019-01157/privacy-act-of-1974-system-of-records) [1]	1. Privacy Act 2. HIPAA Privacy Rule	[1] CMS Staff Meeting
1.1.2.5	International regulations/policies	Information not available/found	Information not available/found	
1.1.2.6	Contractual obligations	Information not available/found	Information not available/found	
1.1.2.7	Repository policies	Information not available/found	Information not available/found	
1.2	**Governance for data linkage, sharing, access, and use based on data collection authorization or applicable regulations/policies (i.e., the origin of the governance)**			
1.2.1	Whether the data can be linked	Does not authorize/specify	Does not authorize/specify	
1.2.2	With what other data can it be linked or can it not be linked (scope of linkage)	Does not authorize/specify	Does not authorize/specify	
1.2.3	Whether data can be shared	CMS is a covered entity under HIPAA and is subject to the Privacy Act when collecting, using, and disclosing data. (https://www.federalregister.gov/documents/2019/02/06/2019-01157/privacy-act-of-1974-system-of-records) [1]	HIPAA Privacy Rule that CMS is subject to under the Privacy Act specifies that de-identified data can be shared.	[1] CMS Staff Meeting
1.2.4	How data can be shared (de-identification status, disclosure review)	CMS is a covered entity under HIPAA and is subject to the Privacy Act when collecting, using, and disclosing data. (https://www.federalregister.gov/documents/2019/02/06/2019-01157/privacy-act-of-1974-system-of-records) [1]	HIPAA Privacy Rule that CMS is subject to under the Privacy Act specifies that T-MSIS data be de-identified of all 18 HIPAA identifiers prior to sharing.	[1] CMS Staff Meeting
1.2.5	How data can be accessed (access type, data use agreement, data access committee/group approval, IRB LOD, etc.)	Does not authorize/specify	Does not authorize/specify	
1.2.6	How data can be used (data use limitations)	Does not authorize/specify	Does not authorize/specify	
1.2.7	Other (specify)	Does not authorize/specify	Does not authorize/specify	
2	**Data Linkage**			
2.1	**Authorizations and Applicable Regulations/Policies**			
2.1.1	**Authorizations**		1. CMS Privacy Board 2. Chief Data Officer 3. Information Exchange Agreement (IEA)	
2.1.1.1	Assent	N/A	N/A	
2.1.1.2	Consent	N/A	N/A	
2.1.1.3	IRB/equivalent Privacy Board determination	For standard T-MSIS Analytic Files (TAFs), the CMS Privacy Board authorizes linkage and use of the TAFs by researchers for research purposes. Linkage cannot occur if the Privacy Board does not authorize the linkage as part of the RIF application approval. [1]	CMS Privacy Board approval of a RIF application specifying linkage authorizes data linkage	[1] CMS Staff Meeting

Dataset 7 - T-MSIS Analytic Files (TAF)				
		Raw Language	**Interpretation**	**Source**
2.1.1.4	Local/state/federal laws	Information not available/found	Information not available/found	
2.1.1.5	Institutional Certification	N/A	N/A	
2.1.1.6	Data originator agreement	N/A	N/A	
2.1.1.7	Repository agreements/policies	DATA USE AGREEMENT FOR USE OF CENTERS FOR MEDICARE AND MEDICAID SERVICES (CMS) RESEARCH IDENTIFIABLE FILES (RIFS): Linking Data. Absent express written authorization from CMS, the Requesting Organization agrees not to link or attempt to link beneficiary level records included in the file(s) listed in Attachment A – RIF Request Application to any other source of information. A RIF Request Application that includes the linkage of specific files that has been approved in accordance with section 3 constitutes express authorization from CMS to link files as described in the protocol. [1] From the RIF Application: "8. Please list any other data files or sources of information that you are planning to use to support your research study. (e.g., Provider of Services (POS) file, AMA Physician Master file, etc.). If you will be linking or attempting to link to the CMS files specified in section 5, please describe how you will be linking the data. Name of additional files Purpose for using the data file in the analysis If linking to CMS data, describe how linkage will occur [2] Data can be linked at beneficiary level to non-CMS data using a beneficiary identifier? --Yes [3] "In addition to the research, Federal agencies are permitted to create linked datasets (CMS data linked with another non-CMS data source) and provide linked datasets to outside researchers. CMS does not allow secondary use of CMS only -(non-linked) files. If the data is linked, the federal agency must take the data into their SORN and in addition to the DUA authorizing the original research disclosure, there must also be an IEA with CMS to specify the terms of the secondary use. Because the other federal agency is not a HIPAA Covered Entity, they must agree to treat the secondary release as if they are a HIPAA Covered Entity and follow a process similar to CMS for releasing the data for research purposes, including entering into a DUA with the researcher." [4]	Information Exchange Agreement (IEA) between CMS and the Participating Agency (on top of the Data Use Agreement) authorizes data linkage for federal entities performing linkage with non-standard TAFs containing PII	[1] https://resdac.org/sites/datadocumentation.resdac.org/files/2022-10/Instructions%20-%20RIF%20Data%20Use%20Agreement.pdf (Accessed: 4/25/23) [2] https://resdac.org/request-form/rif-application (Accessed: 4/25/23) [3] https://resdac.org/articles/differences-between-rif-lds-and-puf-data-files (Accessed: 4/25/23) [4] CMS Staff Meeting
2.1.1.8	Other (specify)	Information not available/found	Information not available/found	
2.1.2	**Applicable Regulations/Policies**			
2.1.2.1	Local regulations/policies	Information not available/found	Information not available/found	
2.1.2.2	Tribal regulations/policies	Information not available/found	Information not available/found	
2.1.2.3	State regulations/policies	Information not available/found	Information not available/found	

Dataset 7 - T-MSIS Analytic Files (TAF)				
		Raw Language	**Interpretation**	**Source**
2.1.2.4	Federal regulations/policies	Information not available/found	Information not available/found	
2.1.2.5	International regulations/policies	Information not available/found	Information not available/found	
2.1.2.6	Contractual obligations	Information not available/found	Information not available/found	
2.1.2.7	Repository policies	DATA USE AGREEMENT FOR USE OF CENTERS FOR MEDICARE AND MEDICAID SERVICES (CMS) RESEARCH IDENTIFIABLE FILES (RIFS): Linking Data. Absent express written authorization from CMS, the Requesting Organization agrees not to link or attempt to link beneficiary level records included in the file(s) listed in Attachment A – RIF Request Application to any other source of information. A RIF Request Application that includes the linkage of specific files that has been approved in accordance with section 3 constitutes express authorization from CMS to link files as described in the protocol. [1] From the RIF Application: "8. Please list any other data files or sources of information that you are planning to use to support your research study. (e.g., Provider of Services (POS) file, AMA Physician Master file, etc.). If you will be linking or attempting to link to the CMS files specified in section 5, please describe how you will be linking the data. Name of additional files Purpose for using the data file in the analysis If linking to CMS data, describe how linkage will occur [2] Data can be linked at beneficiary level to non-CMS data using a beneficiary identifier? --Yes [3] "In addition to the research, Federal agencies are permitted to create linked datasets (CMS data linked with another non-CMS data source) and provide linked datasets to outside researchers. CMS does not allow secondary use of CMS only - (non-linked) files. If the data is linked, the federal agency must take the data into their SORN and in addition to the DUA authorizing the original research disclosure, there must also be an IEA with CMS to specify the terms of the secondary use. Because the other federal agency is not a HIPAA Covered Entity, they must agree to treat the secondary release as if they are a HIPAA Covered Entity and follow a process similar to CMS for releasing the data for research purposes, including entering into a DUA with the researcher." [4]	CMS Research Data Center (ResDAC) policy	[1] https://resdac.org/sites/datadocumentation.resdac.org/files/2022-10/Instructions%20-%20RIF%20Data%20Use%20Agreement.pdf (Accessed: 4/25/23) [2] https://resdac.org/request-form/rif-application (Accessed: 4/25/23) [3] https://resdac.org/articles/differences-between-rif-lds-and-puf-data-files (Accessed: 4/25/23) [4] CMS Staff Meeting
2.2	**Governance for data linkage, sharing, access, and use based on data linkage authorization or applicable regulations/policies (i.e., the origin of the governance)**			

Dataset 7 - T-MSIS Analytic Files (TAF)				
		Raw Language	**Interpretation**	**Source**
2.2.1	Whether the data can be linked	DATA USE AGREEMENT FOR USE OF CENTERS FOR MEDICARE AND MEDICAID SERVICES (CMS) RESEARCH IDENTIFIABLE FILES (RIFS): Linking Data. Absent express written authorization from CMS, the Requesting Organization agrees not to link or attempt to link beneficiary level records included in the file(s) listed in Attachment A – RIF Request Application to any other source of information. A RIF Request Application that includes the linkage of specific files that has been approved in accordance with section 3 constitutes express authorization from CMS to link files as described in the protocol. [1] From the RIF Application: "8. Please list any other data files or sources of information that you are planning to use to support your research study. (e.g., Provider of Services (POS) file, AMA Physician Master file, etc.). If you will be linking or attempting to link to the CMS files specified in section 5, please describe how you will be linking the data. Name of additional files Purpose for using the data file in the analysis If linking to CMS data, describe how linkage will occur [2] Data can be linked at beneficiary level to non-CMS data using a beneficiary identifier? --Yes [3] For standard TAFs, the CMS Privacy Board authorizes linkage and use of the TAFs by researchers for research purposes. Linkage cannot occur if the Privacy Board does not authorize the linkage as part of the RIF application approval. [4] For non-standard TAFs (i.e., with additional variables), authorization includes CMS Privacy Board and the submission of a letter to the Chief Data Officer. A federal agency requesting access to non-standard research data must submit a letter of justification to the CMS Chief Data Officer (CDO) for approval. The CDO will consult with other CMS leadership where appropriate before approving a non-standard data file request (e.g., leadership from the Center for Medicaid & CHIP Services on requests related to Medicaid data). [4]	1. CMS Privacy Board authorizes data linkage for researchers for research purposes via an approved RIF Application that specifies the scope of linkage 2. CMS Privacy Board, Chief Data Officer approval of a letter of justification for linkage, and Information Exchange Agreement (IEA) authorizes data linkage for federal entities performing linkage with non-standard TAFs containing PII	[1] https://resdac.org/sites/datadocumentation.resdac.org/files/2022-10/Instructions%20-%20RIF%20Data%20Use%20Agreement.pdf (Accessed: 4/25/23) [2] https://resdac.org/request-form/rif-application (Accessed: 4/25/23) [3] https://resdac.org/articles/differences-between-rif-lds-and-puf-data-files (Accessed: 4/25/23) [4] CMS Staff Meeting

Dataset 7 - T-MSIS Analytic Files (TAF)				
		Raw Language	**Interpretation**	**Source**
2.2.2	With what other data can it be linked or can it not be linked (scope of linkage)	DATA USE AGREEMENT FOR USE OF CENTERS FOR MEDICARE AND MEDICAID SERVICES (CMS) RESEARCH IDENTIFIABLE FILES (RIFS): Linking Data. Absent express written authorization from CMS, the Requesting Organization agrees not to link or attempt to link beneficiary level records included in the file(s) listed in Attachment A – RIF Request Application to any other source of information. A RIF Request Application that includes the linkage of specific files that has been approved in accordance with section 3 constitutes express authorization from CMS to link files as described in the protocol. [1] From the RIF Application: "8. Please list any other data files or sources of information that you are planning to use to support your research study. (e.g., Provider of Services (POS) file, AMA Physician Master file, etc.). If you will be linking or attempting to link to the CMS files specified in section 5, please describe how you will be linking the data. Name of additional files Purpose for using the data file in the analysis If linking to CMS data, describe how linkage will occur [2] Data can be linked at beneficiary level to non-CMS data using a beneficiary identifier? --Yes [3] For standard TAFs, the CMS Privacy Board authorizes linkage and use of the TAFs by researchers for research purposes. Linkage cannot occur if the Privacy Board does not authorize the linkage as part of the RIF application approval. [4] Researchers can request linkages as part of research protocol. PII, such as SSN, are not released to researchers as part of the standard research files. o For linkages with external datasets (listed in the RIF application), the researchers must submit PII for linkages with the external datasets to CMS. o CMS facilitates the linkage (to BENE_ID) based on the PII provided. o CMS returns the crosswalk of BENE_ID to the requesting researcher. o The data that CMS releases does not contain SSN; CMS releases data with the identifier BENE_ID which is also used as the link key. [4]	1. CMS policy and Data Use Agreement specify that data can only be linked to external data in accordance with the RIF Application approved by the CMS Privacy Board which specifies the scope of linkage 2. Information Exchange Agreement (IEA) between CMS and Participating Agency specifies the scope of linkage for federal entities performing linkage with non-standard TAFs containing PII	[1] https://resdac.org/sites/datadocumentation.resdac.org/files/2022-10/Instructions%20-%20RIF%20Data%20Use%20Agreement.pdf (Accessed: 4/25/23) [2] https://resdac.org/request-form/rif-application (Accessed: 4/25/23) [3] https://resdac.org/articles/differences-between-rif-lds-and-puf-data-files (Accessed: 4/25/23) [4] CMS Staff Meeting

Dataset 7 - T-MSIS Analytic Files (TAF)				
		Raw Language	**Interpretation**	**Source**
		For non-standard TAFs (i.e., with additional variables), authorization includes CMS Privacy Board and the submission of a letter to the Chief Data Officer. A federal agency requesting access to non-standard research data must submit a letter of justification to the CMS Chief Data Officer (CDO) for approval. The CDO will consult with other CMS leadership where appropriate before approving a non-standard data file request (e.g., leadership from the Center for Medicaid & CHIP Services on requests related to Medicaid data). [4]		
2.2.3	Whether data can be shared	In addition to the research, Federal agencies are permitted to create linked datasets (CMS data linked with another non-CMS data source) and provide linked datasets to outside researchers. CMS does not allow secondary use of CMS only - (non-linked) files. If the data is linked, the federal agency must take the data into their SORN and in addition to the DUA authorizing the original research disclosure, there must also be an IEA with CMS to specify the terms of the secondary use. Because the other federal agency is not a HIPAA Covered Entity, they must agree to treat the secondary release as if they are a HIPAA Covered Entity and follow a process similar to CMS for releasing the data for research purposes, including entering into a DUA with the researcher. [1]	CMS policy, Data Use Agreement, and Information Exhange Agreement (IEA) authorize data sharing of linked data by federal entities but prohibit re-sharing of un-linked CMS data	[1] CMS Staff Meeting
2.2.4	How data can be shared (de-identification status, disclosure review)	In addition to the research, Federal agencies are permitted to create linked datasets (CMS data linked with another non-CMS data source) and provide linked datasets to outside researchers. CMS does not allow secondary use of CMS only - (non-linked) files. If the data is linked, the federal agency must take the data into their SORN and in addition to the DUA authorizing the original research disclosure, there must also be an IEA with CMS to specify the terms of the secondary use. Because the other federal agency is not a HIPAA Covered Entity, they must agree to treat the secondary release as if they are a HIPAA Covered Entity and follow a process similar to CMS for releasing the data for research purposes, including entering into a DUA with the researcher. [1]	1. CMS policy specifies that federal entities performing data linkage using non-standard TAFs containing PII must take the data into their SORN under the Privacy Act 2. CMS policy specifies that federal entities performing data linkage must agree to treat secondarily shared data as if the entity is a HIPAA Covered Entity and follow a process similar to CMS for releasing data including entering into a DUA with the researcher	[1] CMS Staff Meeting
2.2.5	How data can be accessed (access type, data use agreement, data access committee/group approval, IRB LOD, etc.)	Does not authorize/specify	Does not authorize/specify	

Dataset 7 - T-MSIS Analytic Files (TAF)				
		Raw Language	**Interpretation**	**Source**
2.2.6	How data can be used (data use limitations)	For standard TAFs, the CMS Privacy Board authorizes linkage and use of the TAFs by researchers for research purposes. Linkage cannot occur if the Privacy Board does not authorize the linkage as part of the RIF application approval. [1]	CMS Privacy Board specifies that TAFs can be used for resesarch purposes	[1] CMS Staff Meeting
2.2.7	Other (specify)	Does not authorize/specify	Does not authorize/specify	
3	**Data Sharing**			
3.1	**Authorizations and Applicable Regulations/Policies**			
3.1.1	**Authorizations**		Information Exchange Agreement (IEA)	
3.1.1.1	Assent	N/A	N/A	
3.1.1.2	Consent	N/A	N/A	
3.1.1.3	IRB/equivalent Privacy Board determination	Information not available/found	Information not available/found	
3.1.1.4	Local/state/federal laws	Information not available/found	Information not available/found	
3.1.1.5	Institutional Certification	Information not available/found	Information not available/found	
3.1.1.6	Data originator agreement	Information not available/found	Information not available/found	
3.1.1.7	Repository agreements/policies	INFORMATION EXCHANGE AGREEMENT BETWEEN THE CENTERS FOR MEDICARE & MEDICAID SERVICES AND Insert the Participating Agency FOR DISCLOSURE OF [Business Owner inserts brief description of exchanged data]: DESCRIPTION OF THE DATA THAT MAY BE DISCLOSED A. Data Covered by this Agreement [Business Owner list and describe the exchanges of data that correspond with the purposes and legal authority provided in the Purpose and Legal Authorities Sections.] 1. ... 2. ... B. System(s) of Records CMS will provide CMS Data from the following SOR(s): [Business Owner list all the CMS Systems of Records, with notice information and routine use information, an example is provided] 1. [Example] Enrollment Data Base (EDB), System No. 09-70-0502; last modified at 73 FR 10249 (February 26, 2008), as amended at April 23, 2013 (78 FR 23938), February 18, 2016 (81 FR 8204) and February 14, 2018 (83 FR 6591). Data maintained in the EDB will be released pursuant to routine use number 2 and 10, as set forth in the SORN. [1]	Information Exchange Agreement (IEA) between CMS and the Participating Agency (on top of the Data Use Agreement) authorizes data sharing of linked data by federal agencies performing linkage with non-standard TAFs	[1] IEA Template
3.1.1.8	Other	Information not available/found	Information not available/found	
3.1.2	**Applicable Regulations/Policies**			
3.1.2.1	Local regulations/policies	Information not available/found	Information not available/found	
3.1.2.2	Tribal regulations/policies	Information not available/found	Information not available/found	
3.1.2.3	State regulations/policies	Information not available/found	Information not available/found	

Dataset 7 - T-MSIS Analytic Files (TAF)				
		Raw Language	**Interpretation**	**Source**
3.1.2.4	Federal regulations/policies	CMS is a covered entity under HIPAA and is subject to the Privacy Act when collecting, using, and disclosing data. (https://www.federalregister.gov/documents/2019/02/06/2019-01157/privacy-act-of-1974-system-of-records) [1]	1. Privacy Act 2. HIPAA Privacy Rule	[1] CMS Staff Meeting
3.1.2.5	International regulations/policies	Information not available/found	Information not available/found	
3.1.2.6	Contractual obligations	Information not available/found	Information not available/found	
3.1.2.7	Repository policies	Information not available/found	Information not available/found	
3.2	**Governance for data linkage, sharing, access, and use based on data sharing authorization or applicable regulations/policies (i.e., the origin of the governance)**			
3.2.1	Whether the data can be linked	Information not available/found	Information not available/found	
3.2.2	With what other data can it be linked or can it not be linked (scope of	Information not available/found	Information not available/found	
3.2.3	Whether data can be shared	In addition to the research, Federal agencies are permitted to create linked datasets (CMS data linked with another non-CMS data source) and provide linked datasets to outside researchers. CMS does not allow secondary use of CMS only - (non-linked) files. If the data is linked, the federal agency must take the data into their SORN and in addition to the DUA authorizing the original research disclosure, there must also be an IEA with CMS to specify the terms of the secondary use. Because the other federal agency is not a HIPAA Covered Entity, they must agree to treat the secondary release as if they are a HIPAA Covered Entity and follow a process similar to CMS for releasing the data for research purposes, including entering into a DUA with the researcher. [1]	CMS policy, Data Use Agreement, and Information Exhange Agreement (IEA) authorize data sharing of linked data by federal entities but prohibit re-sharing of un-linked CMS data	[1] CMS Staff Meeting
3.2.4	How data can be shared (de-identification status, disclosure review)	In addition to the research, Federal agencies are permitted to create linked datasets (CMS data linked with another non-CMS data source) and provide linked datasets to outside researchers. CMS does not allow secondary use of CMS only - (non-linked) files. If the data is linked, the federal agency must take the data into their SORN and in addition to the DUA authorizing the original research disclosure, there must also be an IEA with CMS to specify the terms of the secondary use. Because the other federal agency is not a HIPAA Covered Entity, they must agree to treat the secondary release as if they are a HIPAA Covered Entity and follow a process similar to CMS for releasing the data for research purposes, including entering into a DUA with the researcher. [1]	1. HIPAA specifies that de-identified data can be shared 2. CMS policy specifies that federal entities performing data linkage using non-standard TAFs containing PII must take the data into their SORN under the Privacy Act 3. CMS policy specifies that federal entities performing data linkage must agree to treat secondarily shared data as if the entity is a HIPAA Covered Entity and follow a process similar to CMS for releasing data including entering into a DUA with the researcher	[1] CMS Staff Meeting
3.2.5	How data can be accessed (access type, data use agreement, data access committee/group approval, IRB LOD, etc.)	Information not available/found	Information not available/found	
3.2.6	How data can be used (data use limitations)	Information not available/found	Information not available/found	

Dataset 7 - T-MSIS Analytic Files (TAF)				
		Raw Language	**Interpretation**	**Source**
3.2.7	Other (specify)	Information not available/found	Information not available/found	
4	**Data Access**			
4.1	**Authorizations and Applicable Regulations/Policies**			
4.1.1	**Authorizations**		1. CMS Privacy Board 2. Data Use Agreement	
4.1.1.1	Assent	N/A	N/A	
4.1.1.2	Consent	N/A	N/A	
4.1.1.3	IRB/equivalent Privacy Board determination	Information not available/found	Information not available/found	
4.1.1.4	Local/state/federal laws	Data access is authorized through CMS Privacy Board approval (the CMS Privacy Board consults with the data owner, who reviews all data requests) and under the DUA. [1]	CMS Privacy Board authorizes data access	
4.1.1.5	Institutional Certification	Information not available/found	Information not available/found	
4.1.1.6	Data originator agreement	Information not available/found The CMS Privacy Board consults with the data owner, who reviews all data requests [1], so a data originator agreement may exist.	Information not available/found	
4.1.1.7	Repository agreements/policies	Data access is authorized through CMS Privacy Board approval (the CMS Privacy Board consults with the data owner, who reviews all data requests) and under the DUA. [1]	Data Use Agreement authorizes data access	[1] CMS Staff Meeting
4.1.1.8	Other (specify)	Information not available/found	Information not available/found	
4.1.2	**Applicable Regulations/Policies**			
4.1.2.1	Local regulations/policies	Information not available/found	Information not available/found	
4.1.2.2	Tribal regulations/policies	Information not available/found	Information not available/found	
4.1.2.3	State regulations/policies	Information not available/found	Information not available/found	

Dataset 7 - T-MSIS Analytic Files (TAF)				
		Raw Language	**Interpretation**	**Source**
4.1.2.4	Federal regulations/policies	CMS is a covered entity under HIPAA and is subject to the Privacy Act when collecting, using, and disclosing data. (https://www.federalregister.gov/documents/2019/02/06/2019-01157/privacy-act-of-1974-system-of-records) [1] CMS must ensure that all research requests for protected health information meet the requirements under the Common Rule and the Health Insurance Portability and Accountability Act (HIPAA) Privacy Rule. As a result, researchers must submit the following as part of the research request packet: Common Rule If subject to the Common Rule: » Documentation of institutional review board (IRB) approval of the research AND » Informed consent of the research subjects or IRB waiver of the requirement to obtain informed consent. If exempt from the Common Rule: » A signed and dated statement describing the basis for exemption. HIPAA Privacy Rule Individual authorization of the use of the data for the research OR Documentation that an IRB or a Privacy Board has approved a waiver of research subjects' authorization for use/disclosure of information about them for research purposes. The IRB study title must match the CMS data request study title, otherwise, a brief explanation is required. [2]	1. Privacy Act 2. HIPAA Privacy Rule 3. Common Rule	[1] CMS Staff Meeting [2] https://resdac.org/irb-common-rule-and-hipaa-waiver-approval (Accessed: 4/25/23)
4.1.2.5	International regulations/policies	Information not available/found	Information not available/found	
4.1.2.6	Contractual obligations	Information not available/found	Information not available/found	

Dataset 7 - T-MSIS Analytic Files (TAF)				
		Raw Language	**Interpretation**	**Source**
4.1.2.7	Repository policies	The Research Identifiable File (RIF) Data Use Agreement (DUA) is a legal agreement between CMS and a requesting organization that documents the terms and conditions under which the CMS data may be used, including CMS privacy and security requirements and data release policies. [1] The Research Identifiable File (RIF) application collects information about the requesting organization and the research study, including detailed study aims, data required, and dissemination of findings plan. The RIF application is used by CMS to assess the feasibility of the research and compliance with CMS data use and release policies. [2] The Key Personnel Supplement collects information about each request's key contacts including the requester, collaborating organizations and additional contacts who should be included on notices about the project. [3] The Specifications Worksheet is required for all RIF requests. It collects detailed requester information, study/project data extract details, shipping information, and method of payment. It also includes a Part D event justification tab that is required for all requests that include Part D data. The Specifications Worksheet is used by the data distributor to generate a cost invoice and to collect data extraction information.[4] CMS Information Exchange Agreement (IEA): "This is used to modify the terms of the research DUA to allow redisclosure of linked datasets by federal agencies. However, we are exploring creating a research DUA addendum with standard clauses for the redisclosure scenario instead. This document is also used by the CMS Privacy Office for other types of disclosures (other than research) or exchanges with federal agencies." [5] IEA Template [6]	CMS Research Data Center (ResDAC) policy	[1] https://resdac.org/request-form/rif-data-use-agreement (also see https://resdac.org/sites/datadocumentation.resdac.org/files/2022-10/RIF%20Data%20Use%20Agreement.pdf) (Accessed: 4/25/23)) [2] https://resdac.org/request-form/rif-application (Accessed: 4/25/23) [3] https://resdac.org/request-form/key-personnel-supplement (Accessed: 4/25/23) [4] https://resdac.org/request-form/rif-specifications-worksheet (Accessed: 4/25/23) [5] CMS Email Communication [6] IEA Template
4.2	**Governance for data linkage, sharing, access, and use based on data access authorization or applicable regulations/policies (i.e., the origin of the governance)**			
4.2.1	Whether the data can be linked	For standard TAFs, the CMS Privacy Board authorizes linkage and use of the TAFs by researchers for research purposes. Linkage cannot occur if the Privacy Board does not authorize the linkage as part of the RIF application approval. [1]	CMS Privacy Board authorizes data linkage for researchers for research purposes via an approved RIF Application that specifies the scope of linkage	[1] CMS Staff Meeting

Dataset 7 - T-MSIS Analytic Files (TAF)				
		Raw Language	**Interpretation**	**Source**
4.2.2	With what other data can it be linked or can it not be linked (scope of linkage)	For standard TAFs, the CMS Privacy Board authorizes linkage and use of the TAFs by researchers for research purposes. Linkage cannot occur if the Privacy Board does not authorize the linkage as part of the RIF application approval. [1] Researchers can request linkages as part of research protocol. PII, such as SSN, are not released to researchers as part of the standard research files. o For linkages with external datasets (listed in the RIF application), the researchers must submit PII for linkages with the external datasets to CMS. o CMS facilitates the linkage (to BENE_ID) based on the PII provided. o CMS returns the crosswalk of BENE_ID to the requesting researcher. o The data that CMS releases does not contain SSN; CMS releases data with the identifier BENE_ID which is also used as the link key. [1]	CMS policy and Data Use Agreement specify that data can only be linked to external data in accordance with the RIF Application approved by the CMS Privacy Board which specifies the scope of linkage	[1] CMS Staff Meeting
4.2.3	Whether data can be shared	Does not authorize/specify	Does not authorize/specify	
4.2.4	How data can be shared (de-identification status, disclosure review)	The Chronic Conditions Warehouse (CCW) Virtual Research Data Center (VRDC) output review process exists to help researchers protect Medicare and Medicaid beneficiaries' confidentiality. The purpose of the output review process is to help CCW VRDC users avoid accidental disclosure or the perceived disclosure of confidential information. The CCW analytical team reviews all output requested for download from the VRDC, and ensures it meets all disclosure checks before allowing the user to download it. [1] '-ensure output containing fields representing small cell size (N > 11) to ensure beneficiary or patient information privacy. '-ensure no personal identifiers '-CMS policies prohibit individual values such as extreme observations (e.g., five smallest or five greatest values in a distribution of data). By definition, an extreme observation is a sample of size N = 1. [1] Researchers must abide by the CMS suppression policy and HIPAA. All results in the VRDC go through output review before they can be downloaded. Researchers must aggregate and de-identify results in the enclave prior to downloading the results from the VRDC. [2]	CMS policy specifies that data shared through the VRDC: must undergo the VRDC Review Process which is disclosure review by CCW VRDC staff of analysis outputs prior to removing output data from the VRDC	[1] https://www2.ccwdata.org/documents/10280/19002246/ccw-vrdc-data-output-review-info.pdf (Accessed: 4/25/23) [2] CMS Staff Meeting

Dataset 7 - T-MSIS Analytic Files (TAF)				
		Raw Language	**Interpretation**	**Source**
4.2.5	How data can be accessed (access type, data use agreement, data access committee/group approval, IRB LOD, etc.)	CMS must ensure that all research requests for protected health information meet the requirements under the Common Rule and the Health Insurance Portability and Accountability Act (HIPAA) Privacy Rule. As a result, researchers must submit the following as part of the research request packet: Common Rule If subject to the Common Rule: » Documentation of institutional review board (IRB) approval of the research AND » Informed consent of the research subjects or IRB waiver of the requirement to obtain informed consent. If exempt from the Common Rule: » A signed and dated statement describing the basis for exemption. HIPAA Privacy Rule Individual authorization of the use of the data for the research OR Documentation that an IRB or a Privacy Board has approved a waiver of research subjects' authorization for use/disclosure of information about them for research purposes. The IRB study title must match the CMS data request study title, otherwise, a brief explanation is required. [1] You will need to submit a research request packet that includes a description of the research being conducted along with other materials. The Request Forms Generator will help you determine the documents you need. To begin the request process, email a draft (no signatures) of the documents in their original format (.xls, .pdf, .doc) to resdac@umn.edu 1. Submit draft request packet to ResDAC 2. ResDAC team reviews packet (ResDAC forwards data management plan for review. ResDAC obtains cost invoice) 3. Complete edits and return final signed documents to ResDAC 4. ResDAC submits final packet to CMS Privacy Board 5. CMS emails Privacy Board decision 6. Submit payment for data and notify CMS 7. Data is processed [2] The Data Management Plan Self-Attestation Questionnaire (DMP SAQ) for Federal Agencies allows for documentation that security and privacy controls have been implemented by the federal agency to protect the requested Identifiable Data File (IDFs) in the environment in which the data will be stored.	CMS specifies that for data access, a user: 1. Must submit RIF Data Use Agreement, Attachment A: RIF Application, RIF Application Key Personnel Supplement, RIF Specifications Worksheet, and Data Management Plan Self-Attestation Questionnaire (DMP SAQ) 2. Must obtain review and approval from ResDAC team on the proposed research 3. Must obtain review and approval from CMS' Data Privacy Safeguard Program (DPSP) on the Data Management Plan Self-Attestation Questionnaire (DMP SAQ) 4. Must obtain IRB LOD from the requesting institution 5. Must obtain review and approval from CMS Privacy Board on the proposed research 6. Must access data through the VRDC or through encrypted shipped disks	[1] https://resdac.org/irb-common-rule-and-hipaa-waiver-approval (Accessed: 4/25/23) [2] https://resdac.org/cms-research-identifiable-request-process-timeline (Accessed: 4/25/23) [3] https://resdac.org/request-form/dmp-saq (Accessed: 4/28/23) [4] CMS Staff Meeting

Dataset 7 - T-MSIS Analytic Files (TAF)				
		Raw Language	**Interpretation**	**Source**
		The DMP SAQ for Federal Agencies replaces the former Data Management Plan (DMP) requirement for CMS IDF requests. Unlike the DMP, which was specific to a single study, the DMP SAQ is an organizational-level plan and all studies using the approved environment can be covered by a single DMP SAQ. The Data Management Plan Self-Attestation Questionnaire (DMP SAQ) documents security and privacy controls implemented by the research organization to protect the requested Research Identifiable Files (RIF) in the environment in which the data will be stored. The DMP SAQ replaces the former Data Management Plan (DMP) requirement for CMS RIF requests. Unlike the DMP, which was specific to a single study, the DMP SAQ is an organizational-level plan and all studies using the approved computing environment can be covered by a single DMP SAQ. The DMP SAQ is based on the CMS Acceptable Risk Safeguards (ARS) security and privacy controls. Research organizations attest that the organization complies with CMS ARS security and privacy controls addressed by the questionnaire. Some questions also require additional explanation and evidence. The DMP SAQ recognizes information systems may vary between organizations and allows flexibility through compensating controls or alternative implementations. The important takeaway when implementing the controls is that the intent of the security and privacy control is met. For any control that cannot be met, organizations must provide justification for not being able to implement the control. Approved DMP SAQs are valid for one year, after which organizations will need to recertify and update the DMP SAQ to capture any changes to their environments. Any changes to the organization's environment prior to the recertification date require notification within 15 days of the change. [3] After CMS receives the payment for data request, the request is sent to the contractor who then pulls the requested data extracts/research files as listed in the Specification Worksheet. The data is either encrypted and shipped or made available in the VRDC. If the data is accessed in the VRDC, VRDC staff reach out to the seat holders to begin the onboarding process. This includes identity verification and security training. [4]		

Dataset 7 - T-MSIS Analytic Files (TAF)				
		Raw Language	**Interpretation**	**Source**
4.2.6	How data can be used (data use limitations)	Does not authorize/specify	Does not authorize/specify	
4.2.7	Other (specify)	Does not authorize/specify	Does not authorize/specify	
5	**Data Use**			
5.1	**Authorizations and Applicable Regulations/Policies**			
5.1.1	**Authorizations**			
5.1.1.1	Assent	N/A	N/A	
5.1.1.2	Consent	N/A	N/A	
5.1.1.3	IRB/equivalent Privacy Board determination	Information not available/found	Information not available/found	
5.1.1.4	Local/state/federal laws	Information not available/found	Information not available/found	
5.1.1.5	Institutional Certification	Information not available/found	Information not available/found	
5.1.1.6	Data originator agreement	Information not available/found	Information not available/found	
5.1.1.7	Repository agreements/policies	Data access is authorized through CMS Privacy Board approval (the CMS Privacy Board consults with the data owner, who reviews all data requests) and under the DUA. [1]	Data Use Agreement (DUA) authorizes data use	[1] CMS Staff Meeting
5.1.1.8	Other (specify)	Information not available/found	Information not available/found	
5.1.2	**Applicable Regulations/Policies**			
5.1.2.1	Local regulations/policies	Federal agencies can request TMSIS for research purposes following the standard research process (including approval by the CMS Privacy Board). Federal agencies cannot request CMS data simply to allow secondary use (e.g., to build a database for their grantees), there must be a research use to justify the initial disclosure and findings from the research must be publicly available. [1] For standard TAFs, the CMS Privacy Board authorizes linkage and use of the TAFs by researchers for research purposes. Linkage cannot occur if the Privacy Board does not authorize the linkage as part of the RIF application approval. [1]	T-MSIS policy	[1] CMS Staff Meeting
5.1.2.2	Tribal regulations/policies	Information not available/found	Information not available/found	
5.1.2.3	State regulations/policies	Information not available/found	Information not available/found	
5.1.2.4	Federal regulations/policies	CMS is a covered entity under HIPAA and is subject to the Privacy Act when collecting, using, and disclosing data. (https://www.federalregister.gov/documents/2019/02/06/2019-01157/privacy-act-of-1974-system-of-records) [1]	1. Privacy Act 2. HIPAA Privacy Rule	[1] CMS Staff Meeting
5.1.2.5	International regulations/policies	Information not available/found	Information not available/found	
5.1.2.6	Contractual obligations	Information not available/found	Information not available/found	

Dataset 7 - T-MSIS Analytic Files (TAF)				
		Raw Language	**Interpretation**	**Source**
5.1.2.7	Repository policies	The Research Identifiable File (RIF) Data Use Agreement (DUA) is a legal agreement between CMS and a requesting organization that documents the terms and conditions under which the CMS data may be used, including CMS privacy and security requirements and data release policies. [1] The Research Identifiable File (RIF) application collects information about the requesting organization and the research study, including detailed study aims, data required, and dissemination of findings plan. The RIF application is used by CMS to assess the feasibility of the research and compliance with CMS data use and release policies. [2] The Key Personnel Supplement collects information about each request's key contacts including the requester, collaborating organizations and additional contacts who should be included on notices about the project. [3] The Specifications Worksheet is required for all RIF requests. It collects detailed requester information, study/project data extract details, shipping information, and method of payment. It also includes a Part D event justification tab that is required for all requests that include Part D data. The Specifications Worksheet is used by the data distributor to generate a cost invoice and to collect data extraction information.[4] CMS Information Exchange Agreement (IEA): "This is used to modify the terms of the research DUA to allow redisclosure of linked datasets by federal agencies. However, we are exploring creating a research DUA addendum with standard clauses for the redisclosure scenario instead. This document is also used by the CMS Privacy Office for other types of disclosures (other than research) or exchanges with federal agencies." [5]	CMS Research Data Center (ResDAC) policy	[1] https://resdac.org/request-form/rif-data-use-agreement (also see https://resdac.org/sites/datadocumentation.resdac.org/files/2022-10/RIF%20Data%20Use%20Agreement.pdf) (Accessed: 4/25/23)) [2] https://resdac.org/request-form/rif-application (Accessed: 4/25/23) [3] https://resdac.org/request-form/key-personnel-supplement (Accessed: 4/25/23) [4] https://resdac.org/request-form/rif-specifications-worksheet (Accessed: 4/25/23) [5] CMS Email Communication [6] IEA Template
5.2	**Governance for data linkage, sharing, access, and use based on data access authorization or applicable regulations/policies (i.e., the origin of the governance)**			
5.2.1	Whether the data can be linked	Does not authorize/specify	Does not authorize/specify	
5.2.2	With what other data can it be linked or can it not be linked (scope of linkage)	Does not authorize/specify	Does not authorize/specify	
5.2.3	Whether data can be shared	Does not authorize/specify	Does not authorize/specify	
5.2.4	How data can be shared (de-identification status, disclosure review)	Does not authorize/specify	Does not authorize/specify	
5.2.5	How data can be accessed (access type, data use agreement, data access committee/group approval, IRB LOD, etc.)	Does not authorize/specify	Does not authorize/specify	

Dataset 7 - T-MSIS Analytic Files (TAF)				
		Raw Language	**Interpretation**	**Source**
5.2.6	How data can be used (data use limitations)	Federal agencies can request TMSIS for research purposes following the standard research process (including approval by the CMS Privacy Board). Federal agencies cannot request CMS data simply to allow secondary use (e.g., to build a database for their grantees), there must be a research use to justify the initial disclosure and findings from the research must be publicly available. [1] For standard TAFs, the CMS Privacy Board authorizes linkage and use of the TAFs by researchers for research purposes. Linkage cannot occur if the Privacy Board does not authorize the linkage as part of the RIF application approval. [1] DATA USE AGREEMENT FOR USE OF CENTERS FOR MEDICARE AND MEDICAID SERVICES (CMS) RESEARCH IDENTIFIABLE FILES (RIFS): a) The data files specified in Attachment A – RIF Request Application, as well as any derivative data, will be used solely for the study titled:----------, as described in detail in "Attachment A –RIF Request Application", as modified, if applicable, by submitting a request and receiving CMS approval to amend Attachment A – RIF Request Application. c) Attachment A - RIF Request Application contains a detailed description of the entirety of the research to be done in the above-referenced research study, the research could not practicably be conducted without CMS data, and the requested data is the minimum necessary to achieve the stated research purpose(s). d) As described in Attachment A – RIF Request Application that is submitted to CMS, the researcher believes that the study demonstrates the potential to improve the administration of the Medicare and Medicaid programs [2]	1. T-MSIS policy specifies that there must be a research use to justify the initial disclosure and findings from the research must be publicly available. 2. Data Use Agreement (DUA) specifies that data will be used solely for the study described in detail in the RIF Request Application	[1] CMS Staff Meeting [2] https://resdac.org/sites/datadocumentation.resdac.org/files/2022-10/RIF%20Data%20Use%20Agreement.pdf
5.2.7	Other (specify)	Does not authorize/specify	Does not authorize/specify	
6	**PII Elements**			

Dataset 7 - T-MSIS Analytic Files (TAF)				
		Raw Language	**Interpretation**	**Source**
6.1	PII elements collected	The eligible individual's social security number. For newborns when value is unknown it is not required. For SSN states, in instances where the social security number is not known and a temporary MSIS Identification Number is used, the MSIS Identification Number field should be populated with the temporary MSIS Identification Number and the SSN field should be space-filled, or blank. When the SSN becomes known, the MSIS Identification Number field should continue to be populated with the temporary MSIS Identification Number and the SSN field should be populated with the newly acquired SSN for at least one monthly submission of the Eligible File so that T-MSIS can associated the temporary MSIS Identification Number and the social security number. The last name of the individual to whom the services were provided. (The patients name should be captured as it appears on the claim record, it does not need to be the same as it appears on the eligibility file. The MSIS Identification Number will be used to associate a claim record with the appropriate eligibility data.) The first name of the individual to whom the services were provided. (The patients name should be captured as it appears on the claim record, it does not need to be the same as it appears on the eligibility file. The MSIS Identification Number will be used to associate a claim record with the appropriate eligibility data.)	SSN, First Name, Middle Initial, Last Name, Address, City, State, Zip code, date of birth, sex	[1] https://www.medicaid.gov/tmsis/dataguide/data-elements/ (Accessed: 4/25/23)

Dataset 7 - T-MSIS Analytic Files (TAF)				
		Raw Language	**Interpretation**	**Source**
		Individual's middle initial; middle initial component of full name (e.g. First Name, Middle Initial, Last Name). The first line of a potentially multi-line physical street or mailing address for a given entity (e.g. person, organization, agency, etc.). The city component of an address associated with a given entity (e.g. person, organization, agency, etc.). The ANSI state numeric code for the U.S. state, Territory, or the District of Columbia code for where the individual eligible to receive healthcare services resides. (The state for the type of address indicated in Address Type.) U.S. ZIP Code component of an address associated with a given entity (e.g. person, organization, agency, etc.) Date of birth of the individual to whom the services were provided. A patient's age should not be greater than 112 years. Either individual's biological sex or their self-identified sex. [1]		
6.2	PII elements holder (i.e., party that holds the PII)	(See data dictionary definitions in 6.1)	PII elements are part of the research identifiable dataset from TMSIS	[1] https://www.medicaid.gov/tmsis/dataguide/data-elements/ (Accessed: 4/25/23)
6.3	Use of common data model, if any, for data collection	Information not available/found	Information not available/found	
7	**Prior Data Linkages**			
7.1	**Dataset linked with other datasets**			
7.1.1	Name of other linked dataset	NCHS survey data from participants who consented to linkage Approval for the linkage was provided by the NCHS Research Ethics Review Board (ERB). [1]	Linkage was performed only on consented data with approval from NCHS ERB	[1] https://www.cdc.gov/nchs/data/datalinkage/nchs-cms-tmsis-linkage-methodology.pdf (Accessed: 4/25/23)
7.1.2	Other dataset type (clinical, EHR, survey, claims, SDOH, etc.)	NCHS has recently linked data from following surveys to 2014-2019 CMS T-MSIS enrollment and claims data: • 1994-2018 National Health Interview Survey (NHIS) • 1999-2018 Continuous National Health and Nutrition Examination Survey (NHANES) • Third National Health and Nutrition Examination Survey (NHANES III) • 2004 National Nursing Home Survey (NNHS) [1]	Survey data from NCHS (NHIS, NHANES, and NNHS)	[1] https://www.cdc.gov/nchs/data/datalinkage/nchs-cms-tmsis-linkage-methodology.pdf (Accessed: 4/25/23)
7.1.3	Other dataset source(s)	NCHS	Source was NCHS datasets	[1] https://www.cdc.gov/nchs/data/datalinkage/nchs-cms-tmsis-linkage-methodology.pdf (Accessed: 4/25/23)

Dataset 7 - T-MSIS Analytic Files (TAF)				
		Raw Language	**Interpretation**	**Source**
7.1.4	Linking methodology (PPRL or non-PPRL); linkage technology	The NCHS survey participant records and the CMS T-MSIS enrollment database were linked using both deterministic and probabilistic approaches. For the probabilistic approach, scoring was conducted according to the Fellegi-Sunter method. Following this, a selection process was implemented with the goal of selecting pairs believed to match (i.e., representing the same individual between the data sources). 1. Deterministic linkage joined records on exact SSN, with links validated by comparing other identifying fields (i.e., first name, last name, day of birth, etc.) 2. Probabilistic linkage identified likely matches, or links, between all records. All records were probabilistically linked and scored as follows: a. Formed pairs via blocking b. Scored pairs c. Modeled probability – assigned estimated probability that pairs are matches 3. Pairs were selected that were believed to represent the same individual between data sources (i.e., they are a match). Deterministic matches (from step 1) were assigned a match probability of 1 and records selected from the probabilistic match (step 2) were assigned the modeled match probability. For each NCHS survey participant record that was linked, CMS extracted the T-MSIS claims information and sent the data to NCHS following secure data transfer procedures. [1]	Non-PPRL method was used - both deterministic using direct identifiers and probabilistic using various PII	[1] https://www.cdc.gov/nchs/data/datalinkage/nchs-cms-tmsis-linkage-methodology.pdf (Accessed: 4/25/23)
7.1.5	PII elements used for the linkage	Linkage-eligible NCHS survey participant records were linked to the CMS T-MSIS enrollment database using the following identifiers: SSN (9 digits or last 4 digits, depending on the survey and year of the survey), first name, last name, middle initial, month of birth, day of birth, year of birth, 5-digit ZIP code of residence, state of residence, and sex. [1]	Broad set of PIIs were used: SSN (9 digits or last 4 digits, depending on the survey and year of the survey), first name, last name, middle initial, month of birth, day of birth, year of birth, 5-digit ZIP code of residence, state of residence, and sex.	[1] https://www.cdc.gov/nchs/data/datalinkage/nchs-cms-tmsis-linkage-methodology.pdf (Accessed: 4/25/23)
7.1.6	Entity resolver (data originator or data linker or third party)	NCHS	NCHS	[1] https://www.cdc.gov/nchs/data/datalinkage/nchs-cms-tmsis-linkage-methodology.pdf (Accessed: 4/25/23)
7.1.7	Party performing the linkages	NCHS	NCHS	[1] https://www.cdc.gov/nchs/data/datalinkage/nchs-cms-tmsis-linkage-methodology.pdf (Accessed: 4/25/23)

Dataset 7 - T-MSIS Analytic Files (TAF)				
		Raw Language	**Interpretation**	**Source**
7.1.8	Linkage quality assessment	Subsequent to performing the record linkage analysis an error analysis was performed. There are two type of errors that were estimated: • Type I Error: Among pairs that are linked, what percentage of them were not true matches • Type II Error: Among true matches, how many were not linked [1]	Linkage error analysis was performed - details in the pdf file	[1] https://www.cdc.gov/nchs/data/datalinkage/nchs-cms-tmsis-linkage-methodology.pdf (Accessed: 4/25/23)
7.1.9	Linked data sharing method (linkage maps or pre-linked dataset)	To ensure confidentiality, NCHS provides safeguards including the removal of all personal identifiers from analytic linked files. Additionally, the linked data files are only made available in secure facilities for approved research projects. Researchers who wish to access the Linked NCHS-CMS T-MSIS data files must submit a research proposal to the NCHS Research Data Center (RDC) to obtain permission to access the restricted use files. All researchers must submit a research proposal to determine if their projects are feasible and to gain access to these restricted data files. The proposal provides a framework which allows RDC staff to identify potential disclosure risks. More information regarding the RDC and instructions for submitting an RDC proposal are available from: https://www.cdc.gov/rdc/ (accessed September 19, 2022). [1]	The linked data files are made available only through the RDC for approved research projects.	[1] https://www.cdc.gov/nchs/data/datalinkage/nchs-cms-tmsis-linkage-methodology.pdf (Accessed: 4/25/23)

TABLE 1: USE CASE 3 - DATASET GOVERNANCE

[Legend: Blank cell in Table 1 = information not available/found; N/A = information confirmed to not exist]

	Governance	Dataset 1 - N3C					Dataset 2 - PEDSnet					Dataset 3 - RADx-UP					Dataset 4 - EPA				
		Data Collection	Data Linkage	Data Sharing	Data Access	Data Use	Data Collection	Data Linkage	Data Sharing	Data Access	Data Use	Data Collection	Data Linkage	Data Sharing	Data Access	Data Use	Data Collection	Data Linkage	Data Sharing	Data Access	Data Use
1	**Authorization/s**																				
1.1	Assent	N/A	N/A	N/A	N/A	N/A	Assent authorizes data collection.					Assent from children authorizes data collection.	Does not authorize/ specify	Assent from children authorizes data sharing.	Does not authorize/specify	Does not authorize/specify	N/A	N/A	N/A	N/A	N/A
1.2	Consent	N/A	N/A	N/A	N/A	N/A	Consent authorizes data collection.	Consent (when obtained) authorizes data linkage.				Consent from parents authorizes data collection.	Does not authorize/ specify	Consent from parents authorizes data sharing.	Does not authorize/specify	Does not authorize/specify	N/A	N/A	N/A	N/A	N/A
1.3	IRB/equivalent Privacy Board determination	N/A		Two IRB determinations authorize data sharing: 1. Institutional IRB or external central IRB approval for transfer of data from participating institutions to the NCATS N3C Platform 2. NIH IRB waiver of consent for sharing through the NCATS N3C Platform	Letter of Determination from user's Insitutional IRB authorizes data access	Letter of Determination from user's institutional IRB authorizes data access	CHOP IRB or local/site-level IRB approval authorizes data collection.	IRB authorizes data linkage for research conducted under a waiver of consent.		Human Subjects Review (IRB review or determination) authorizes data access in two possible paths: 1. Non-Human Subjects Research (NHSR) determination: no further review/MRA required 2. Human Subjects Research (HSR) determination: requester IRB approval with IRB reliance for site providing data (NPRA MRA or SMART IRB MRA is also required		AHARO Health Centers/Comprehensive Health Center IRB authorizes data collection.	AHARO Health Centers/Comprehensive Health Center IRB authorizes data linkage.			Does not authorize/specify	N/A	N/A	N/A	N/A	N/A
1.4	Local/state/federal laws	HIPAA authorizes health care providers to release data to N3C	N/A					PEDSnet Steering Committee approval authorizes data linkage as part of a specific study research plan.		Network Participation Approval (PEDSnet Executive Committee Approval) authorizes data access.	Steering Committee authorizes data use.	Hawaii DOE Data Governance and Analytics Branch authorizes data collection.		FERPA authorizes data sharing.			Clean Air Act authorizes data collection by state/local/tribal air pollution control agencies for reporting to the EPA.		Clean Air Act authorizes data sharing.		
1.5	Institutional Certification	N/A	N/A		N/A									RADx Institutional Certification authorizes data sharing.		The RADx Institutional Certification authorizes data use.	N/A	N/A	N/A	N/A	N/A
1.6	Data originator agreement	N/A	Two data originator agreements authorize data linkage: 1. The Linkage Honest Broker Agreement (LHBA) authorizes data linkage 2. Participating PPRL sites authorize data linkage for particular external datasets	Data Transfer Agreement (DTA) executed with NCATS authorizes data transfer to NCATS N3C Platform			Master Reliance Agreement authorizes data collection under one IRB.	Individual PEDSnet sites authorize data linkage on a study-by-study basis.	Two agreements authorize data sharing: 1. PEDSnet Participation and DUA signed by PEDSnet members authorizes data sharing between PEDSnet members. 2. PEDSnet sites data release vote authorizes data sharing (i.e., release).	Institutional participation approval (prospective site PI approval) authorizes data access.	Two agreements authorize data use: 1. PEDSnet Site PI written affirmation authorizes data use for other investigators 2. PEDSnet site participation vote authorizes data use for individual studies			Agreement For Disclosure And Transfer Of Confidential Information And Protected Health Information authorizes data sharing.						N/A
1.7	Repository agreements/policies	N/A	The External Dataset Committee in the Tools and Resource subgroup and NCATS approval authorizes data linkage		Two repository agreements authorize data access: 1. Data Use Agreement 2. Data Use Request	Two repository agreements authorize data use: 1. Data Use Agreement 2. Data Use Request			PEDSnet Master Data Use Agreement authorizes data sharing.	Legal review (comprised of PEDSnet Data use agreement and Responsible Use of Data Agreement) authorizes data access.	Legal review (comprised of PEDSnet Data use agreement and Responsible Use of Data Agreement) authorizes data use.				Three repository agreements authorize data access: 1. RADx Data Use Certification (DUC) Agreement 2. Genomic Data Use Coder of Conduct 3. RADxSM Data User Code of Conduct	RADx Data Use Certification Agreement authorizes data use.	N/A	AQS being in the public domain authorizes data linkage.		AQS being in the public domain authorizes data access.	AQS being in the public domain authorizes data use.
1.8	Other (specify)		The External Dataset Committee in the Tools and Resource subgroup and NCATS approval authorizes data linkage				Waiver of consent authorizes data collection for routine delivery of health care.			Network Participation Approval (PEDSnet Executive Committee Approval) authorizes data access.	Steering Committee authorizes data use.										
2	**Applicable Regulations/Policies**																				
2.1	Local regulations/ policies		N3C policy designation of external datasets	N3C policies	Information not available/found	N3C Data User Code of Conduct			PEDSnet policy		PEDSnet policy										

TABLE 1: USE CASE 3 - DATASET GOVERNANCE

[Legend: Blank cell in Table 1 = information not available/found; N/A = information confirmed to not exist]

	Governance	Dataset 1 - N3C					Dataset 2 - PEDSnet					Dataset 3 - RADx-UP					Dataset 4 - EPA				
		Data Collection	Data Linkage	Data Sharing	Data Access	Data Use	Data Collection	Data Linkage	Data Sharing	Data Access	Data Use	Data Collection	Data Linkage	Data Sharing	Data Access	Data Use	Data Collection	Data Linkage	Data Sharing	Data Access	Data Use
2.2	Tribal regulations/policies			1. NIH Guidance on the Implementation of the HHS Tribal Consultation Policy 2. Tribal Consultation Report		1. NIH Guidance on the Implementation of the HHS Tribal Consultation Policy 2. Tribal Consultation Report															
2.3	State regulations/policies								State regulations												
2.4	Federal regulations/policies	HIPAA Privacy Rule		HHS Tribal Consultation Policy	Certificate of Confidentiality	HHS Tribal Consultation Policy			1. Federal regulations 2. 45 CFR 46 (Common Rule) 3. HIPAA Privacy Rule												
2.5	International regulations/policies																				
2.6	Contractual Obligations				Obligations from contract between NCATS and Palantir																
2.7	Repository policies		N3C policy designation of external datasets		N3C policies, including N3C Results Download Policy	N3C policies				PEDSnet policy	PEDSnet policy			RADx Data Hub policy	RADx Data Hub policy	RADx Data Hub policy	N/A	AQS policy		AQS policy	AQS policy
3	Data Linking/Sharing/Access/Use Governance Based on Authorizations and Applicable Regulations/Policies																				
3.1	Whether the data can be linked	Does not authorize/specify	1. LHBA specifies that the data can be linked 2. Participating PPRL sites specify that data can be linked with particular external datasets 2. The External Dataset Committee in the Tools and Resource subgroup and NCATS approval specifies that data can be linked	Does not authorize/specify	Does not authorize/specify	Does not authorize/specify	Consent (when obtained) specifies that data can be linked.	1. Consent (when obtained) specifies that data can be linked. 2. PEDSnet Steering Committee approval specifies that data can be linked according to the approved research plan. 3. Individual PEDSnet sites, through a study particpation vote, specify that the sites can participation in data linkage on a study-by-study basis. 4. IRB specifies that PEDSnet data can be linked for research conducted under a waiver of consent.	Does not authorize/specify	Does not authorize/specify	Individual PEDSnet sites, through a study particpation vote, specify participation in data linkage on a study-by-study basis	1. Parental informed consent and assent do not specify linkage. Raw language referring to "other research studies" is interpreted by the study PI as leaving the option open for data linkage. 2. AHARO Health Centers/Comprehensive Health Center IRB specifies that data can be linked at an individual level only if the IRB approves the linkage.	1. Parental informed consent and assent do not specify linkage. Raw language refering to "other research studies" is interpreted by the study PI as leaving the option open for data linkage. 2. AHARO Health Centers/Comprehensive Health Center IRB specifices that data can be linked at an individual level only if the IRB approves the linkage.	Parental informed consent and assent do not specify linkage. Raw language refering to "other research studies" is interpreted by the study PI as leaving the option open for data linkage.	Does not authorize/specify	Does not authorize/specify	Does not authorize/specify	N/A	Does not authorize/specify		Does not authorize/specify
3.2	With what other data can it be linked or can it not be linked (scope of linkage)	Does not authorize/specify	1. N3C policy designation of external datasets for linking specifies that: a. External datasets must be classified as Class 0, 2, 3, or 4 to be considered for N3C linkage. A dataset which is categorized as class 2 can be imported but will require hashing b. Class 1 linkages are not permitted 2. Participating PPRL sites specify linkages with external datasets on a case-by-case basis 3. External Dataset Committee in the Tools and Resource subgroup and NCATS determines the scope of linkage by approving external datasets for import and linkage within N3C	Common Rule	Does not authorize/specify	Does not authorize/specify	Governing IRB protocols specify that data can be linked using PPRL for research conducted under a waiver of consent.	1. IRB specifies that data can be linked using PPRL for research conducted under a waiver of consent. 2. Individual PEDSnets study sites specify the scope of data linkage on a study-by-study basis.	Does not authorize/specify	Does not authorize/specify	Does not authorize/specify	AHARO Health Centers/ Comprehensive Health Center IRB specifies that any data linkages at an individual level or outside of general research purposes must be approved by the AHARO Health Centers IRB.	AHARO Health Centers/ Comprehensive Health Center IRB specifies that any data linkages at an individual level or outside of general research purposes must be approved by the AHARO Health Centers IRB.	Does not authorize/specify	Does not authorize/specify	Does not authorize/specify	Does not authorize/specify	N/A	Does not authorize/specify	Does not authorize/specify	Does not authorize/specify

TABLE 1: USE CASE 3 - DATASET GOVERNANCE

[Legend: Blank cell in Table 1 = information not available/found; N/A = information confirmed to not exist]

	Governance	Dataset 1 - N3C					Dataset 2 - PEDSnet					Dataset 3 - RADx-UP					Dataset 4 - EPA				
		Data Collection	Data Linkage	Data Sharing	Data Access	Data Use	Data Collection	Data Linkage	Data Sharing	Data Access	Data Use	Data Collection	Data Linkage	Data Sharing	Data Access	Data Use	Data Collection	Data Linkage	Data Sharing	Data Access	Data Use
3.3	Whether data can be shared	Does not authorize/specify	Does not authorize/specify	1. NIH IRB waiver of consent specifies that data can be shared 2. Data Transfer Agreement specifies that data can be shared 3. NIH Guidance on the Implementation of the HHS Tribal Consultation Policy, the HHS Tribal Consultation Policy, and the Tribal Consultation Report specify that AI/AN data can be shared	Does not authorize/specify	Does not authorize/specify	Does not authorize/specify	Does not authorize/specify	Does not authorize/specify	1. PEDSnet sites data release vote specifies whether data can be shared (i.e., release). 2. The PEDSnet Master DUA specifies whether data can be shared. 3. 45 CFR 46 (Common Rule) and HIPAA Privacy Rule specifies that de-identified data can be shared.	Does not authorize/specify	Does not authorize/specify	Parental informed consent specifies that data will be shared with other researchers.	Parental informed consent specifies that data will be shared with other researchers.	The following specify that data can be shared: 1. Parental informed consent language 2. FERPA 3. Agreement For Disclosure And Transfer Of Confidential Information And Protected Health Information (study data, including PHI, to be sent to Duke (CDCC) and for Duke to provide de-identified project data for the awarding agency.) 4. Study registration in dbGaP 5. RADx Institutional Certification (in the RADx Data Hub by the CDCC/Duke)	Does not authorize/specify	RADx Institutional Certification specifies that the CDCC (Duke in the case of RADx UP) submits data to the RADx Data Hub.	Does not authorize/specify	N/A	Clean Air Act specifies that ambient air data can be shared.	Does not authorize/specify	Does not authorize/specify
3.4	How data can be shared (de-identification status, disclosure review)	HIPAA specifies that N3C data partners (health care providers) can release limited EHR dataset (with no direct PII) for research purposes	Does not authorize/specify	1. N3C policy specifies limited datasets (LDS), de-identified, and synthetic datasets are shared. 2. NIH Guidance on the Implementation of the HHS Tribal Consultation Policy, the HHS Tribal Consultation Policy, and the Tribal Consultation Report specify that (a). AI/AN data will be a standalone category. With this change, AI/AN data will be available in any N3C analysis that provides race and ethnicity distribution.(b). ZIP codes must be removed entirely for all geographic units containing 20,000 or fewer people, and full five-digit ZIP codes of predominantly AI/AN community will never be shown.	Does not authorize/specify	Does not authorize/specify	Does not authorize/specify	Does not authorize/specify	1. State and federal regulations specify that HIV-related data and reproductive and mental health health care data for minors must be removed before sharing data. 2. PEDSnet regulations specify that individual level data must be de-identified using the Safe Harbor method of de-identification of PHI before sharing data. 3. PEDSnet policy specifies that a risk review is performed on the requested datasets as well as data transformations, such as date shifts, replacement labels for free text fields and geographic information, and removing HIV/pregnancy/ mental health data prior to data sharing.	Does not authorize/specify	Does not authorize/specify	Does not authorize/specify	Does not authorize/specify	1. RADx policy specifies that the study be registered in dbGaP prior to sharing through RADx Data Hub. 2. The RADx Institutional Certification specifies that all data shared in an NIH designated repository must be de-identified. 3. The RADx DCC works with study teams to de-identify zip codes, shift dates, and adjust ages into categories for specific ages.	Does not authorize/specify	1. RADx Institutional Certification specifies that all data shared in the RADx Data Hub must be-identified. 2. RADx Institutional Certification specifies that all data be shared to an NIH-designated repository [RADx Data Hub].	Clean Air Act specifies that ambient air data be shared through EPA's Air Quality System (AQS).	N/A	Clean Air Act specifies that full geographic identifiers including site address, zip code, CBSA, county, and state are shared.	AQS policy specifies that full geographic identifiers including site address, zip code, CBSA, county, and state are shared.	Does not authorize/specify

TABLE 1: USE CASE 3 - DATASET GOVERNANCE

[Legend: Blank cell in Table 1 = information not available/found; N/A = information confirmed to not exist]

	Governance	Dataset 1 - N3C					Dataset 2 - PEDSnet					Dataset 3 - RADx-UP					Dataset 4 - EPA				
		Data Collection	Data Linkage	Data Sharing	Data Access	Data Use	Data Collection	Data Linkage	Data Sharing	Data Access	Data Use	Data Collection	Data Linkage	Data Sharing	Data Access	Data Use	Data Collection	Data Linkage	Data Sharing	Data Access	Data Use
3.5	How data can be accessed (access type, data use agreement, data access committee/ group approval, IRB LOD, etc.)	HIPAA Privacy Rule specifies that N3C data partners (health care providers) must enter into a data use agreement with limited EHR dataset receipients	N3C specifies that for data access, there must be: 1. For N3C Class 0 or Class 2 linkages: a. Existing institutional N3C Data Use Agreement b. Dual authentication and authorization c. Signed institutional linkage honest broker agreement for multiple datasets d. Approved data use request (DUR) by the federally staffed data access committee (DAC) e. Local institutions IRB letter of determination f. Interconnect agreement (for Class 0 only) 2. For N3C Class 3 and 4 linkages: a. Approved data use request (DUR) by the federally staffed data access committee (DAC)	The DTA specifies that users who access the data will access the data within the NCATS N3C Platform	N3C policies specify that to access the Limited Dataset, the user: 1) Must complete N3C registration and create a N3C Data Enclave account 2) Must execute an Institutional Data Use Agreement 3) Must submit Data Use Request (DUR) for approval by N3C Data Access Committee 4) Must complete NIH IT training, attest to the N3C Data User Code of Conduct, and complete Human Subjects Research Protection training at their home institution 5) Must provide IRB letter of determination for data access 6) Must access the data within the N3C Enclave 7) Must abide by the N3C Results Download Policy when downloading results from the N3C enclave	N3C policy specifies that to access the Limited Dataset, the user: 1) Must complete N3C registration and create a N3C Data Enclave account 2) Must execute an Institutional Data Use Agreement 3) Must submit Data Use Request (DUR) for approval by N3C Data Access Committee 4) Must complete NIH IT training , attest to the N3C Data User Code of Conduct, and complete Human Subjects Research Protection training 5) Must provide IRB letter of determination for data access 6) Must access the data within the N3C Enclave	Does not authorize/specify	Does not authorize/specify	Does not authorize/specify	PEDSnet human subjects review, network participation review, institutional participation approval, and legal review specify that the requester: 1. Must submit request form for approval by the Research Committee 2. Must undergo IRB review/determination (Human Subjects Review) 3. If IRB determines the proposed study is NHSR, then no further review/MRA required 4. If IRB determines the proposed study HSR, the requester must provide IRB approval with IRB reliance for site providing data (NPRA MRA or SMART IRB MRA is also required) 5. Must sign DUA and RUD (Responsible Use of Data) (Legal Review) 6. Must receive prospective site PI approval (Institutional Participation Approval) 7. Must receive PEDSnet Executive Committee approval (Network Participation Approval) 8. Must access the data through a workspace within the PEDSnet cloud enclave--OR--to have the data transfered to their institution, the PEDSnet Study Approval request should specify, pending approval from all PEDSnet institutions providing data for the request.	PEDSnet policy specifies that the requester: 1. Must sign DUA and RUD (Responsible Use of Data) (Legal Review) 2. Must access the data through a workspace within the PEDSnet cloud enclave--OR--to have the data transfered to their institution, the PEDSnet Study Approval request should specify, pending approval from all PEDSnet institutions providing data for the request	Does not authorize/specify	Does not authorize/specify	RADx Institutional Certification specifies that all individual-level data are controlled access.	The dbGaP data access request process specifies RADx Data Hub data access. The requirements/steps specify that the user/ eligible investigator: 1) Must have an eRA commons or Login.gov account 2) Must submit a Data Access Request (DAR). Data Use Certification (DUC) Agreement, the Genomic Data User Code of Conduct, and the RADx SM Data User Code of Conduct are signed as part of the DAR process 3) Must ensure that the Signing Official from investigator's institution reviews, approves, and co-signs the request 4) Must receive approval from Data Access Committee 5) Must access the controlled access data through RADx Data Hub Jupyter Notebooks	RADx Institutional Certification specifies that all individual-level data are controlled access.	Does not authorize/specify	N/A	Does not authorize/specify	AQS being in the public domain specifies that the data is open access.	Does not authorize/specify

TABLE 1: USE CASE 3 - DATASET GOVERNANCE

[Legend: Blank cell in Table 1 = information not available/found; N/A = information confirmed to not exist]

	Governance	Dataset 1 - N3C					Dataset 2 - PEDSnet					Dataset 3 - RADx-UP					Dataset 4 - EPA				
		Data Collection	Data Linkage	Data Sharing	Data Access	Data Use	Data Collection	Data Linkage	Data Sharing	Data Access	Data Use	Data Collection	Data Linkage	Data Sharing	Data Access	Data Use	Data Collection	Data Linkage	Data Sharing	Data Access	Data Use
3.6	How data can be used (including data use limitations)	HIPAA privacy regulations specify that de-identified data can be used for general research purposes.	Does not authorize/specify	1. The Tribal Consultation Report specifies that data users will be asked to attest that they understand the N3C contains no Tribal affiliation data and that use of AI/AN data and ZIP code information to make assumptions about Tribal affiliation is not valid or appropriate 2. The DTA specifies that data must be used only for research purposes and public health activities related to the COVID-19 pandemic	N3C DUA specifies that the data must be used exclusively for the Research Project proposed, and/or comparative studies using data from individuals infected with pathogens such as SARS, MERS, and H1N1 to support comparative studies.	1. N3C DUA specifies that the data must be used exclusively for the Research Project proposed, and/or comparative studies using data from individuals infected with pathogens such as SARS, MERS, and H1N1 to support comparative studies. 2. N3C Data User Code of Conduct specifies that N3C data must only be used for COVID-19 general research purposes. 3. NIH Guidance on the Implementation of the HHS Tribal Consultation Policy, the HHS Tribal Consultation Policy, and the Tribal Consultation Report specify that data users will be asked to attest that they understand the N3C contains no Tribal affiliation data and that use of AI/AN data and ZIP code information to make assumptions about Tribal affiliation is not valid or appropriate. 4. User must also comply with the N3C Community Guiding Principles and the Attribution and Publication Principles.	Does not authorize/specify	PEDSnet site participation vote specifies data use for individual studies.	Does not authorize/specify	Does not authorize/specify	PEDSnet policy specifies two data use limitations: 1. Data can only be used for the purposes specified and approved by the Steering Committee. Namely, using data from real-world clinical settings for research, quality measurement, and improvement/ advancement of child health, particularly studies that inform or directly address clinical decision making, including retrospective observational studies 2. Cannot be used for commercial sale	Does not authorize/specify	Does not authorize/specify	RADx Institutional Certification specifies that the data can be used for general research purposes.	Does not authorize/specify	RADx Institutional Certification specifies that the use of data is for general research purposes.	Does not authorize/specify	N/A	Does not authorize/specify	Does not authorize/specify	No data use limitations since the data is open access
3.7	Other (specify)	Does not authorize/specify	Does not authorize/specify	Does not authorize/specify	1. Contract between NCATS and Palantir specifies that Palantir contractors with access to the NCATS GovCloud instance to implement and maintain the NCATS N3C Data Enclave are subject to all relevant NIH-specified clearances, non-disclosure agreements, training, rules and restrictions and are not allowed to independently access NCATS N3C Data Enclave data, remove it from the enclave, or use it for commercial purposes. 2. Certificate of Confidentiality protects N3C data from certain types of disclosures.	Does not authorize/specify	Does not authorize/specify	Does not authorize/specify	Does not authorize/specify	Does not authorize/specify	Does not authorize/specify	Does not authorize/specify	Does not authorize/specify	Does not authorize/specify	Does not authorize/specify	Does not authorize/specify	Does not authorize/specify	Does not authorize/specify	Does not authorize/specify	Does not authorize/specify	N/A	Does not authorize/specify	Does not authorize/specify	Does not authorize/specify

TABLE 2: USE CASE 3 - DATASET LINKAGE DETERMINATION

Definitions: Limitations are restrictions on data linkage and use (e.g., dataset must only be linked with other disease-relevant data, dataset must be used in a physical enclave, etc.). **Controls** are processes established to ensure compliance with governance for data sharing, access, and use (e.g., user must access data in a physical enclave, user must sign data use agreement, user must receive data access committee approval, etc.). **Authorization gaps** exist when there is no explicit authorization or the authorization is not available or found in the information collected by the project team for various data life cycle stages (data collection, linking, sharing, access and use).

		Can the datasets be linked?	What limitations do the linked datasets inherit?	What controls do the linked dataset require?	What authorization gaps exist?
1	N3C (Dataset 1) and PEDSnet (Dataset 2) linkage	**Yes, N3C and PEDSnet can be linked provided:** **A. N3C staff:** 1. Shares N3C limited datasets (LDS), de-identified datasets, or synthetic datasets [Control 1a] 2. Removes ZIP codes entirely for all geographic units containing 20,000 or fewer people and replaces full five-digit ZIP codes of predominantly AI/AN communities with partial ZIP codes [Control 1b] 3. Has waiver of consent from NIH IRB for sharing data through the NCATS N3C Platform [Control 1c] **B. Data providers:** 4. Execute a Data Transfer Agreement (DTA) with NCATS [Control 1D] 5. Obtain institutional or external IRB approval [Control 1E] **C. PEDSnet staff:** 6. Removes HIV-related data and reproductive and mental health care data for minors [Control 2a] 7. De-identifies individual level data using the Safe Harbor method of de-identification of PHI [Control 2b] 8. Performs a risk review on the requested datasets as well as data transformations, such as date shifts, replacement labels for free text fields and geographic information, and removing HIV/pregnancy/mental health data [Control 2c] **D. The researcher/user:** 1. Uses the linked N3C and PEDSnet data for general COVID-19 research purposes specified and approved by PEDSnet participating sites and the PEDSnet Steering Committee [Limitations 1b and 2b] 2. Does not use the linked data to make assumptions about Tribal affiliation [Limitation 1c] 3. Complies with the N3C Community Guiding Principles and the Attribution and Publication Principles [Limitation 1e] 4. Executes the Institutional Data Use Agreements (DUA) with NCATS and PEDSnet and Responsible Use of Data Agreement (RUD) with PEDSnet [Controls 1f and 2f] 5. Submits Data Use Request (DUR) for approval by N3C Data Access Committee and request form for approval by the PEDSnet Research Committee [Controls 1g and 2d] 6. Completes NIH IT training, attests to the N3C Data User Code of Conduct, abides by N3C Results Download Policy, and completes Human Subjects Research Protection training to access N3C data [Control 1h] 7. Provides IRB letter of determination for N3C data access and if determined to be Human Subjects Research, provide IRB approval with IRB reliance for site providing data (NPRA Master Reliance Agreement (MRA) or SMART IRB MRA) for PEDSnet [Controls 1i and 2e] 8. Obtains approvals from PEDSnet prospective site PI and PEDSnet Executive Committee on the proposed linkage	**Researchers/users must:** 1a. Use N3C data within the N3C Enclave [N3C] 1b. Use N3C data for COVID-19 general research purposes [N3C] 1c. Not use AI/AN data and ZIP code information to make assumptions about Tribal affiliation [N3C] 1d. Work with N3C staff to link Class 2 or Class 0 data using PPRL [N3C] 1e. Comply with the N3C Community Guiding Principles and the Attribution and Publication Principles [N3C] 2a. Use the data in a workspace within the PEDSnet cloud enclave--OR--at their own institution if approved to have the data transferred to their institution by all PEDSnet institutions providing data for the request [PEDSnet] 2b. Use the data for purposes specified and approved by participating sites and the Steering Committee, namely using data from real-world clinical settings for research, quality measurement, and improvement/advancement of child health, particularly studies that inform or directly address clinical decision making, including retrospective observational studies. [PEDSnet] 2c. Work with PEDSnet staff to link data conducted under a waiver of consent using PPRL [PEDSnet]	**For sharing:** **A. N3C staff must:** 1a. Share N3C limited datasets (LDS), de-identified datasets, or synthetic datasets [N3C] 1b. Remove ZIP codes entirely for all geographic units containing 20,000 or fewer people and replace full five-digit ZIP codes of predominantly AI/AN communities with partial ZIP codes [N3C] 1c. Have waiver of consent from NIH IRB for sharing data through the NCATS N3C Platform [N3C] **B. Data providers must:** 1d. Execute a Data Transfer Agreement (DTA) with NCATS [N3C] 1e. Obtain institutional or external IRB approval [N3C] **For sharing, PEDSnet staff must:** 2a. Remove HIV-related data and reproductive and mental health care data for minors [PEDSnet] 2b. De-identify individual level data using the Safe Harbor method of de-identification of PHI [PEDSnet] 2c. Perform a risk review on the requested datasets as well as data transformations, such as date shifts, replacement labels for free text fields and geographic information, and removing HIV/pregnancy/mental health data [PEDSnet] **For accessing N3C, researchers/users must:** 1f. Execute Institutional Data Use Agreement (DUA) with NCATS [N3C] 1g. Submit Data Use Request (DUR) for approval by N3C Data Access Committee [N3C] 1h. Complete NIH IT training, attest to the N3C Data User Code of Conduct, abide by the N3C Results Download Policy, and complete Human Subjects Research Protection training [N3C] 1i. Provide IRB letter of determination for data access [N3C] 1j. Access the data within the N3C Enclave [N3C] **For linking N3C data, researchers/users must:** 1k. Work with N3C staff to verify and complete the following requirements for N3C Class 0 or Class 2 linkages [N3C]: i. Existing institutional N3C Data Use Agreement ii. Dual authentication and authorization iii. Signed institutional linkage honest broker agreement for multiple datasets iv. Approved data use request (DUR) by the federally staffed Data Access Committee (DAC) v. Local institution's IRB letter of determination vi. Interconnect agreement (for Class 0 only) 1l. Have agreement from participating sites for linkage with the external dataset [N3C] 1m. Have approval from the External Dataset Committee in the Tools and Resource subgroup and NCATS for linkage [N3C]	No authorization gaps exist

TABLE 2: USE CASE 3 - DATASET LINKAGE DETERMINATION

Definitions: Limitations are restrictions on data linkage and use (e.g., dataset must only be linked with other disease-relevant data, dataset must be used in a physical enclave, etc.). **Controls** are processes established to ensure compliance with governance for data sharing, access, and use (e.g., user must access data in a physical enclave, user must sign data use agreement, user must receive data access committee approval, etc.). **Authorization gaps** exist when there is no explicit authorization or the authorization is not available or found in the information collected by the project team for various data life cycle stages (data collection, linking, sharing, access and use).

		Can the datasets be linked?	What limitations do the linked datasets inherit?	What controls do the linked dataset require?	What authorization gaps exist?
		[Controls 2g and 2h] 9. Works with N3C staff to obtain Class 2 designation for PEDSnet data so that it can be linked using PPRL with N3C data [Limitations 1d and 2c] - **Assumption** 10. If PEDSnet data are designated as Class 2, uses/accesses N3C data within the N3C Enclave and obtains approval from PEDSnet staff to export PEDSnet data into the N3C Enclave [Limitations 1a and 2a, Controls 1j and 2i] - **Assumption for PEDSnet** 11. Ensures they have an existing institutional N3C Data Use Agreement, dual authentication and authorization, signed institutional linkage honest broker agreement for multiple datasets, an approved data use request (DUR) by the data access committee (DAC), and local institution's IRB letter of determination for N3C Class 2 or Class 0 designation for PEDSnet linkage. If PEDSnet data are designated as Class 0, ensures they also have an Interconnect agreement. [Control 1k] 12. Has approval from participating sites for linkage with the external dataset and the External Dataset Committee in the Tools and Resource subgroup and NCATS for PEDSnet linkage [Limitations 1l and 1m]		**For accessing PEDSnet, researchers/users must:** 2d. Submit request form for approval by the Research Committee [PEDSnet] 2e. Undergo IRB review/determination (Human Subjects Review, HSR) [PEDSnet] i. If IRB determines the proposed study is NHSR, then no further review/MRA required [PEDSnet] ii. If IRB determines the proposed study HSR, the requester must provide IRB approval with IRB reliance for site providing data (NPRA MRA or SMART IRB MRA) [PEDSnet] 2f. Sign DUA (Data Use Agreement) and RUD (Responsible Use of Data) (Legal Review) [PEDSnet] 2g. Receive prospective site PI approval (Institutional Participation Approval) [PEDSnet] 2h. Receive PEDSnet Executive Committee approval (Network Participation Approval) [PEDSnet] 2i. Access the data through a workspace within the PEDSnet cloud enclave--OR--have the data transferred to their institution, the PEDSnet Study Approval request should specify, pending approval from all PEDSnet institutions providing data for the request [PEDSnet]	
2	N3C (Dataset 1) and RADx-UP (Dataset 3) linkage	**Yes, N3C and RADx-UP can be linked provided:** **A. N3C staff:** 1. Shares N3C limited datasets (LDS), de-identified datasets, or synthetic datasets [Control 1a] 2. Removes ZIP codes entirely for all geographic units containing 20,000 or fewer people and replaces full five-digit ZIP codes of predominantly AI/AN communities with partial ZIP codes [Control 1b] 3. Has waiver of consent from NIH IRB for sharing data through the NCATS N3C Platform [Control 1c] **B. Data providers:** 4. Execute a Data Transfer Agreement (DTA) with NCATS [Control 1D] 5. Obtain institutional or external IRB approval [Control 1E] **C. RADx Data Hub staff:** 6. Ensures the studies are registered in dbGaP [Control 3a] 7. Ensures that the data is de-identified by working with study teams to de-identify zip codes, shift dates, and adjust ages into categories for specific ages [Control 3b] **D. The researcher/user:** 1. Uses the linked N3C and RADx-UP data for general COVID-19 research purposes [Limitations 1b and 3a]	**Researchers/users must:** 1a. Use N3C data within the N3C Enclave [N3C] 1b. Use N3C data for COVID-19 general research purposes [N3C] 1c. Not use AI/AN data and ZIP code information to make assumptions about Tribal affiliation [N3C] 1d. Work with N3C staff to link Class 2 or Class 0 data using PPRL [N3C] 1e. Comply with the N3C Community Guiding Principles and the Attribution and Publication Principles [N3C] 3a. Use RADx-UP data for general research purposes [RADx-UP]	**For sharing:** **A. N3C staff must:** 1a. Share N3C limited datasets (LDS), de-identified datasets, or synthetic datasets [N3C] 1b. Remove ZIP codes entirely for all geographic units containing 20,000 or fewer people and replace full five-digit ZIP codes of predominantly AI/AN communities with partial ZIP codes [N3C] 1c. Have waiver of consent from NIH IRB sharing data through the NCATS N3C Platform [N3C] **B. Data providers must:** 1d. Execute a Data Transfer Agreement (DTA) with NCATS [N3C] 1e. Obtain institutional or external IRB approval [N3C] **For sharing, RADx Data Hub stuff must:** 3a. Ensure the studies are registered in dbGaP [RADx-UP] 3b. Ensure that the data is de-identified by working with study teams to de-identify zip codes, shift dates, and adjust ages into categories for specific ages [RADx-UP] **For accessing N3C, researchers/users must:** 1f. Execute Institutional Data Use Agreement (DUA) with NCATS [N3C] 1g. Submit Data Use Request (DUR) for approval by N3C Data Access Committee [N3C] 1h. Complete NIH IT training, attest to the N3C Data User	No authorization gaps exist

TABLE 2: USE CASE 3 - DATASET LINKAGE DETERMINATION

Definitions: Limitations are restrictions on data linkage and use (e.g., dataset must only be linked with other disease-relevant data, dataset must be used in a physical enclave, etc.). **Controls** are processes established to ensure compliance with governance for data sharing, access, and use (e.g., user must access data in a physical enclave, user must sign data use agreement, user must receive data access committee approval, etc.). **Authorization gaps** exist when there is no explicit authorization or the authorization is not available or found in the information collected by the project team for various data life cycle stages (data collection, linking, sharing, access and use).

		Can the datasets be linked?	What limitations do the linked datasets inherit?	What controls do the linked dataset require?	What authorization gaps exist?
		2. Does not use the linked data to make assumptions about Tribal affiliation [Limitation 1c] 3. Complies with the N3C Community Guiding Principles and the Attribution and Publication Principles [Limitation 1e] 4. Executes the Institutional Data Use Agreements (DUA) with NCATS [Control 1f] 5. Has an eRA commons or Login.gov account [Control 3c] 6. Submits Data Use Request (DUR) for approval by N3C Data Access Committee and a Data Access Request (DAR), which includes the Data Use Certification (DUC) Agreement, the Genomic Data User Code of Conduct, and the RADx SM Data User Code of Conduct, for RADx-UP [Controls 1g and 3d] 7. Ensure the Signing Official from the investigator's institution reviews, approves, and co-signs the request [Control 3e] 8. Completes NIH IT training, attests to the N3C Data User Code of Conduct, abides by the N3C Results Download Policy, and completes Human Subjects Research Protection training to access N3C data [Control 1h] 9. Provides IRB letter of determination for N3C data access [Control 1i] 10. Obtains approvals from the AHARO Center/Comprehensive Health Center IRB and the RADx Data Hub Data Access Committee for the proposed linkage [Controls 3h and 3f] 11. Works with N3C staff to obtain Class 2 designation for RADx-UP data so that it can be linked using PPRL with N3C [Limitation 1d] - Assumption 12. If RADx-UP data are designated as Class 2, uses/accesses N3C data within the N3C Enclave and obtains approval from RADx Data Hub staff to export RADx-UP data into the N3C Enclave [Limitation 1a, Controls 1j and 3g] - Assumption for RADx-UP 13. Ensures they have an existing institutional N3C Data Use Agreement, dual authentication and authorization, signed institutional linkage honest broker agreement for multiple datasets, an approved data use request (DUR) by the data access committee (DAC), and local institutions IRB letter of determination for N3C Class 2 or Class 0 designation for RADx-UP linkage. If RADx-UP data are designated as Class 0, ensures they also have an Interconnect agreement. [Control 1k] 14. Has approval from participating sites for linkage with the external dataset and the External Dataset Committee in the Tools and Resource subgroup and NCATS for RADx-UP linkage [Limitations 1l and 1m]		Code of Conduct, abide by the N3C Results Download Policy, and complete Human Subjects Research Protection training [N3C] 1i. Provide IRB letter of determination for data access [N3C] 1j. Access the data within the N3C Enclave [N3C] **For linking N3C data, researchers/users must:** 1k. Work with N3C staff to verify and complete the following requirements for N3C Class 0 or Class 2 linkages [N3C]: i. Existing institutional N3C Data Use Agreement ii. Dual authentication and authorization iii. Signed institutional linkage honest broker agreement for multiple datasets iv. Approved data use request (DUR) by the federally staffed Data Access Committee (DAC) v. Local institution's IRB letter of determination vi. Interconnect agreement (for Class 0 only) 1l. Have agreement from participating sites for linkage with the external dataset [N3C] 1m. Have approval from the External Dataset Committee in the Tools and Resource subgroup and NCATS for linkage [N3C] **For accessing RADx-UP data, researchers/users must:** 3c. Have an eRA commons or Login.gov account [RADx-UP] 3d. Submit a Data Access Request (DAR), which includes the Data Use Certification (DUC) Agreement, the Genomic Data User Code of Conduct, and the RADx SM Data User Code of Conduct [RADx-UP] 3e. Ensure the Signing Official from the investigator's institution reviews, approves, and co-signs the request [RADx-UP] 3f. Receive approval from the Data Access Committee [RADx-UP] 3g. Access the data through RADx Data Hub Jupyter Notebooks [RADx-UP] **For linking RADx-UP data, researchers/users must:** 3h. Work with AHARO Center/Comprehensive Health Center IRB to obtain approval for individual level data linkages [RADx-UP]	

TABLE 2: USE CASE 3 - DATASET LINKAGE DETERMINATION

Definitions: Limitations are restrictions on data linkage and use (e.g., dataset must only be linked with other disease-relevant data, dataset must be used in a physical enclave, etc.). **Controls** are processes established to ensure compliance with governance for data sharing, access, and use (e.g., user must access data in a physical enclave, user must sign data use agreement, user must receive data access committee approval, etc.). **Authorization gaps** exist when there is no explicit authorization or the authorization is not available or found in the information collected by the project team for various data life cycle stages (data collection, linking, sharing, access and use).

		Can the datasets be linked?	What limitations do the linked datasets inherit?	What controls do the linked dataset require?	What authorization gaps exist?
3	N3C (Dataset 1) and EPA (Dataset 4) linkage	**Yes, N3C and EPA can be linked provided:** **A. N3C staff:** 1. Shares N3C limited datasets (LDS), de-identified datasets, or synthetic datasets [Control 1a] 2. Removes ZIP codes entirely for all geographic units containing 20,000 or fewer people and replaces full five-digit ZIP codes of predominantly AI/AN communities with partial ZIP codes [Control 1b] 3. Has waiver of consent from NIH IRB for sharing data through the NCATS N3C Platform [Control 1c] **B. Data providers:** 4. Execute a Data Transfer Agreement (DTA) with NCATS [Control 1d] 5. Obtain institutional or external IRB approval [Control 1e] **C. EPA staff:** 6. Hosts ambient air data, containing full geographic identifiers including site address, zip code, CBSA, county and state, through EPA's Air Quality System (AQS) [Control 4a] **C. The researcher/user:** 1. Uses the linked N3C and EPA data for general COVID-19 research purposes [Limitation 1b] 2. Does not use the linked data to make assumptions about Tribal affiliation [Limitation 1c] 3. Complies with the N3C Community Guiding Principles and the Attribution and Publication Principles [Limitation 1e] 4. Uses/accesses the linked N3C and EPA data within the N3C enclave [Limitation 1a, Control 1j] 5. Executes the Institutional Data Use Agreements (DUA) with NCATS [Controls 1f] 6. Submits Data Use Request (DUR) for approval by N3C Data Access Committee [Control 1g] 7. Completes NIH IT training, attests to the N3C Data User Code of Conduct, abides by the N3C Results Download Policy, and completes Human Subjects Research Protection training to access N3C data [Control 1h] 8. Provides IRB letter of determination for N3C data access [Control 1i] 9. Has approval from participating sites for linkage with the external dataset and the External Dataset Committee in the Tools and Resource subgroup and NCATS for EPA linkage [Limitations 1l and 1m] Note: Controls 1k, 4a, and 4c are not required for this linkage as EPA Air Quality Data has already been brought into the N3C Enclave and is already available for linkage to N3C.	**Researchers/users must:** 1a. Use N3C data within the N3C Enclave [N3C] 1b. Use N3C data for COVID-19 general research purposes [N3C] 1c. Not use AI/AN data and ZIP code information to make assumptions about Tribal affiliation [N3C] 1d. Work with N3C staff to link Class 2 or Class 0 data using PPRL [N3C] 1e. Comply with the N3C Community Guiding Principles and the Attribution and Publication Principles [N3C]	**For sharing:** **A. N3C staff must:** 1a. Share N3C limited datasets (LDS), de-identified datasets, or synthetic datasets [N3C] 1b. Remove ZIP codes entirely for all geographic units containing 20,000 or fewer people and replace full five-digit ZIP codes of predominantly AI/AN communities with partial ZIP codes [N3C] 1c. Have waiver of consent from NIH IRB for sharing data through the NCATS N3C Platform [N3C] **B. Data providers must:** 1d. Execute a Data Transfer Agreement (DTA) with NCATS [N3C] 1e. Obtain institutional or external IRB approval [N3C] **For sharing, EPA staff must:** 4a. Host ambient air data, which contains full geographic identifiers including site address, zip code, CBSA, county, and state, through EPA's Air Quality System (AQS) [EPA] **For accessing N3C, researchers/users must:** 1f. Execute Institutional Data Use Agreement (DUA) with NCATS [N3C] 1g. Submit Data Use Request (DUR) for approval by N3C Data Access Committee [N3C] 1h. Complete NIH IT training, attest to the N3C Data User Code of Conduct, abide by the N3C Results Download Policy, and complete Human Subjects Research Protection training [N3C] 1i. Provide IRB letter of determination for data access [N3C] 1j. Access the data within the N3C Enclave [N3C] **For linking N3C data, researchers/users must:** 1k. Work with N3C staff to verify and complete the following requirements for N3C Class 0 or Class 2 linkages [N3C]: i. Existing institutional N3C Data Use Agreement ii. Dual authentication and authorization iii. Signed institutional linkage honest broker agreement for multiple datasets iv. Approved data use request (DUR) by the federally staffed Data Access Committee (DAC) v. Local institution's IRB letter of determination vi. Interconnect agreement (for Class 0 only) 1l. Have agreement from participating sites for linkage with the external dataset [N3C] 1m. Have approval from the External Dataset Committee in the Tools and Resource subgroup and NCATS for linkage [N3C] **For accessing EPA data, researchers/users can:** 4c. Obtain data from AQS, an open access repository	No authorization gaps exist

TABLE 2: USE CASE 3 - DATASET LINKAGE DETERMINATION

Definitions: Limitations are restrictions on data linkage and use (e.g., dataset must only be linked with other disease-relevant data, dataset must be used in a physical enclave, etc.). **Controls** are processes established to ensure compliance with governance for data sharing, access, and use (e.g., user must access data in a physical enclave, user must sign data use agreement, user must receive data access committee approval, etc.). **Authorization gaps** exist when there is no explicit authorization or the authorization is not available or found in the information collected by the project team for various data life cycle stages (data collection, linking, sharing, access and use).

		Can the datasets be linked?	What limitations do the linked datasets inherit?	What controls do the linked dataset require?	What authorization gaps exist?
4	PEDSnet (Dataset 2) and RADx-UP (Dataset 3) linkage	**Yes, PEDSnet and RADx-UP can be linked provided:** **A. PEDSnet staff:** 1. Removes HIV-related data and reproductive and mental health care data for minors [Control 2a] 2. De-identifies individual level data using the Safe Harbor method of de-identification of PHI [Control 2b] 3. Performs a risk review on the requested datasets as well as data transformations, such as date shifts, replacement labels for free text fields and geographic information, and removing HIV/pregnancy/mental health data [Control 2c] **B. RADx Data Hub staff:** 4. Ensures the studies are registered in dbGaP [Control 3a] 5. Ensures that the data is de-identified by working with study teams to de-identify zip codes, shift dates, and adjust ages into categories for specific ages [Control 3b] **C. The researcher/user:** 1. Uses the linked PEDSnet and RADx-UP data for general research purposes specified and approved by PEDSnet participating sites and the PEDSnet Steering Committee [Limitations 2b and 3a] 2. Has an eRA commons or Login.gov account [Control 3c] 3. Submits request form for approval by the Research Committee and submits a Data Access Request (DAR), which includes the Data Use Certification (DUC) Agreement, the Genomic Data User Code of Conduct, and the RADx SM Data User Code of Conduct [Controls 2d and 3d] 4. Ensures that the Signing Official from investigator's institution reviews, approves, and co-signs the request [Control 3e] 5. Undergoes IRB review to receive a letter of determination for data access and if determined to be Human Subjects Research, provide IRB approval with IRB reliance for site providing data (NPRA Master Reliance Agreement (MRA) or SMART IRB MRA is also required) for PEDSnet [Control 2e] 6. Signs PEDSnet DUA and Responsible Use of Data Agreement (RUD) with PEDSnet [Control 2f] 7. Receives PEDSnet prospective site PI approval and PEDSnet Executive Committee approval, AHARO Center/ Comprehensive Health Center IRB and RADx Data Access Committee approval for the proposed PPRL linkage [Limitation 2c; Controls 2g, 2h, 3h, and 3f] 8. Works with PEDSnet staff and RADx Data Hub staff to determine whether RADx-UP data downloaded through Jupyter Notebooks can be transferred to the PEDSnet cloud enclave workspace --OR-- obtains approval to have PEDSnet data transferred to the user's institution to use with the downloaded RADx-UP data [Controls 2i and 3g] -- **Assumption** 9. Uses the linked PEDSnet and RADx-UP data within the PEDSnet cloud enclave workspace or at the user's institution with approval [Limitation 2a] -- **Assumption**	**Researchers/users must:** 2a. Use the data in a workspace within the PEDSnet cloud enclave--OR--at their own institution if approved to have the data transferred to their institution by all PEDSnet institutions providing data for the request [PEDSnet] 2b. Use the data for purposes specified and approved by participating sites and the Steering Committee, namely using data from real-world clinical settings for research, quality measurement, and improvement/advancement of child health, particularly studies that inform or directly address clinical decision making, including retrospective observational studies. [PEDSnet] 2c. Work with PEDSnet staff to link data conducted under a waiver of consent using PPRL [PEDSnet] 3a. Use RADx-UP data for general research purposes [RADx-UP]	**For sharing, PEDSnet staff must:** 2a. Remove HIV-related data and reproductive and mental health care data for minors [PEDSnet] 2b. De-identify individual level data using the Safe Harbor method of de-identification of PHI [PEDSnet] 2c. Perform a risk review on the requested datasets as well as data transformations, such as date shifts, replacement labels for free text fields and geographic information, and removing HIV/pregnancy/ mental health data [PEDSnet] **For sharing, RADx Data Hub staff must:** 3a. Ensure the studies are registered in dbGaP [RADx-UP] 3b. Ensure that the data is de-identified by working with study teams to de-identify zip codes, shift dates, and adjust ages into categories for specific ages [RADx-UP] **For accessing PEDSnet, researchers/users must:** 2d. Submit request form for approval by the Research Committee [PEDSnet] 2e. Undergo IRB review/determination (Human Subjects Review, HSR) [PEDSnet] i. If IRB determines the proposed study is NHSR, then no further review/MRA required [PEDSnet] ii. If IRB determines the proposed study HSR, the requester must provide IRB approval with IRB reliance for site providing data (NPRA MRA or SMART IRB MRA) [PEDSnet] 2f. Sign DUA (Data Use Agreement) and RUD (Responsible Use of Data) (Legal Review) [PEDSnet] 2g. Receive prospective site PI approval (Institutional Participation Approval) [PEDSnet] 2h. Receive PEDSnet Executive Committee approval (Network Participation Approval) [PEDSnet] 2i. Access the data through a workspace within the PEDSnet cloud enclave--OR--have the data transferred to their institution, the PEDSnet Study Approval request should specify, pending approval from all PEDSnet institutions providing data for the request [PEDSnet] **For accessing RADx-UP data, researchers/users must:** 3c. Have an eRA commons or Login.gov account [RADx-UP] 3d. Submit a Data Access Request (DAR), which includes the Data Use Certification (DUC) Agreement, the Genomic Data User Code of Conduct, and the RADx SM Data User Code of Conduct [RADx-UP] 3e. Ensure the Signing Official from the investigator's institution reviews, approves, and co-signs the request [RADx-UP] 3f. Receive approval from the Data Access Committee [RADx-UP] 3g. Access the data through RADx Data Hub Jupyter Notebooks [RADx-UP] **For linking RADx-UP data, researchers/users must:** 3h. Work with AHARO Center/Comprehensive Health Center IRB to obtain approval for individual level data linkages [RADx-UP]	No authorization gaps exist

TABLE 2: USE CASE 3 - DATASET LINKAGE DETERMINATION

Definitions: Limitations are restrictions on data linkage and use (e.g., dataset must only be linked with other disease-relevant data, dataset must be used in a physical enclave, etc.). **Controls** are processes established to ensure compliance with governance for data sharing, access, and use (e.g., user must access data in a physical enclave, user must sign data use agreement, user must receive data access committee approval, etc.). **Authorization gaps** exist when there is no explicit authorization or the authorization is not available or found in the information collected by the project team for various data life cycle stages (data collection, linking, sharing, access and use).

		Can the datasets be linked?	What limitations do the linked datasets inherit?	What controls do the linked dataset require?	What authorization gaps exist?
5	PEDSnet (Dataset 2) and EPA (Dataset 4) linkage	**Yes, PEDSnet and EPA can be linked provided:** **A. PEDSnet staff:** 1. Removes HIV-related data and reproductive and mental health care data for minors [Control 2a] 2. De-identifies individual level data using the Safe Harbor method of de-identification of PHI [Control 2b] 3. Performs a risk review on the requested datasets as well as data transformations, such as date shifts, replacement labels for free text fields and geographic information, and removing HIV/pregnancy/mental health data [Control 2c] **B. EPA staff:** 4. Hosts ambient air data, containing full geographic identifiers including site address, zip code, CBSA, county and state, through EPA's Air Quality System (AQS) [Control 4a] **C. The researcher/user:** 1. Uses the linked PEDSnet and EPA data for purposes specified and approved by PEDSnet participating sites and the PEDSnet Steering Committee [Limitation 2b] 2. Obtains the EPA data from AQS, an open access repository [Control 4c] 3. Submits request form for approval by the Research Committee [Control 2d] 4. Undergoes IRB review to receive a letter of determination for data access and if determined to be Human Subjects Research, provide IRB approval with IRB reliance for site providing data (NPRA Master Reliance Agreement (MRA) or SMART IRB MRA is also required) for PEDSnet [Control 2e] 5. Signs PEDSnet DUA and Responsible Use of Data Agreement (RUD) with PEDSnet [Control 2f] 6. Receives prospective site PI approval and PEDSnet Executive Committee approval for the proposed linkage [Controls 2g and 2h] 7. Works with PEDSnet staff to determine whether EPA data can be transferred to the PEDSnet cloud enclave workspace --OR-- obtain approval to have PEDSnet data transferred to the user's institution to use with the EPA data [Control 2i] -- **Assumption** 8. Uses the linked PEDSnet and EPA data within the PEDSnet cloud enclave workspace or at the user's institution with approval [Limitation 2a] -- **Assumption** **Note**: Limitation 2c is not required as EPA data is location based and would not be linked through PPRL	**Researchers/users must:** 2a. Use the data in a workspace within the PEDSnet cloud enclave--OR--at their own institution if approved to have the data transferred to their institution by all PEDSnet institutions providing data for the request [PEDSnet] 2b. Use the data for purposes specified and approved by participating sites and the Steering Committee, namely using data from real-world clinical settings for research, quality measurement, and improvement/advancement of child health, particularly studies that inform or directly address clinical decision making, including retrospective observational studies. [PEDSnet] 2c. Work with PEDSnet staff to link data conducted under a waiver of consent using PPRL [PEDSnet]	**For sharing, PEDSnet staff must:** 2a. Remove HIV-related data and reproductive and mental health care data for minors [PEDSnet] 2b. De-identify individual level data using the Safe Harbor method of de-identification of PHI [PEDSnet] 2c. Perform a risk review on the requested datasets as well as data transformations, such as date shifts, replacement labels for free text fields and geographic information, and removing HIV/pregnancy/ mental health data [PEDSnet] **For sharing, EPA staff must:** 4a. Host ambient air data, which contains full geographic identifiers including site address, zip code, CBSA, county, and state, through EPA's Air Quality System (AQS) [EPA] **For accessing PEDSnet, researchers/users must:** 2d. Submit request form for approval by the Research Committee [PEDSnet] 2e. Undergo IRB review/determination (Human Subjects Review) [PEDSnet] i. If IRB determines the proposed study is NHSR, then no further review/MRA required [PEDSnet] ii. If IRB determines the proposed study HSR, the requester must provide IRB approval with IRB reliance for site providing data (NPRA MRA or SMART IRB MRA) [PEDSnet] 2f. Sign DUA (Data Use Agreement) and RUD (Responsible Use of Data) (Legal Review) [PEDSnet] 2g. Receive prospective site PI approval (Institutional Participation Approval) [PEDSnet] 2h. Receive PEDSnet Executive Committee approval (Network Participation Approval) [PEDSnet] 2i. Access the data through a workspace within the PEDSnet cloud enclave--OR--have the data transferred to their institution, the PEDSnet Study Approval request should specify, pending approval from all PEDSnet institutions providing data for the request [PEDSnet] **For accessing EPA data, researchers/users can:** 4c. Obtain data from AQS, an open access repository	No authorization gaps exist

TABLE 2: USE CASE 3 - DATASET LINKAGE DETERMINATION

Definitions: Limitations are restrictions on data linkage and use (e.g., dataset must only be linked with other disease-relevant data, dataset must be used in a physical enclave, etc.). **Controls** are processes established to ensure compliance with governance for data sharing, access, and use (e.g., user must access data in a physical enclave, user must sign data use agreement, user must receive data access committee approval, etc.). **Authorization gaps** exist when there is no explicit authorization or the authorization is not available or found in the information collected by the project team for various data life cycle stages (data collection, linking, sharing, access and use).

		Can the datasets be linked?	What limitations do the linked datasets inherit?	What controls do the linked dataset require?	What authorization gaps exist?
6	RADx-UP (Dataset 3) and EPA (Dataset 4) linkage	**Yes, RADx-UP and EPA can be linked provided:** **A. RADx Data Hub staff:** 1. Ensures the studies are registered in dbGaP [Control 3a] 2. Ensures that the data is de-identified by working with study teams to de-identify zip codes, shift dates, and adjust ages into categories for specific ages [Control 3b] **B. EPA staff:** 3. Hosts ambient air data, containing full geographic identifiers including site address, zip code, CBSA, county and state, through EPA's Air Quality System (AQS) [Control 4a] **C. The researcher/user:** 1. Uses the linked RADx-UP and EPA data for general research purposed [Limitation 3a] 2. Obtains the EPA data from AQS, an open access repository [Control 4c] 3. Has an eRA commons or Login.gov account [Control 3c] 4. Submits a Data Access Request (DAR), which includes the Data Use Certification (DUC) Agreement, the Genomic Data User Code of Conduct, and the RADx SM Data User Code of Conduct [Control 3d] 5. Ensures the Signing Official from the investigator's institution reviews, approves, and co-signs the request [Control 3e] 6. Receives approval from the AHARO Center/Comprehensive Health Center IRB and the Data Access Committee for the proposed linkage [Controls 3h and 3f] 7. Access the linked data through RADx Data Hub Jupyter Notebooks [Control 3g]	**Researchers/users must:** 3a. Use RADx-UP data for general research purposes [RADx-UP]	**For sharing, RADx Data Hub staff must:** 3a. Ensure the studies are registered in dbGaP [RADx-UP] 3b. Ensure that the data is de-identified by working with study teams to de-identify zip codes, shift dates, and adjust ages into categories for specific ages [RADx-UP] **For sharing, EPA staff must:** 4a. Host ambient air data, which contains full geographic identifiers including site address, zip code, CBSA, county, and state, through EPA's Air Quality System (AQS) [EPA] **For accessing RADx-UP data, researchers/users must:** 3c. Have an eRA commons or Login.gov account [RADx-UP] 3d. Submit a Data Access Request (DAR), which includes the Data Use Certification (DUC) Agreement, the Genomic Data User Code of Conduct, and the RADx SM Data User Code of Conduct [RADx-UP] 3e. Ensure the Signing Official from the investigator's institution reviews, approves, and co-signs the request [RADx-UP] 3f. Receive approval from the Data Access Committee [RADx-UP] 3g. Access the data through RADx Data Hub Jupyter Notebooks [RADx-UP] **For linking RADx-UP data, researchers/users must:** 3h. Work with AHARO Center/Comprehensive Health Center IRB to obtain approval for individual level data linkages [RADx-UP] **For accessing EPA data, researchers/users can:** 4c. Obtain data from AQS, an open access repository	No authorization gaps exist
7	All datasets (N3C, PEDSnet, RADx-UP, EPA)	**Yes, N3C, PEDSnet, RADx-UP, and EPA can be linked provided:** **A. N3C/PEDSnet/RADx Data Hub/EPA staff:** 1. Shares N3C, PEDSnet, and RADx data de-identified of all direct identifers - for N3C, limited datasets or synthetic datasets can also be shared; ZIP codes entirely for all geographic units containing 20,000 or fewer people should be removed; and full five-digit ZIP codes of predominantly AI/AN communities should be replaced with partial ZIP codes - for PEDSnet, HIV-related data and reproductive and mental health care data for minors should be removed - for EPA, full geographic identifiers including site address, zip code, CBSA, county, and state are shared [Controls 1a, 1b, 2a, 2b, 3b, and 4a] 2. Has waiver of consent from NIH IRB for sharing data through the NCATS N3C Platform [Control 1c] 3. Performs risk review prior to sharing of PEDSnet data (data transformations, such as date shifts, replacement labels for free text fields and geographic information, and removing HIV/pregnancy/mental health data) [Control 2c] 4. Ensures the RADx studies are registered in dbGaP [Control 3a] **B. N3C Data Providers:** 5. Execute a Data Transfer Agreement (DTA) with NCATS	**Researchers/users must:** 1a. Use N3C data within the N3C Enclave [N3C] 1b. Use N3C data for COVID-19 general research purposes [N3C] 1c. Not use AI/AN data and ZIP code information to make assumptions about Tribal affiliation [N3C] 1d. Work with N3C staff to link Class 2 or Class 0 data using PPRL [N3C] 1e. Comply with the N3C Community Guiding Principles and the Attribution and Publication Principles [N3C] 2a. Use the data in a workspace within the PEDSnet cloud enclave--OR--at their own institution if approved to have the data transferred to their institution by all PEDSnet institutions providing data for the request [PEDSnet] 2b. Use the data for purposes specified and approved by participating sites and the Steering Committee, namely using data from real-world clinical settings for research, quality measurement, and improvement/advancement of child health, particularly studies that inform or directly address clinical decision making, including retrospective observational studies. [PEDSnet] 2c. Work with PEDSnet staff to link data conducted under a waiver of consent using PPRL [PEDSnet] 3a. Use RADx-UP data for general research purposes [RADx-UP]	**For sharing:** **A. N3C staff must:** 1a. Share N3C limited datasets (LDS), de-identified datasets, or synthetic datasets [N3C] 1b. Remove ZIP codes entirely for all geographic units containing 20,000 or fewer people and replace full five-digit ZIP codes of predominantly AI/AN communities with partial ZIP codes [N3C] 1c. Have waiver of consent from NIH IRB for sharing data through the NCATS N3C Platform [N3C] **B. Data providers must:** 1d. Execute a Data Transfer Agreement (DTA) with NCATS [N3C] 1e. Obtain institutional or external IRB approval [N3C] **For sharing, PEDSnet staff must:** 2a. Remove HIV-related data and reproductive and mental health care data for minors [PEDSnet] 2b. De-identify individual level data using the Safe Harbor method of de-identification of PHI [PEDSnet] 2c. Perform a risk review on the requested datasets as well as data transformations, such as date shifts, replacement labels for free text fields and geographic information, and removing HIV/pregnancy/mental health data [PEDSnet] **For sharing, RADx Data Hub stuff must:**	No authorization gaps exist

TABLE 2: USE CASE 3 - DATASET LINKAGE DETERMINATION

Definitions: Limitations are restrictions on data linkage and use (e.g., dataset must only be linked with other disease-relevant data, dataset must be used in a physical enclave, etc.). **Controls** are processes established to ensure compliance with governance for data sharing, access, and use (e.g., user must access data in a physical enclave, user must sign data use agreement, user must receive data access committee approval, etc.). **Authorization gaps** exist when there is no explicit authorization or the authorization is not available or found in the information collected by the project team for various data life cycle stages (data collection, linking, sharing, access and use).

		Can the datasets be linked?	What limitations do the linked datasets inherit?	What controls do the linked dataset require?	What authorization gaps exist?
		[Control 1d] 6. Obtain institutional or external IRB approval [Control 1e] **C. The researcher/user:** 1. Uses the linked N3C and PEDSnet data for general COVID-19 research purposes specified and approved by the PEDSnet participating sites and PEDSnet Steering Committee [Limitations 1b, 2b, and 3a] 2. Does not use the linked data to make assumptions about Tribal affiliation [Limitation 1c] 3. Complies with the N3C Community Guiding Principles and the Attribution and Publication Principles [Limitation 1e] 4. Has an eRA commons or Login.gov account [Control 3c] 5. Executes the Institutional Data Use Agreements (DUA) with NCATS and PEDSnet and Responsible Use of Data Agreement (RUD) with PEDSnet [Controls 1f and 2f] 6. Submits Data Use Request (DUR) for approval by N3C Data Access Committee, request form for approval by the PEDSnet Research Committee, and Data Access Request (DAR), which includes the Data Use Certification (DUC) Agreement, the Genomic Data User Code of Conduct, and the RADx SM Data User Code of Conduct [Controls 1h, 2d, and 3d] 7. Ensures the Signing Official from the investigator's institution reviews, approves, and co-signs the request [Control 3e] 8. Completes NIH IT training, attests to the N3C Data User Code of Conduct, and completes Human Subjects Research Protection training to access N3C data [Control 1i] 9. Provides IRB letter of determination for N3C data access and if determined to be Human Subjects Research, provide IRB approval with IRB reliance for site providing data (NPRA Master Reliance Agreement (MRA) or SMART IRB MRA) for PEDSnet [Controls 1k and 2e] 10. Obtains approvals from PEDSnet prospective site PI approval and PEDSnet Executive Committee, the AHARO Center/Comprehensive Health Center IRB and RADx Data Hub Data Access Committee on the proposed linkage [Controls 2g, 2h, 3h, and 3f] 11. Works with N3C staff to obtain Class 2 or Class 0 designation for PEDSnet and RADx-UP so that all three datasets can be linked using PPRL [Limitations 1d and 2c] - **Assumption** 12. If non-N3C data are designated as Class 2, uses/accesses N3C data within the N3C Enclave and obtains approval from PEDSnet staff to export PEDSnet data into the N3C Enclave and RADx Data Hub staff to export RADx-UP data into the N3C Enclave [Limitations 1a and 2a; Controls 1j, 2i, and 3g] - **Assumption for PEDSnet and RADx-UP**		3a. Ensure the studies are registered in dbGaP [RADx-UP] 3b. Ensure that the data is de-identified by working with study teams to de-identify zip codes, shift dates, and adjust ages into categories for specific ages [RADx-UP] **For sharing, EPA staff must:** 4a. Host ambient air data, which contains full geographic identifiers including site address, zip code, CBSA, county, and state, through EPA's Air Quality System (AQS) [EPA] **For accessing N3C, researchers/users must:** 1f. Execute Institutional Data Use Agreement (DUA) with NCATS [N3C] 1g. Submit Data Use Request (DUR) for approval by N3C Data Access Committee [N3C] 1h. Complete NIH IT training, attest to the N3C Data User Code of Conduct, abide by the N3C Results Download Policy, and complete Human Subjects Research Protection training [N3C] 1i. Provide IRB letter of determination for data access [N3C] 1j. Access the data within the N3C Enclave [N3C] **For accessing PEDSnet, researchers/users must:** 2d. Submit request form for approval by the Research Committee [PEDSnet] 2e. Undergo IRB review/determination (Human Subjects Review) [PEDSnet] i. If IRB determines the proposed study is NHSR, then no further review/MRA required [PEDSnet] ii. If IRB determines the proposed study HSR, the requester must provide IRB approval with IRB reliance for site providing data (NPRA MRA or SMART IRB MRA) [PEDSnet] 2f. Sign DUA (Data Use Agreement) and RUD (Responsible Use of Data) (Legal Review) [PEDSnet] 2g. Receive prospective site PI approval (Institutional Participation Approval) [PEDSnet] 2h. Receive PEDSnet Executive Committee approval (Network Participation Approval) [PEDSnet] 2i. Access the data through a workspace within the PEDSnet cloud enclave--OR--have the data transferred to their institution, the PEDSnet Study Approval request should specify, pending approval from all PEDSnet institutions providing data for the request [PEDSnet] **For accessing RADx-UP data, researchers/users must:** 3c. Have an eRA commons or Login.gov account [RADx-UP] 3d. Submit a Data Access Request (DAR), which includes the Data Use Certification (DUC) Agreement, the Genomic Data User Code of Conduct, and the RADx SM Data User Code of Conduct [RADx-UP]	

TABLE 2: USE CASE 3 - DATASET LINKAGE DETERMINATION

Definitions: Limitations are restrictions on data linkage and use (e.g., dataset must only be linked with other disease-relevant data, dataset must be used in a physical enclave, etc.). **Controls** are processes established to ensure compliance with governance for data sharing, access, and use (e.g., user must access data in a physical enclave, user must sign data use agreement, user must receive data access committee approval, etc.). **Authorization gaps** exist when there is no explicit authorization or the authorization is not available or found in the information collected by the project team for various data life cycle stages (data collection, linking, sharing, access and use).

		Can the datasets be linked?	What limitations do the linked datasets inherit?	What controls do the linked dataset require?	What authorization gaps exist?
		13. Ensures they have an existing institutional N3C Data Use Agreement, dual authentication and authorization, signed institutional linkage honest broker agreement for multiple datasets, an approved data use request (DUR) by the data access committee (DAC), and local institutions IRB letter of determination for N3C Class 2 or Class 0 designation for PEDSnet and RADx-UP linkage. If non-N3C data are designated as Class 0, ensures they also have an Interconnect agreement. [Control 1k] 14. Has approval from participating sites for linkage with the external dataset and the External Dataset Committee in the Tools and Resource subgroup and NCATS for EPA linkage [Limitations 1l and 1m] Note: Controls 4a and 4c are not required for this linkage as EPA Air Quality Data has already been brought into the N3C Enclave and is already available for linkage to N3C.		3e. Ensure the Signing Official from the investigator's institution reviews, approves, and co-signs the request [RADx-UP] 3f. Receive approval from the Data Access Committee [RADx-UP] 3g. Access the data through RADx Data Hub Jupyter Notebooks [RADx-UP] **For accessing EPA data, researchers can:** 4c. Obtain data from AQS, an open access repository **For linking N3C data, researchers/users must:** 1k. Work with N3C staff to verify and complete the following requirements for N3C Class 0 or Class 2 linkages [N3C]: i. Existing institutional N3C Data Use Agreement ii. Dual authentication and authorization iii. Signed institutional linkage honest broker agreement for multiple datasets iv. Approved data use request (DUR) by the federally staffed Data Access Committee (DAC) v. Local institution's IRB letter of determination vi. Interconnect agreement (for Class 0 only) 1l. Have agreement from participating sites for linkage with the external dataset [N3C] 1m. Have approval from the External Dataset Committee in the Tools and Resource subgroup and NCATS for linkage [N3C] **For linking RADx-UP data, researchers/users must:** 3h. Work with AHARO Center/Comprehensive Health Center IRB to obtain approval for individual level data linkages [RADx-UP]	

USE CASE 3 - GOVERNANCE INFORMATION				
Use Case 3: SARS-CoV-2 Vaccination and Asthma-Related School Absence - Does SARS-CoV-2 vaccination result in reduced asthma-related school absences at 3/6/12+ months post-vaccination?				
Dataset 8 - N3C (Limited Data Set)				
	Dataset Source	National COVID Cohort Collaborative (N3C)		
	Dataset Source Agency	NIH, NCATS		
	Dataset Type (Clinical, EHR, Survey, SDOH, etc.)	Clinical, EHR Data		
	Information Sources	Website, NCATS's Responses to N3C Questions, Information from predecessor report, NCATS staff email communication and meeting		
Dataset 8 - N3C				
		Raw Language	**Interpretation**	**Source**
1	**Data Collection**			
1.1	**Authorizations and Applicable Regulations/Policies**			
1.1.1	**Authorizations**		1. HIPAA	
1.1.1.1	Assent	N/A	N/A	
1.1.1.2	Consent	N/A -- Participating institutions do not obtain consent from individual patients for the data they send to the N3C. The 1996 Health Insurance Portability and Accountability Act (HIPAA) allows medical and health care institutions to release data for research without obtaining an individual's authorization if direct identifying information is removed and appropriate oversight and agreements are in place. Under the HIPAA Privacy Rule requirements, these institutions can release what is called a Limited Data Set. This is what participating health sites send to the N3C. [1]	N/A	[1] https://ncats.nih.gov/n3c/about (Accessed 4/21/23)
1.1.1.3	IRB/equivalent Privacy Board determination	N/A	N/A	
1.1.1.4	Local/state/federal laws	The 1996 Health Insurance Portability and Accountability Act (HIPAA) allows medical and health care institutions to release data for research without obtaining an individual's authorization if direct identifying information is removed and appropriate oversight and agreements are in place. [1] Under the HIPAA privacy regulations for a Limited Data Set, de-identified health information may be used and disclosed for research purposes. [2]	HIPAA authorizes health care providers to release data to N3C	[1] https://ncats.nih.gov/n3c/about (Accessed 4/21/23) [2] https://ncats.nih.gov/n3c/about/program-faq#use-the-data (Accessed 4/21/23)
1.1.1.5	Institutional Certification	N/A	N/A	
1.1.1.6	Data originator agreement	N/A	N/A	
1.1.1.7	Repository agreements/policies	N/A	N/A	
1.1.1.8	Other (specify)	Information not available/found	Information not available/found	
1.1.2	**Applicable Regulations/Policies**			
1.1.2.1	Local regulations/policies	Information not available/found	Information not available/found	
1.1.2.2	Tribal regulations/policies	Information not available/found	Information not available/found	
1.1.2.3	State regulations/policies	Information not available/found	Information not available/found	
1.1.2.4	Federal regulations/policies	Under the HIPAA privacy regulations for a Limited Data Set, de-identified health information may be used and disclosed for research purposes. [1] Data-contributing sites abide by the HIPAA Privacy Rule; Data are provided as a HIPAA-defined limited data set [2]	HIPAA Privacy Rule	[1] https://ncats.nih.gov/n3c/about/program-faq#use-the-data (Accessed 4/21/23) [2] NCATS Email Communication
1.1.2.5	International regulations/policies	Information not available/found	Information not available/found	
1.1.2.6	Contractual obligations	Information not available/found	Information not available/found	
1.1.2.7	Repository policies	Information not available/found	Information not available/found	

Dataset 8 - N3C				
		Raw Language	Interpretation	Source
1.2	**Governance for data linkage, sharing, access, and use based on data collection authorization or applicable regulations/policies (i.e., the origin of the governance)**			
1.2.1	Whether the data can be linked	Does not authorize/specify	Does not authorize/specify	
1.2.2	With what other data can it be linked or can it not be linked (scope of linkage)	Does not authorize/specify	Does not authorize/specify	
1.2.3	Whether data can be shared	Does not authorize/specify	Does not authorize/specify	
1.2.4	How data can be shared (de-identification status, disclosure review)	The 1996 Health Insurance Portability and Accountability Act (HIPAA) allows medical and health care institutions to release data for research without obtaining an individual's authorization if direct identifying information is removed and appropriate oversight and agreements are in place. [1] Under the HIPAA privacy regulations for a Limited Data Set, de-identified health information may be used and disclosed for research purposes. [2]	HIPAA specifies that N3C data partners (health care providers) can release limited EHR dataset (with no direct PII) for research purposes	[1] https://ncats.nih.gov/n3c/about (Accessed 4/21/23) [2] https://ncats.nih.gov/n3c/about/program-faq#use-the-data (Accessed 4/21/23)
1.2.5	How data can be accessed (access type, data use agreement, data access committee/group approval, IRB LOD, etc.)	"Under the HIPAA Privacy Rule "A covered entity may use or disclose a limited data set that meets the requirements of paragraphs (e)(2) and (e)(3) of this section, if the covered entity enters into a data use agreement with the limited data set recipient, in accordance with paragraph (e)(4) of this section…A covered entity may use or disclose a limited data set under paragraph (e)(1) of this section only for the purposes of research, public health, or health care operations." [1]	HIPAA Privacy Rule specifies that N3C data partners (health care providers) must enter into a data use agreement with limited EHR dataset receipients	[1] https://www.hhs.gov/sites/default/files/ocr/privacy/hipaa/administrative/combined/hipaa-simplification-201303.pdf (Accessed 11/9/23)
1.2.6	How data can be used (data use limitations)	The 1996 Health Insurance Portability and Accountability Act (HIPAA) allows medical and health care institutions to release data for research without obtaining an individual's authorization if direct identifying information is removed and appropriate oversight and agreements are in place. [1] Under the HIPAA privacy regulations for a Limited Data Set, de-identified health information may be used and disclosed for research purposes. [2]	HIPAA privacy regulations specify that de-identified data can be used for general research purposes.	[1] https://ncats.nih.gov/n3c/about (Accessed 4/21/23) [2] https://ncats.nih.gov/n3c/about/program-faq#use-the-data (Accessed 4/21/23)
1.2.7	Other (specify)	Does not authorize/specify	Does not authorize/specify	
2	**Data Linkage**			
2.1	**Authorizations and Applicable Regulations/Policies**			
2.1.1	**Authorizations**		1. Linkage Honest Broker Agreement 2. External Dataset Committee in the Tools and Resource subgroup and NCATS	
2.1.1.1	Assent	N/A	N/A	
2.1.1.2	Consent	N/A	N/A	
2.1.1.3	IRB/equivalent Privacy Board determination	Information not available/found	Information not available/found	
2.1.1.4	Local/state/federal laws	N/A	N/A	
2.1.1.5	Institutional Certification	N/A	N/A	

Dataset 8 - N3C				
		Raw Language	**Interpretation**	**Source**
2.1.1.6	Data originator agreement	All organizations contributing data to the N3C Data Enclave must have an approved Data Transfer Agreement (DTA). In addition to the DTA, these organizations have the option of signing the Linkage Honest Broker Agreement (LHBA) to participate in the PPRL pilot. The LHBA is an agreement between the organization, NCATS and The Regenstrief Institute, which serves as the linkage honest broker. A linkage honest broker in the PPRL's infrastructure is a party that holds de-identified tokens and operates a service that matches tokens generated across disparate data sets to formulate a single Match ID for a specific use case. The data remains under the complete control of the organizations that provide data to N3C and is never accessible by or under the control of the linkage honest broker. PPRL enables three functions within N3C: Deduplication of patient records, linkage of a patient's records from different sources and cohort discovery. Deduplication is a requirement for any organization that participates in the LHBA because of its importance to the data quality of the N3C Data Enclave and its scientific mission. Organizations participating in the LHBA have the option of participating in linking multiple data sets and cohort discovery as well. [1] Data partners who choose to participate in PPRL with external datasets must agree ... to link to the given external dataset PPRL Site Permissions Dashboard [2]	Two data originator agreements authorize data linkage: 1. The Linkage Honest Broker Agreement (LHBA) authorizes data linkage 2. Participating PPRL sites authorize data linkage for particular external datasets	[1] https://ncats.nih.gov/n3c/about/data-overview#privacy-preserving-record-linkage (Accessed 4/21/23) [2] Predecessor report (https://www.nichd.nih.gov/sites/default/files/inline-files/NICHD_ODSS_PPRL_for_Pediatric_COVID-19_Studies_Public_Final_Report_508.pdf)
2.1.1.7	Repository agreements/policies	Users that need to utilize an external dataset will first request the external dataset for consideration to be imported. Once the request is received, the External Dataset Committee in the Tools and Resource subgroup and NCATS will review external database requests to include in the N3C Dataset. This review process will be done in three phases: Initial Review Process Overall Dataset Review Process External Dataset Importation [1]	The External Dataset Committee in the Tools and Resource subgroup and NCATS approval authorizes data linkage	[1] https://docs.google.com/document/d/1QJi_sNi0wnZFV3ghTBubI7d3kFLtdkVfQwsLQJOFil8/edit (Accessed: 11/17/23)
2.1.1.8	Other (specify)	Information not available/found	Information not available/found	
2.1.2	**Applicable Regulations/Policies**			
2.1.2.1	Local regulations/policies	Information not available/found	Information not available/found	
2.1.2.2	Tribal regulations/policies	Information not available/found	Information not available/found	
2.1.2.3	State regulations/policies	Information not available/found	Information not available/found	
2.1.2.4	Federal regulations/policies	Information not available/found	Information not available/found	
2.1.2.5	International regulations/policies	Information not available/found	Information not available/found	
2.1.2.6	Contractual obligations	Information not available/found	Information not available/found	

Dataset 8 - N3C				
		Raw Language	**Interpretation**	**Source**
2.1.2.7	Repository policies	All External Datasets under consideration are given a classification code, which defines a dataset's risk for re-identification. The current classes are defined below: Class 0: Linkage is not available at this time, but defined in anticipation of the implementation of a Privacy-Preserving Record Linkage (PPRL e.g., hashed identifiers) managed by a third-party honest broker. ○ The assumption here is that Enclave data for a given patient has an internally generated ID, the external team has their own ID, and the honest broker has both (but nothing else), and only shares a hash of the other ID with each team. ○ Examples here include connecting Enclave data with separate repositories (at the individual patient level) of sequence data, imaging, etc. Class 1: Linkages leading to immediate re-identification of patients. ○ Linked datasets directly containing data that increases the risk or the reality of re-identification. A simple example here is a table of HIPAA-sensitive variables matching against Limited Dataset records. ○ Data fitting this class will not be permitted in the N3C Data Enclave, but we define it for sake of completeness. Class 2: Linkages leading to high-confidence heuristic re-identification of patients. ○ Example: a linkage of zip code and age (plus or minus an approximate data of infection) potentially leading to high-confidence heuristic reidentification of nursing home residents. Class 3: Linkages leading to data sufficiently aggregated to reasonably mitigate the risk of re-identification. ○ Example: a patient with a specific comorbidity known to be associated with a given genetic defect, knowing the pool of persons with that genetic defect fails to provide sufficient discrimination (assuming that that pool is of a sufficient cardinality.) Note that any additional linkage information (e.g., a person's height) might provide additional discriminating power in a sparse information space to allow re-identification. Class 4: Linkages or use of data not involving individual persons. ○ Examples: a dataset mapping zip code to the climate zone or a dataset mapping drugs to known side effects. Data sources that would extend specific value sets, such as ClinVar data for specific genomic mutations are another example. [1]	N3C policy designation of external datasets	[1] NCATS Response to N3C Questions
2.2	**Governance for data linkage, sharing, access, and use based on data linkage authorization or applicable regulations/policies (i.e., the origin of the governance)**			

Dataset 8 - N3C				
		Raw Language	**Interpretation**	**Source**
2.2.1	Whether the data can be linked	All organizations contributing data to the N3C Data Enclave must have an approved Data Transfer Agreement (DTA). In addition to the DTA, these organizations have the option of signing the Linkage Honest Broker Agreement (LHBA) to participate in the PPRL pilot. The LHBA is an agreement between the organization, NCATS and The Regenstrief Institute, which serves as the linkage honest broker. A linkage honest broker in the PPRL's infrastructure is a party that holds de-identified tokens and operates a service that matches tokens generated across disparate data sets to formulate a single Match ID for a specific use case. The data remains under the complete control of the organizations that provide data to N3C and is never accessible by or under the control of the linkage honest broker. PPRL enables three functions within N3C: Deduplication of patient records, linkage of a patient's records from different sources and cohort discovery. Deduplication is a requirement for any organization that participates in the LHBA because of its importance to the data quality of the N3C Data Enclave and its scientific mission. Organizations participating in the LHBA have the option of participating in linking multiple data sets and cohort discovery as well. [1] Data partners who choose to participate in PPRL with external datasets must agree ... to link to the given external dataset PPRL Site Permissions Dashboard [2] Users that need to utilize an external dataset will first request the external dataset for consideration to be imported. Once the request is received, the External Dataset Committee in the Tools and Resource subgroup and NCATS will review external database requests to include in the N3C Dataset. This review process will be done in three phases: Initial Review Process Overall Dataset Review Process External Dataset Importation [3]	1. LHBA specifies that the data can be linked 2. Participating PPRL sites specify that data can be linked with particular external datasets 3. The External Dataset Committee in the Tools and Resource subgroup and NCATS approval specifies that data can be linked	[1] NCATS Response to N3C Questions [2] Predecessor report (https://www.nichd.nih.gov/sites/default/files/inline-files/NICHD_ODSS_PPRL_for_Pediatric_COVID-19_Studies_Public_Final_Report_508.pdf) [3] https://docs.google.com/document/d/1QJi_sNi0wnZFV3ghTBubI7d3kFLtdkVfQwsLQJOFil8/edit

Dataset 8 - N3C				
		Raw Language	**Interpretation**	**Source**
2.2.2	With what other data can it be linked or can it not be linked (scope of linkage)	**Linking Multiple Datasets:** Though there are many types of datasets and ways to link to them, the Linkage Honest Broker Agreement applies only to datasets that are within N3C Data Enclave and requires linkage using the hash/PPRL.N3C has developed an external dataset classification system (See description below Multi-Dataset Linkage Classification) the LHBA only applies to datasets classified as class “0”and Class “2”. Linkages to external datasets that do not require the hash or PPRL are not covered by this agreement. The difference between Class 0 datasets and Class 2 datasets is Class 0 datasets originate from different enclaves and allows for a temporary extension of the N3C Data Enclave to accommodate this requirement. If additional computational resources are required for large datasets,the N3C Data Enclave will utilize NCATS High PC Performance Computing (HPC) services for data processing. **Multi-Dataset Linkage Classification Summary** ● Class 0: Linkages using cryptographic hash codes (tokens) managed by a third-party linkage honest broker to connect multiple Enclaves. ● Class 1: Linkages leading to immediate re-identification of patients and is not permitted with the N3C ● Class 2: Linkages using cryptographic hash codes (tokens)within a single enclave leading to higher confidence re-identification of patients. ● Class 3: Linkages leading to data sufficiently aggregated to reasonably mitigate the risk of re-identification ● Class 4: Linkages or use of data not involving individual persons. [1] All External Datasets under consideration are given a classification code, which defines a dataset’s risk for re-identification. The current classes are defined below: Class 0: Linkage is not available at this time, but defined in anticipation of the implementation of a Privacy-Preserving Record Linkage (PPRL e.g., hashed	1. N3C policy designation of external datasets for linking specifies that: a. External datasets must be classified as Class 0, 2, 3, or 4 to be considered for N3C linkage. A dataset which is categorized as class 2 can be imported but will require hashing b. Class 1 linkages are not permitted 2. Participating PPRL sites specify linkages with external datasets on a case-by-case basis 3. External Dataset Committee in the Tools and Resource subgroup and NCATS determines the scope of linkage by approving external datasets for import and linkage within N3C	[1] https://zenodo.org/record/5165257 (Accessed 4/19/23) [2] NCATS Response to N3C Questions [3] https://docs.google.com/document/d/1QJi_sNi0wnZFV3ghTBubI7d3kFLtdkVfQwsLQJOFil8/edit (Accessed: 11/17/23)

Dataset 8 - N3C				
		Raw Language	**Interpretation**	**Source**
		identifiers) managed by a third-party honest broker. ○ The assumption here is that Enclave data for a given patient has an internally generated ID, the external team has their own ID, and the honest broker has both (but nothing else), and only shares a hash of the other ID with each team. ○ Examples here include connecting Enclave data with separate repositories (at the individual patient level) of sequence data, imaging, etc. Class 1: Linkages leading to immediate re-identification of patients. ○ Linked datasets directly containing data that increases the risk or the reality of re-identification. A simple example here is a table of HIPAA-sensitive variables matching against Limited Dataset records. ○ Data fitting this class will not be permitted in the N3C Data Enclave, but we define it for sake of completeness. Class 2: Linkages leading to high-confidence heuristic re-identification of patients. ○ Example: a linkage of zip code and age (plus or minus an approximate data of infection) potentially leading to high-confidence heuristic re-identification of nursing home residents. Class 3: Linkages leading to data sufficiently aggregated to reasonably mitigate the risk of re-identification. ○ Example: a patient with a specific co-morbidity known to be associated with a given genetic defect knowing the pool of persons with that genetic defect fails to provide sufficient discrimination (assuming that that pool is of a sufficient cardinality.) Note that any additional linkage information (e.g., a person's height provide additional discriminating power in a sparse information space to allow re-identification. Class 4: Linkages or use of data not involving individual persons. ○ Examples: a dataset mapping zip code to the climate zone or a dataset mapping drugs to known side effects. Data sources that would extend specific value sets, such as ClinVar data for specific genomic mutations are another example. [2] Users that need to utilize an external dataset will first request the external dataset for consideration to be imported. Once the request is received, the External Dataset Committee in the Tools and Resource subgroup and NCATS will review external database requests to include in the N3C Dataset. This review process will be done in three phases: Initial Review Process Overall Dataset Review Process External Dataset Importation [3]		
2.2.3	Whether data can be shared	Does not authorize/specify	Does not authorize/specify	
2.2.4	How data can be shared (de-identification status, disclosure review)	Does not authorize/specify	Does not authorize/specify	

Dataset 8 - N3C				
		Raw Language	**Interpretation**	**Source**
2.2.5	How data can be accessed (access type, data use agreement, data access committee/group approval, IRB LOD, etc.)	Class 1 (not allowed). Linkages leading to immediate re-identification of patients. Access to various classes of data require sets of agreements to be in place for the data requesters. All datasets that use the PPRL (Class 0 and 2) are considered level 3 data and as such must work through their institutional policies and require a letter of determination when submitting an N3C Data Use Request (DUR). Class 2 dataset linkages require existing institutional N3C Data Use Agreement, Dual authentication and authorization, a signed institutional linkage honest broker agreement for multiple datasets, an approved data use request (DUR) by the federally staffed data access committee (DAC), and local institutions IRB letter of determination. IRBs must clearly have reviewed the DURs proposed protocol and the specific use of multiple datasets beyond N3C EHR-derived data. Class 2 dataset linkage are contained within the single N3C Data Enclave. An example of a class 2 multiple linkage datasets would be if N3C data is linked to Mortality data that was sent to N3C. For class 0 dataset linkages, that connect more than one enclave, an additional interconnect agreement will be in place. The interconnect agreement will be agreement between two trusted enclaves in order to instantiate what is referred to as a temporary virtual or ephemeral workbench. The workbench is ephemeral because it is short-lived for a specific task and then destroyed when an investigator's work is completed. Classes 3 and 4 require a DUR for the study approved by the DAC. [1]	N3C specifies that for data access, there must be: 1. For N3C Class 0 or Class 2 linkages: a. Existing institutional N3C Data Use Agreement b. Dual authentication and authorization c. Signed institutional linkage honest broker agreement for multiple datasets d. Approved data use request (DUR) by the federally staffed data access committee (DAC) e. Local institutions IRB letter of determination f. Interconnect agreement (for Class 0 only) 2. For N3C Class 3 and 4 linkages: a. Approved data use request (DUR) by the federally staffed data access committee (DAC)	[1] https://covid.cd2h.org/PPRL (Accessed 4/28/23)
2.2.6	How data can be used (data use limitations)	Does not authorize/specify	Does not authorize/specify	
2.2.7	Other (specify)	Does not authorize/specify	Does not authorize/specify	
3	**Data Sharing**			
3.1	**Authorizations and Applicable Regulations/Policies**			
3.1.1	**Authorizations**		1. Institutional IRB or external central IRB approval 2. Waiver of consent 3. Data Transfer Agreement	
3.1.1.1	Assent	N/A	N/A	
3.1.1.2	Consent	N/A	N/A	
3.1.1.3	IRB/equivalent Privacy Board determination	The N3C does not contain direct identifying information, and additional measures have been put in place to protect patient privacy. As a result, NCATS received a waiver of consent from an NIH Institutional Review Board, conforming to the Federal Policy for the Protection of Human Subjects (“Common Rule”). [1] Transfer of data from participating institutions to the NCATS N3C platform covered under a cIRB, unless an institution chose to utilize their own IRB for data transfer to NCATS [2]	Two IRB determinations authorize data sharing: 1. Institutional IRB or external central IRB approval for transfer of data from participating institutions to the NCATS N3C Platform 2. NIH IRB waiver of consent for sharing through the NCATS N3C Platform	[1] https://ncats.nih.gov/n3c/about (Accessed 4/21/23) [2] NCATS Email Communication
3.1.1.4	Local/state/federal laws	Information not available/found	Information not available/found	

Dataset 8 - N3C				
		Raw Language	**Interpretation**	**Source**
3.1.1.5	Institutional Certification	Information not available/found	Information not available/found	
3.1.1.6	Data originator agreement	Participating partners and other collaborators provide data to the N3C after they execute a Data Transfer Agreement with NCATS. [1] “Recipient shall not use the Data except as authorized under this Agreement. The Data will be deposited in the NIH COVID-19 Data Warehouse and made accessible to users under a separate data use agreement to support the response to the COVID-19 pandemic.” [2] Institutions that wish to contribute data to the N3C must execute a Data Transfer Agreement (DTA) with NCATS. The DTA provides terms and conditions for data transfer and outlines the general terms of data use. [3] The NCATS Data Transfer Agreement: is under the stewardship of NCATS and is an agreement between NCATS and participating institutions that are willing to provide de-identified data to the N3C Data Enclave. [4]	Data Transfer Agreement (DTA) executed with NCATS authorizes data transfer to NCATS N3C Platform	[1] https://ncats.nih.gov/n3c/about (Accessed 4/21/23) [2] https://ncats.nih.gov/sites/default/files/NCATS_Data_Transfer_Agreement_508.pdf (Accessed 4/21/23) [3] https://ncats.nih.gov/n3c/resources/data-contribution (Accessed 4/21/23) [4] NCATS's Responses to N3C Questions
3.1.1.7	Repository agreements/policies	Information not available/found	Information not available/found	
3.1.1.8	Other	Information not available/found	Information not available/found	
3.1.2	**Applicable Regulations/Policies**			
3.1.2.1	Local regulations/policies	NCATS asks medical institutions and health care organizations to contribute this information as a limited data set, pursuant to the requirements in the HIPAA Privacy Rule. A limited data set is defined as protected health information that excludes certain direct identifiers of an individual or of relatives, employers or household members of the individual — but may include city, state, ZIP code and elements of dates. A limited data set can be disclosed only for purposes of research, public health or health care operations. Three levels of data are available for analysis: Limited Data Set (LDS): Consists of patient data that retain the following protected health information — dates of service patient ZIP code De-identified Data Set: Consists of patient data from the LDS with the following changes — Dates of service are algorithmically shifted to protect patient privacy. Patient ZIP codes are truncated to the first three digits or removed entirely if the ZIP code represents fewer than 20,000 individuals or represents Tribal lands. Synthetic Data Set: Consists of data that are computationally derived from the LDS and that resemble patient information statistically but are not actual patient data. [1]	N3C policies	[1] https://ncats.nih.gov/n3c/about/data-overview (Accessed 8/9/23)

Dataset 8 - N3C				
		Raw Language	**Interpretation**	**Source**
3.1.2.2	Tribal regulations/policies	In alignment with the National Institutes of Health Guidance on the Implementation of the HHS Tribal Consultation Policy, NCATS sought input from Tribal Nations on whether and how to provide AI/AN data within the N3C. Now that the Tribal Consultation has been conducted, AI/AN data is available for research in a manner that reflects the input that NIH received through the consultation process. The Tribal Consultation report (PDF-1.3MB) details how this data is now being made available. [1] Based on the feedback from Tribal Consultation, NCATS will take the following steps to make AI/AN data available for research through its standard Data Use Request (DUR) process: 1. AI/AN data will be moved back to a standalone category. With this change, AI/AN data will be available in any N3C analysis that provides race and ethnicity distribution. 2. ZIP codes that overlap with Tribal communities will be available for research in the following manner: ZIP codes for all geographic units containing 20,000 or fewer people are removed entirely. This is standard practice for all geographic units and will be applied the same way when AI/AN data are restored to a separate category. For example, if a ZIP code of “01234” represents a community of 20,000 or fewer individuals, the user will see a ZIP code of “00000.” Currently, specific ZIP codes representing rural populations predominantly with AI/AN-identifying individuals are hidden. These will now be visible in both the limited data set and de-identified data. This means that AI/AN data will be managed as others are managed, with the exception that the full five-digit ZIP codes are never shown. For example, if a ZIP code of “01234” represents a predominantly AI/AN community, the user will see only a partial ZIP code of “012.” 3. The N3C Data User Code of Conduct will be modified so that data users will be asked to attest that they understand the N3C contains no Tribal affiliation data and that use of AI/AN data and ZIP code information to make assumptions about Tribal affiliation is not valid or appropriate. This statement will be included as a reminder when a DUR is received, a Data Use Agreement is executed, and data are accessed in the N3C platform and during publication processes. 4. The N3C will continue to engage NIH’s Tribal Health Research Office and Tribal Nations as issues for discussion arise. [2] It is important to note that the N3C does not have EHR data from the Indian Health Service, nor does it contain information about Tribal affiliation. [3]	1. NIH Guidance on the Implementation of the HHS Tribal Consultation Policy 2. Tribal Consultation Report	[1] https://ncats.nih.gov/n3c/about/program-faq# (Accessed 4/21/23) [2] https://ncats.nih.gov/n3c/about/tribal-consultation (Accessed 4/21/23) [3] https://dpcpsi.nih.gov/sites/default/files/N3C-Consultation-Report-508.pdf (Accessed 4/21/23)
3.1.2.3	State regulations/policies	Information not available/found	Information not available/found	

Dataset 8 - N3C				
		Raw Language	**Interpretation**	**Source**
3.1.2.4	Federal regulations/policies	In alignment with the National Institutes of Health Guidance on the Implementation of the HHS Tribal Consultation Policy, NCATS sought input from Tribal Nations on whether and how to provide AI/AN data within the N3C. Now that the Tribal Consultation has been conducted, AI/AN data is available for research in a manner that reflects the input that NIH received through the consultation process. The Tribal Consultation report (PDF-1.3MB) details how this data is now being made available. [1]	HHS Tribal Consultation Policy	[1] https://ncats.nih.gov/n3c/about/program-faq# (Accessed 4/21/23)
3.1.2.5	International regulations/policies	Information not available/found	Information not available/found	
3.1.2.6	Contractual obligations	Information not available/found	Information not available/found	
3.1.2.7	Repository policies	Information not available/found	Information not available/found	
3.2	**Governance for data linkage, sharing, access, and use based on data sharing authorization or applicable regulations/policies (i.e., the origin of the governance)**			
3.2.1	Whether the data can be linked	Does not authorize/specify	Does not authorize/specify	
3.2.2	With what other data can it be linked or can it not be linked (scope of linkage)	...the N3C platform is subject to the common rule. [1] The N3C does not contain direct identifying information, and additional measures have been put in place to protect patient privacy. As a result, NCATS received a waiver of consent from an NIH Institutional Review Board, conforming to the Federal Policy for the Protection of Human Subjects (“Common Rule”). [2]	Common Rule	[1] NCATS Email Communication [2] https://ncats.nih.gov/n3c/about (Accessed 4/21/23)
3.2.3	Whether data can be shared	The N3C does not contain direct identifying information, and additional measures have been put in place to protect patient privacy. As a result, NCATS received a waiver of consent from an NIH Institutional Review Board, conforming to the Federal Policy for the Protection of Human Subjects (“Common Rule”). [1] Participating partners and other collaborators provide data to the N3C after they execute a Data Transfer Agreement with NCATS. [1] “Recipient shall not use the Data except as authorized under this Agreement. The Data will be deposited in the NIH COVID-19 Data Warehouse and made accessible to users under a separate data use agreement to support the response to the COVID-19 pandemic.” [2] In alignment with the National Institutes of Health Guidance on the Implementation of the HHS Tribal Consultation Policy, NCATS sought input from Tribal Nations on whether and how to provide AI/AN data within the N3C. Now that the Tribal Consultation has been conducted, AI/AN data is available for research in a manner that reflects the input that NIH received through the consultation process. The Tribal Consultation report (PDF-1.3MB) details how this data is now being made available. [3] Based on the feedback from Tribal Consultation, NCATS will take the following steps to make AI/AN data available for research through its standard Data Use Request (DUR) process:	1. NIH IRB waiver of consent specifies that data can be shared 2. Data Transfer Agreement specifies that data can be shared 3. NIH Guidance on the Implementation of the HHS Tribal Consultation Policy, the HHS Tribal Consultation Policy, and the Tribal Consultation Report specify that AI/AN data can be shared	[1] https://ncats.nih.gov/n3c/about (Accessed 4/21/23) [2] https://ncats.nih.gov/sites/default/files/NCATS_Data_Transfer_Agreement_508.pdf (Accessed 4/21/23) [3] https://ncats.nih.gov/n3c/about/program-faq# (Accessed 4/21/23) [4] https://ncats.nih.gov/n3c/about/tribal-consultation (Accessed 4/21/23) [5] https://dpcpsi.nih.gov/sites/default/files/N3C-Consultation-Report-508.pdf (Accessed 4/21/23)

Dataset 8 - N3C				
		Raw Language	**Interpretation**	**Source**
		1. AI/AN data will be moved back to a standalone category. With this change, AI/AN data will be available in any N3C analysis that provides race and ethnicity distribution. 2. ZIP codes that overlap with Tribal communities will be available for research in the following manner: ZIP codes for all geographic units containing 20,000 or fewer people are removed entirely. This is standard practice for all geographic units and will be applied the same way when AI/AN data are restored to a separate category. For example, if a ZIP code of “01234” represents a community of 20,000 or fewer individuals, the user will see a ZIP code of “00000.” Currently, specific ZIP codes representing rural populations predominantly with AI/AN-identifying individuals are hidden. These will now be visible in both the limited data set and de-identified data. This means that AI/AN data will be managed as others are managed, with the exception that the full five-digit ZIP codes are never shown. For example, if a ZIP code of “01234” represents a predominantly AI/AN community, the user will see only a partial ZIP code of “012.” 3. The N3C Data User Code of Conduct will be modified so that data users will be asked to attest that they understand the N3C contains no Tribal affiliation data and that use of AI/AN data and ZIP code information to make assumptions about Tribal affiliation is not valid or appropriate. This statement will be included as a reminder when a DUR is received, a Data Use Agreement is executed, and data are accessed in the N3C platform and during publication processes. 4. The N3C will continue to engage NIH’s Tribal Health Research Office and Tribal Nations as issues for discussion arise. [4] It is important to note that the N3C does not have EHR data from the Indian Health Service, nor does it contain information about Tribal affiliation. [5]		

Dataset 8 - N3C				
		Raw Language	**Interpretation**	**Source**
3.2.4	How data can be shared (de-identification status, disclosure review)	Three levels of data are available for analysis: Limited Data Set (LDS): Consists of patient data that retain the following protected health information — dates of service patient ZIP code De-identified Data Set: Consists of patient data from the LDS with the following changes — Dates of service are algorithmically shifted to protect patient privacy. Patient ZIP codes are truncated to the first three digits or removed entirely if the ZIP code represents fewer than 20,000 individuals or represents Tribal lands. Synthetic Data Set: Consists of data that are computationally derived from the LDS and that resemble patient information statistically but are not actual patient data. [1] In alignment with the National Institutes of Health Guidance on the Implementation of the HHS Tribal Consultation Policy, NCATS sought input from Tribal Nations on whether and how to provide AI/AN data within the N3C. Now that the Tribal Consultation has been conducted, AI/AN data is available for research in a manner that reflects the input that NIH received through the consultation process. The Tribal Consultation report (PDF-1.3MB) details how this data is now being made available. [2] Based on the feedback from Tribal Consultation, NCATS will take the following steps to make AI/AN data available for research through its standard Data Use Request (DUR) process: 1. AI/AN data will be moved back to a standalone category. With this change, AI/AN data will be available in any N3C analysis that provides race and ethnicity distribution. 2. ZIP codes that overlap with Tribal communities will be available for research in the following manner: ZIP codes for all geographic units containing 20,000 or fewer people are removed entirely. This is standard practice for all geographic units and will be applied the same way when AI/AN data are restored to a separate category. For example, if a ZIP code of “01234” represents a community of 20,000 or fewer individuals, the user will see a ZIP code of “00000.” Currently, specific ZIP codes representing rural populations predominantly with AI/AN-identifying individuals are hidden. These will now be visible in both the limited data set and de-identified data. This means that AI/AN data will be managed as others are managed, with the exception that the full five-digit ZIP codes are never shown. For example, if a ZIP code of “01234” represents a predominantly AI/AN community, the user will see only a partial ZIP code of “012.”	1. N3C policy specifies limited datasets (LDS), de-identified, and synthetic datasets are shared. 2. NIH Guidance on the Implementation of the HHS Tribal Consultation Policy, the HHS Tribal Consultation Policy, and the Tribal Consultation Report specify that (a). AI/AN data will be a standalone category. With this change, AI/AN data will be available in any N3C analysis that provides race and ethnicity distribution. (b). ZIP codes must be removed entirely for all geographic units containing 20,000 or fewer people, and full five-digit ZIP codes of predominantly AI/AN community will never be shown.	[1] https://ncats.nih.gov/n3c/about/data-overview (Accessed 4/21/23) [2] https://ncats.nih.gov/n3c/about/program-faq# (Accessed 4/21/23) [3] https://ncats.nih.gov/n3c/about/tribal-consultation (Accessed 4/21/23) [4] https://dpcpsi.nih.gov/sites/default/files/N3C-Consultation-Report-508.pdf (Accessed 4/21/23)

Dataset 8 - N3C				
		Raw Language	**Interpretation**	**Source**
		3. The N3C Data User Code of Conduct will be modified so that data users will be asked to attest that they understand the N3C contains no Tribal affiliation data and that use of AI/AN data and ZIP code information to make assumptions about Tribal affiliation is not valid or appropriate. This statement will be included as a reminder when a DUR is received, a Data Use Agreement is executed, and data are accessed in the N3C platform and during publication processes. 4. The N3C will continue to engage NIH's Tribal Health Research Office and Tribal Nations as issues for discussion arise. [3] It is important to note that the N3C does not have EHR data from the Indian Health Service, nor does it contain information about Tribal affiliation. [4]		
3.2.5	How data can be accessed (access type, data use agreement, data access committee/group approval, IRB LOD, etc.)	Recipient shall not use the Data except as authorized under this Agreement. The Data will be deposited in the NIH COVID-19 Data Warehouse and made accessible to usersNCATS Data Transfer Agreement FINAL 4-7-2021 2 under a separate data use agreement to support the response to the COVID-19 pandemic. Users who access the Data will: • only analyze the Data within the NIH COVID-19 Data Warehouse platform. • not be able to download or remove the Data from the NIH COVID-19 Warehouse in any form. • share the results of analyses within the platform to the extent possible • make no effort to contact or identify individuals who are or may be the sources or subjects of the Data • agree to acknowledge the NIH COVID-19 Data Warehouse in all publications and oral disclosures that rely on the Data • be used only for research purposes and public health activities related to the COVID-19 pandemic [1]	The DTA specifies that users who access the data will access the data within the NCATS N3C Plaform	[1] https://ncats.nih.gov/sites/default/files/NCATS_Data_Transfer_Agreement_508.pdf (Accessed 4/21/23)

Dataset 8 - N3C				
		Raw Language	**Interpretation**	**Source**
3.2.6	How data can be used (data use limitations)	3. The N3C Data User Code of Conduct will be modified so that data users will be asked to attest that they understand the N3C contains no Tribal affiliation data and that use of AI/AN data and ZIP code information to make assumptions about Tribal affiliation is not valid or appropriate. This statement will be included as a reminder when a DUR is received, a Data Use Agreement is executed, and data are accessed in the N3C platform and during publication processes. [1] Recipient shall not use the Data except as authorized under this Agreement. The Data will be deposited in the NIH COVID-19 Data Warehouse and made accessible to users NCATS Data Transfer Agreement FINAL 4-7-2021 2 under a separate data use agreement to support the response to the COVID-19 pandemic. Users who access the Data will: • only analyze the Data within the NIH COVID-19 Data Warehouse platform. • not be able to download or remove the Data from the NIH COVID-19 Warehouse in any form. • share the results of analyses within the platform to the extent possible • make no effort to contact or identify individuals who are or may be the sources or subjects of the Data • agree to acknowledge the NIH COVID-19 Data Warehouse in all publications and oral disclosures that rely on the Data • be used only for research purposes and public health activities related to the COVID-19 pandemic [2]	1. The Tribal Consultation Report specifies that data users will be asked to attest that they understand the N3C contains no Tribal affiliation data and that use of AI/AN data and ZIP code information to make assumptions about Tribal affiliation is not valid or appropriate 2. The DTA specifies that data must be used only for research purposes and public health activities related to the COVID-19 pandemic	[1] https://ncats.nih.gov/n3c/about/tribal-consultation (Accessed 4/21/23) [2] https://ncats.nih.gov/sites/default/files/NCATS_Data_Transfer_Agreement_508.pdf (Accessed 4/21/23)
3.2.7	Other (specify)	Does not authorize/specify	Does not authorize/specify	
4	**Data Access**			
4.1	**Authorizations and Applicable Regulations/Policies**			
4.1.1	**Authorizations**		1. Letter of Determination 2. Data Use Agreement 3. Data Use Request	
4.1.1.1	Assent	N/A	N/A	
4.1.1.2	Consent	N/A	N/A	
4.1.1.3	IRB/equivalent Privacy Board determination	Local IRB letter of determination is required for accessing data (see Appendix Table 2) [1] If requesting access to the Limited Data Set, researchers will need to provide a copy of their institution's Human Research Protection Program IRB determination letter. [2]	Letter of Determination from user's Insitutional IRB authorizes data access	[1] Predecessor report (https://www.nichd.nih.gov/sites/default/files/inline-files/NICHD_ODSS_PPRL_for_Pediatric_COVID-19_Studies_Public_Final_Report_508.pdf) [2] https://ncats.nih.gov/n3c/about/applying-for-access (Accessed 4/21/23)
4.1.1.4	Local/state/federal laws	Information not available/found	Information not available/found	
4.1.1.5	Institutional Certification	N/A	N/A	
4.1.1.6	Data originator agreement	Information not available/found	Information not available/found	

Dataset 8 - N3C				
		Raw Language	**Interpretation**	**Source**
4.1.1.7	Repository agreements/policies	Before researchers can request access to the data, their institutions must execute a Data Use Agreement (DUA) with NCATS. Once a DUA is in place, researchers can submit a Data Use Request (DUR) through the N3C Data Enclave. In the DUR, researchers will need to include, among other information, the project research title, names of project personnel, a non-confidential research statement, the project proposal and the requested data access level. Additional DUR requirements include reviewing and agreeing to comply with the N3C Data User Code of Conduct. The N3C Data Access Committee reviews and approves DURs. [1] Data Access Committee (DAC), referred to as the NCATS N3C DAC, is composed of representatives from the NIH or other Federal Agencies and is responsible for reviewing Data Use Requests (DUR) and verifying that conditions for accessing Data are met. [2] Data Use Request (DUR) a document that User(s) must complete and submit to the DAC for review prior to accessing the NCATS N3C Data Enclave. For each individual project, the DUR, outlined in Appendix A, will contain overarching use objectives and designs, scientific goals, any testable hypotheses, and an outline of plans for using the Data. [2]	Two repository agreements authorize data access: 1. Data Use Agreement 2. Data Use Request	[1] https://ncats.nih.gov/n3c/about/applying-for-access (Accessed 4/21/23) [2] https://ncats.nih.gov/sites/default/files/NCATS_N3C_Data_Use_Agreement.pdf (Accessed 4/21/23)
4.1.1.8	Other (specify)	Information not available/found	Information not available/found	
4.1.2	**Applicable Regulations/Policies**			
4.1.2.1	Local regulations/policies	Information not available/found	Information not available/found	
4.1.2.2	Tribal regulations/policies	Information not available/found	Information not available/found	
4.1.2.3	State regulations/policies	Information not available/found	Information not available/found	
4.1.2.4	Federal regulations/policies	Data residing in NCATS N3C Data Enclave is protected by a Certificate of Confidentiality pursuant to section 301(d) of the Public Health Service Act. The Certificate of Confidentiality prohibits disclosure, including in any federal, state, or local criminal, civil, administrative, legislative, or other proceeding, of identifiable, sensitive information collected or used during research unless disclosure is made pursuant to a statutory exception. The Certificate of Confidentiality also protects all copies of Data from the NCATS N3C Data Enclave in perpetuity as explained in the NIH Policy for Issuing Certificates of Confidentiality [1]	Certificate of Confidentiality	[1] https://grants.nih.gov/grants/guide/notice-files/NOT-OD-17109.html (Accessed 10/23/23)
4.1.2.5	International regulations/policies	Information not available/found	Information not available/found	
4.1.2.6	Contractual obligations	NCATS uses Palantir for its software and expertise in the platform's execution. Palantir is hosted by NCATS within this instance, and no data can leave this enclave or be accessed by the company for its use. All contractors with access to the NCATS GovCloud instance to implement and maintain the NCATS N3C Data Enclave are subject to all relevant NIH-specified clearances, non-disclosure agreements, training, rules and restrictions. Contractors are not allowed to independently access NCATS N3C Data Enclave data, remove it from the enclave or use it for commercial purposes. [1]	Obligations from contract between NCATS and Palantir	[1] NCATS Email Communication

Dataset 8 - N3C				
		Raw Language	**Interpretation**	**Source**
4.1.2.7	Repository policies	Before researchers can request access to the data, their institutions must execute a Data Use Agreement (DUA) with NCATS. Once a DUA is in place, researchers can submit a Data Use Request (DUR) through the N3C Data Enclave. In the DUR, researchers will need to include, among other information, the project research title, names of project personnel, a non-confidential research statement, the project proposal and the requested data access level. Additional DUR requirements include reviewing and agreeing to comply with the N3C Data User Code of Conduct. The N3C Data Access Committee reviews and approves DURs. [1] Data Access Committee (DAC), referred to as the NCATS N3C DAC, is composed of representatives from the NIH or other Federal Agencies and is responsible for reviewing Data Use Requests (DUR) and verifying that conditions for accessing Data are met. [2] Data Use Request (DUR) a document that User(s) must complete and submit to the DAC for review prior to accessing the NCATS N3C Data Enclave. For each individual project, the DUR, outlined in Appendix A, will contain overarching use objectives and designs, scientific goals, any testable hypotheses, and an outline of plans for using the Data. [2] Investigators using the rich data compiled in the N3C Data Enclave are expected to generate quantitative results in the forms of tables, figures, parameter estimates, and aggregated statistics. These results may have broad impact and will be shared, typically in the form of manuscripts, reports, and visualizations for websites and seminars. Additionally, many artifacts such as workflows, value sets, and phenotyping algorithms may be developed in the Enclave and made useful to others by being exported for presentation on public websites such as GitHub, or incorporated as supplementary materials to publications. The purpose of this policy is to establish a process to permit the download of quantitative artifacts used in analyses that are necessary for sharing results, while at the same time ensuring protection and security of the data in the Enclave and compliance with the NCATS Data Transfer and Data Use Agreements. [3]	N3C policies, including N3C Results Download Policy	[1] https://ncats.nih.gov/n3c/about/applying-for-access (Accessed 4/21/23) [2] https://ncats.nih.gov/sites/default/files/NCATS_N3C_Data_Use_Agreement.pdf (Accessed 4/21/23) [3] https://zenodo.org/record/7942069 (Accessed 10/23/23)
4.2	**Governance for data linkage, sharing, access, and use based on data access authorization or applicable regulations/policies (i.e., the origin of the governance)**			
4.2.1	Whether the data can be linked	Does not authorize/specify	Does not authorize/specify	
4.2.2	With what other data can it be linked or can it not be linked (scope of linkage)	Does not authorize/specify	Does not authorize/specify	
4.2.3	Whether data can be shared	Does not authorize/specify	Does not authorize/specify	
4.2.4	How data can be shared (de-identification status, disclosure review)	Does not authorize/specify	Does not authorize/specify	
4.2.5	How data can be accessed (access type, data use agreement, data access committee/group approval, IRB LOD, etc.)	An N3C Data Use Agreement (DUA) is executed by NCATS and a research institution (or directly with a researcher in the case of a citizen scientist who is not affiliated with an institution). To submit a DUA to NCATS, institutions can download this form and email it to NCATSPartnerships@mail.nih.gov (link sends e-mail). Note: DUAs can only be signed by Authorized Institutional Officials who have the	N3C policies specify that to access the Limited Dataset, the user: 1) Must complete N3C registration and create a N3C Data Enclave account 2) Must execute an Institutional Data Use Agreement	[1] https://ncats.nih.gov/n3c/resources/data-access (Accessed 4/21/23) [2] https://ncats.nih.gov/n3c/about/applying-for-access (Accessed 4/21/23)

Dataset 8 - N3C				
		Raw Language	**Interpretation**	**Source**
		authority to bind all users at their institution to the terms of the DUA. With the exception of citizen scientists not associated with institutions, individual researchers cannot sign their own DUAs with NCATS. It is strongly recommended that researchers ensure that their home institution has executed a DUA with NCATS before they begin the process of applying for data access. See the list of institutions(link is external) with active DUAs. DUAs will remain in effect for five years from the DUA Effective Date. They will automatically expire at the end of this period unless terminated or renewed. [1] In order to access N3C data, researchers must submit an online Data Use Request (DUR) through the NCATS N3C Data Enclave for each project they want to start or join. Information researchers need to provide in the DUR includes: The project title Names of project personnel A non-confidential research statement The project proposal The requested data access level The N3C Data Access Committee (DAC) reviews and approves DURs. Once the DAC approves a DUR, access to the N3C Data Enclave workspace will be effective for one year starting from the date access is granted. A DUA must be in place for the entire term of a DUR. DURs will be renewable. When users renew their DURs, they will need to attest at that time that their training for access to the N3C Data Enclave is up to date. N3C Data User Code of Conduct The N3C Data User Code of Conduct states core activities of data users and prohibitions on certain activities involving data or other resources accessible through the N3C Data Enclave. Researchers must agree to the N3C Data User Code of Conduct when they apply to access and use N3C data. NIH Information Security and Information Management Training The N3C Data Enclave is hosted by NCATS, and all researchers must complete the "Information Security, Counterintelligence, Privacy Awareness, Records Management Refresher, Emergency Preparedness Refresher" course, which can be accessed at NIH's information security training website, before submitting a DUR. It will take approximately 60-90 minutes to complete the entire course and users should save evidence of completion for their records (a screenshot or copy of the certificate of completion). [1] If requesting access to the Limited Data Set, researchers will need to provide a copy of their institution's Human Research Protection Program IRB determination letter. [2]	3) Must submit Data Use Request (DUR) for approval by N3C Data Access Committee 4) Must complete NIH IT training, attest to the N3C Data User Code of Conduct, and complete Human Subjects Research Protection training at their home institution 5) Must provide IRB letter of determination for data access 6) Must access the data within the N3C Enclave 7) Must abide by the N3C Results Download Policy when downloading results from the N3C enclave	[3] NCATS Response to N3C Questions [4] Predecessor report (https://www.nichd.nih.gov/sites/default/files/inlinefiles/NICHD_ODSS_PPRL_for_Pediatric_COVID-19_Studies_Public_Final_Report_508.pdf) [5] NCATS Email Communication [6] https://zenodo.org/records/7942069 (Accessed 10/23/23)

Dataset 8 - N3C				
		Raw Language	**Interpretation**	**Source**
		Data are not allowed to leave the Enclave; analyses must be done within the platform and analysis results can be shared broadly. [3] Local IRB letter of determination is required for accessing data (see Appendix Table 2) [4] Researchers who request access to de-identified data or to the Limited Data Set must have completed their home institution's human subjects research training requirements. Researchers will be required to provide the date they completed training in their Data Use Request. [5] Investigators using the rich data compiled in the N3C Data Enclave are expected to generate quantitative results in the forms of tables, figures, parameter estimates, and aggregated statistics. These results may have broad impact and will be shared, typically in the form of manuscripts, reports, and visualizations for websites and seminars. Additionally, many artifacts such as workflows, value sets, and phenotyping algorithms may be developed in the Enclave and made useful to others by being exported for presentation on public websites such as GitHub, or incorporated as supplementary materials to publications. The purpose of this policy is to establish a process to permit the download of quantitative artifacts used in analyses that are necessary for sharing results, while at the same time ensuring protection and security of the data in the Enclave and compliance with the NCATS Data Transfer and Data Use Agreements. [6]		
4.2.6	How data can be used (data use limitations)	User(s) agree(s) to use the Data exclusively for the Research Project proposed, and/or comparative studies using data from individuals infected with pathogens such as SARS, MERS, and H1N1 to support comparative studies, described in the DUR. Any other use of the Data is prohibited. [1]	N3C DUA specifies that the data must be used exclusively for the Research Project proposed, and/or comparative studies using data from individuals infected with pathogens such as SARS, MERS, and H1N1 to support comparative studies.	[1] https://ncats.nih.gov/sites/default/files/NCATS_N3C_Data_Use_Agreement.pdf (Accessed 4/21/23)

Dataset 8 - N3C				
		Raw Language	**Interpretation**	**Source**
4.2.7	Other (specify)	NCATS uses Palantir for its software and expertise in the platform's execution. Palantir is hosted by NCATS within this instance, and no data can leave this enclave or be accessed by the company for its use. All contractors with access to the NCATS GovCloud instance to implement and maintain the NCATS N3C Data Enclave are subject to all relevant NIH-specified clearances, non-disclosure agreements, training, rules and restrictions. Contractors are not allowed to independently access NCATS N3C Data Enclave data, remove it from the enclave or use it for commercial purposes. [1] Data residing in NCATS N3C Data Enclave is protected by a Certificate of Confidentiality pursuant to section 301(d) of the Public Health Service Act. The Certificate of Confidentiality prohibits disclosure, including in any federal, state, or local criminal, civil, administrative, legislative, or other proceeding, of identifiable, sensitive information collected or used during research unless disclosure is made pursuant to a statutory exception. The Certificate of Confidentiality also protects all copies of Data from the NCATS N3C Data Enclave in perpetuity as explained in the NIH Policy for Issuing Certificates of Confidentiality [2]	1. Contract between NCATS and Palantir specifies that Palantir contractors with access to the NCATS GovCloud instance to implement and maintain the NCATS N3C Data Enclave are subject to all relevant NIH-specified clearances, non-disclosure agreements, training, rules and restrictions and are not allowed to independently access NCATS N3C Data Enclave data, remove it from the enclave, or use it for commercial purposes. 2. Certificate of Confidentiality protects N3C data from certain types of disclosures.	[1] NCATS Email Communication [2] https://grants.nih.gov/grants/guide/notice-files/NOT-OD-17-109.html (Accessed 10/23/23)
5	**Data Use**			
5.1	**Authorizations and Applicable Regulations/Policies**			
5.1.1	**Authorizations**		1. Letter of Determination 2. Data Use Agreement 3. Data Use Request	
5.1.1.1	Assent	N/A	N/A	
5.1.1.2	Consent	N/A	N/A	
5.1.1.3	IRB/equivalent Privacy Board determination	Local IRB letter of determination is required for accessing data (see Appendix Table 2) [1] If requesting access to the Limited Data Set, researchers will need to provide a copy of their institution's Human Research Protection Program IRB determination letter. [2]	Letter of Determination from user's institutional IRB authorizes data access	[1] Predecessor report (https://www.nichd.nih.gov/sites/default/files/inline-files/NICHD_ODSS_PPRL_for_Pediatric_COVID-19_Studies_Public_Final_Report_508.pdf) [2] https://ncats.nih.gov/n3c/about/applying-for-access (Accessed 4/21/23)
5.1.1.4	Local/state/federal laws	Information not available/found	Information not available/found	
5.1.1.5	Institutional Certification	Information not available/found	Information not available/found	
5.1.1.6	Data originator agreement	Information not available/found	Information not available/found	

Dataset 8 - N3C				
		Raw Language	**Interpretation**	**Source**
5.1.1.7	Repository agreements/policies	Before researchers can request access to the data, their institutions must execute a Data Use Agreement (DUA) with NCATS. Once a DUA is in place, researchers can submit a Data Use Request (DUR) through the N3C Data Enclave. In the DUR, researchers will need to include, among other information, the project research title, names of project personnel, a non-confidential research statement, the project proposal and the requested data access level. Additional DUR requirements include reviewing and agreeing to comply with the N3C Data User Code of Conduct. The N3C Data Access Committee reviews and approves DURs. [1] Data Access Committee (DAC), referred to as the NCATS N3C DAC, is composed of representatives from the NIH or other Federal Agencies and is responsible for reviewing Data Use Requests (DUR) and verifying that conditions for accessing Data are met. [2] Data Use Request (DUR) a document that User(s) must complete and submit to the DAC for review prior to accessing the NCATS N3C Data Enclave. For each individual project, the DUR, outlined in Appendix A, will contain overarching use objectives and designs, scientific goals, any testable hypotheses, and an outline of plans for using the Data. [2] User(s) agree(s) to use the Data exclusively for the Research Project proposed, and/or comparative studies using data from individuals infected with pathogens such as SARS, MERS, and H1N1 to support comparative studies, described in the DUR. Any other use of the Data is prohibited. [2]	Two repository agreements authorize data use: 1. Data Use Agreement 2. Data Use Request	[1] https://ncats.nih.gov/n3c/about/applying-for-access (Accessed 4/21/23) [2] https://ncats.nih.gov/sites/default/files/NCATS_N3C_Data_Use_Agreement.pdf (Accessed 4/21/23)
5.1.1.8	Other (specify)	Information not available/found	Information not available/found	
5.1.2	**Applicable Regulations/Policies**			
5.1.2.1	Local regulations/policies	N3C data may be used only for COVID-19 research purposes. [1] Use data only for the COVID-19-related research projects or comparative studies defined in the Data Use Request (DUR) approved by the N3C DAC for public health purposes and research to inform decision making. [2]	N3C Data User Code of Conduct	[1] NCATS Response to N3C Questions [2] https://ncats.nih.gov/n3c/resources/data-user-code-of-conduct (Accessed 4/21/23)

Dataset 8 - N3C				
		Raw Language	**Interpretation**	**Source**
5.1.2.2	Tribal regulations/policies	In alignment with the National Institutes of Health Guidance on the Implementation of the HHS Tribal Consultation Policy, NCATS sought input from Tribal Nations on whether and how to provide AI/AN data within the N3C. Now that the Tribal Consultation has been conducted, AI/AN data is available for research in a manner that reflects the input that NIH received through the consultation process. The Tribal Consultation report (PDF-1.3MB) details how this data is now being made available. [1] Based on the feedback from Tribal Consultation, NCATS will take the following steps to make AI/AN data available for research through its standard Data Use Request (DUR) process: 1. AI/AN data will be moved back to a standalone category. With this change, AI/AN data will be available in any N3C analysis that provides race and ethnicity distribution. 2. ZIP codes that overlap with Tribal communities will be available for research in the following manner: ZIP codes for all geographic units containing 20,000 or fewer people are removed entirely. This is standard practice for all geographic units and will be applied the same way when AI/AN data are restored to a separate category. For example, if a ZIP code of “01234” represents a community of 20,000 or fewer individuals, the user will see a ZIP code of “00000.” Currently, specific ZIP codes representing rural populations predominantly with AI/AN-identifying individuals are hidden. These will now be visible in both the limited data set and de-identified data. This means that AI/AN data will be managed as others are managed, with the exception that the full five-digit ZIP codes are never shown. For example, if a ZIP code of “01234” represents a predominantly AI/AN community, the user will see only a partial ZIP code of “012.” 3. The N3C Data User Code of Conduct will be modified so that data users will be asked to attest that they understand the N3C contains no Tribal affiliation data and that use of AI/AN data and ZIP code information to make assumptions about Tribal affiliation is not valid or appropriate. This statement will be included as a reminder when a DUR is received, a Data Use Agreement is executed, and data are accessed in the N3C platform and during publication processes. 4. The N3C will continue to engage NIH’s Tribal Health Research Office and Tribal Nations as issues for discussion arise. [2] It is important to note that the N3C does not have EHR data from the Indian Health Service, nor does it contain information about Tribal affiliation. [3]	1. NIH Guidance on the Implementation of the HHS Tribal Consultation Policy 2. Tribal Consultation Report	[1] https://ncats.nih.gov/n3c/about/program-faq# (Accessed 4/21/23) [2] https://ncats.nih.gov/n3c/about/tribal-consultation (Accessed 4/21/23) [3] https://dpcpsi.nih.gov/sites/default/files/N3C-Consultation-Report-508.pdf (Accessed 4/21/23)
5.1.2.3	State regulations/policies	Information not available/found	Information not available/found	

Dataset 8 - N3C				
		Raw Language	**Interpretation**	**Source**
5.1.2.4	Federal regulations/policies	In alignment with the National Institutes of Health Guidance on the Implementation of the HHS Tribal Consultation Policy, NCATS sought input from Tribal Nations on whether and how to provide AI/AN data within the N3C. Now that the Tribal Consultation has been conducted, AI/AN data is available for research in a manner that reflects the input that NIH received through the consultation process. The Tribal Consultation report (PDF-1.3MB) details how this data is now being made available. [1]	HHS Tribal Consultation Policy	[1] https://ncats.nih.gov/n3c/about/program-faq# (Accessed 4/21/23)
5.1.2.5	International regulations/policies	Information not available/found	Information not available/found	
5.1.2.6	Contractual obligations	Information not available/found	Information not available/found	
5.1.2.7	Repository policies	Before researchers can request access to the data, their institutions must execute a Data Use Agreement (DUA) with NCATS. Once a DUA is in place, researchers can submit a Data Use Request (DUR) through the N3C Data Enclave. In the DUR, researchers will need to include, among other information, the project research title, names of project personnel, a non-confidential research statement, the project proposal and the requested data access level. Additional DUR requirements include reviewing and agreeing to comply with the N3C Data User Code of Conduct. The N3C Data Access Committee reviews and approves DURs. [1] Data Access Committee (DAC), referred to as the NCATS N3C DAC, is composed of representatives from the NIH or other Federal Agencies and is responsible for reviewing Data Use Requests (DUR) and verifying that conditions for accessing Data are met. [2] Data Use Request (DUR) a document that User(s) must complete and submit to the DAC for review prior to accessing the NCATS N3C Data Enclave. For each individual project, the DUR, outlined in Appendix A, will contain overarching use objectives and designs, scientific goals, any testable hypotheses, and an outline of plans for using the Data. [2] This Community Guiding Principles document outlines expectations of every N3C community member, which includes, institutions and individuals who use N3C resources, expertise, and data. The N3C community is committed to providing a welcoming and inspiring environment for all community members. These principles were established by the N3C Community in an unprecedented effort to encourage and enable a highly collaborative research environment to address one of the worst health crises of our time. As such, we expect these guiding principles to be honored by each community member. [3] Purpose: This document provides community-driven guidelines and approaches that all users within the N3C research community uphold, and it addresses attribution and publication principles regarding N3C Community dissemination of research. These publication and attribution principles apply to N3C analysis reports, data, resources, abstracts, presentations, preprints, manuscripts, and other publications arising from the use of content in N3C Data Enclave. This document will be reviewed and updated regularly as N3C evolves, but at least annually by the N3C Governance Workstream and modified or recertified as deemed necessary. [4]	N3C policies	[1] https://ncats.nih.gov/n3c/about/applying-for-access (Accessed 4/21/23) [2] https://ncats.nih.gov/sites/default/files/NCATS_N3C_Data_Use_Agreement.pdf (Accessed 4/21/23) [3] https://zenodo.org/record/3979610 (Accessed 10/23/23) [4] https://zenodo.org/record/7787523 (Accessed 10/23/23)

Dataset 8 - N3C				
		Raw Language	**Interpretation**	**Source**
5.2	**Governance for data linkage, sharing, access, and use based on data access authorization or applicable regulations/policies (i.e., the origin of the governance)**			
5.2.1	Whether the data can be linked	Does not authorize/specify	Does not authorize/specify	
5.2.2	With what other data can it be linked or can it not be linked (scope of linkage)	Does not authorize/specify	Does not authorize/specify	
5.2.3	Whether data can be shared	Does not authorize/specify	Does not authorize/specify	
5.2.4	How data can be shared (de-identification status, disclosure review)	Does not authorize/specify	Does not authorize/specify	
5.2.5	How data can be accessed (access type, data use agreement, data access committee/group approval, IRB LOD, etc.)	An N3C Data Use Agreement (DUA) is executed by NCATS and a research institution (or directly with a researcher in the case of a citizen scientist who is not affiliated with an institution). To submit a DUA to NCATS, institutions can download this form and email it to NCATSPartnerships@mail.nih.gov (link sends e-mail). Note: DUAs can only be signed by Authorized Institutional Officials who have the authority to bind all users at their institution to the terms of the DUA. With the exception of citizen scientists not associated with institutions, individual researchers cannot sign their own DUAs with NCATS. It is strongly recommended that researchers ensure that their home institution has executed a DUA with NCATS before they begin the process of applying for data access. See the list of institutions(link is external) with active DUAs. DUAs will remain in effect for five years from the DUA Effective Date. They will automatically expire at the end of this period unless terminated or renewed. [1] If requesting access to the Limited Data Set, researchers will need to provide a copy of their institution's Human Research Protection Program IRB determination letter. [2] In order to access N3C data, researchers must submit an online Data Use Request (DUR) through the NCATS N3C Data Enclave for each project they want to start or join. Information researchers need to provide in the DUR includes: The project title Names of project personnel	N3C policy specifies that to access the Limited Dataset, the user: 1) Must complete N3C registration and create a N3C Data Enclave account 2) Must execute an Institutional Data Use Agreement 3) Must submit Data Use Request (DUR) for approval by N3C Data Access Committee 4) Must complete NIH IT training, attest to the N3C Data User Code of Conduct, and complete Human Subjects Research Protection training 5) Must provide IRB letter of determination for data access 6) Must access the data within the N3C Enclave	[1] https://ncats.nih.gov/n3c/resources/data-access (Accessed 4/21/23) [2] https://ncats.nih.gov/n3c/about/applying-for-access (Accessed 4/21/23) [3] NCATS Response to N3C Questions [4] Predecessor report (https://www.nichd.nih.gov/sites/default/files/inline-files/NICHD_ODSS_PPRL_for_Pediatric_COVID-19_Studies_Public_Final_Report_508.pdf)

Dataset 8 - N3C				
		Raw Language	**Interpretation**	**Source**
		A non-confidential research statement The project proposal The requested data access level The N3C Data Access Committee (DAC) reviews and approves DURs. Once the DAC approves a DUR, access to the N3C Data Enclave workspace will be effective for one year starting from the date access is granted. A DUA must be in place for the entire term of a DUR. DURs will be renewable. When users renew their DURs, they will need to attest at that time that their training for access to the N3C Data Enclave is up to date. N3C Data User Code of Conduct The N3C Data User Code of Conduct states core activities of data users and prohibitions on certain activities involving data or other resources accessible through the N3C Data Enclave. Researchers must agree to the N3C Data User Code of Conduct when they apply to access and use N3C data. NIH Information Security and Information Management Training The N3C Data Enclave is hosted by NCATS, and all researchers must complete the "Information Security, Counterintelligence, Privacy Awareness, Records Management Refresher, Emergency Preparedness Refresher" course, which can be accessed at NIH's information security training website, before submitting a DUR. It will take approximately 60-90 minutes to complete the entire course and users should save evidence of completion for their records (a screenshot or copy of the certificate of completion). [1] Data are not allowed to leave the Enclave; analyses must be done within the platform and analysis results can be shared broadly. [3] Local IRB letter of determination is required for accessing data (see Appendix Table 2) [4]		

Dataset 8 - N3C				
		Raw Language	**Interpretation**	**Source**
5.2.6	How data can be used (data use limitations)	User(s) agree(s) to use the Data exclusively for the Research Project proposed, and/or comparative studies using data from individuals infected with pathogens such as SARS, MERS, and H1N1 to support comparative studies, described in the DUR. Any other use of the Data is prohibited. [1] N3C data may be used only for COVID-19 research purposes. [2] Use data only for the COVID-19-related research projects or comparative studies defined in the Data Use Request (DUR) approved by the N3C DAC for public health purposes and research to inform decision making. [3] Based on the feedback from Tribal Consultation, NCATS will take the following steps to make AI/AN data available for research through its standard Data Use Request (DUR) process: 1. AI/AN data will be moved back to a standalone category. With this change, AI/AN data will be available in any N3C analysis that provides race and ethnicity distribution. 2. ZIP codes that overlap with Tribal communities will be available for research in the following manner: ZIP codes for all geographic units containing 20,000 or fewer people are removed entirely. This is standard practice for all geographic units and will be applied the same way when AI/AN data are restored to a separate category. For example, if a ZIP code of “01234” represents a community of 20,000 or fewer individuals, the user will see a ZIP code of “00000.” Currently, specific ZIP codes representing rural populations predominantly with AI/AN-identifying individuals are hidden. These will now be visible in both the	1. N3C DUA specifies that the data must be used exclusively for the Research Project proposed, and/or comparative studies using data from individuals infected with pathogens such as SARS, MERS, and H1N1 to support comparative studies. 2. N3C Data User Code of Conduct specifies that N3C data must only be used for COVID-19 general research purposes. 3. NIH Guidance on the Implementation of the HHS Tribal Consultation Policy, the HHS Tribal Consultation Policy, and the Tribal Consultation Report specify that data users will be asked to attest that they understand the N3C contains no Tribal affiliation data and that use of AI/AN data and ZIP code information to make assumptions about Tribal affiliation is not valid or appropriate. 4. User must also comply with the N3C Community Guiding Principles and the Attribution and Publication Principles.	[1] https://ncats.nih.gov/sites/default/files/NCATS_N3C_Data_Use_Agreement.pdf (Accessed 4/21/23) [2] NCATS Response to N3C Questions [3] https://ncats.nih.gov/n3c/resources/data-user-code-of-conduct (Accessed 4/21/23) [4] https://ncats.nih.gov/n3c/about/tribal-consultation (Accessed 4/21/23) [5] https://dpcpsi.nih.gov/sites/default/files/N3C-Consultation-Report-508.pdf (Accessed 4/21/23) [6] https://zenodo.org/record/3979610 (Accessed 10/23/23) [7] https://zenodo.org/record/7787523 (Accessed 10/23/23)

Dataset 8 - N3C				
		Raw Language	**Interpretation**	**Source**
		limited data set and de-identified data. This means that AI/AN data will be managed as others are managed, with the exception that the full five-digit ZIP codes are never shown. For example, if a ZIP code of “01234” represents a predominantly AI/AN community, the user will see only a partial ZIP code of “012.” 3. The N3C Data User Code of Conduct will be modified so that data users will be asked to attest that they understand the N3C contains no Tribal affiliation data and that use of AI/AN data and ZIP code information to make assumptions about Tribal affiliation is not valid or appropriate. This statement will be included as a reminder when a DUR is received, a Data Use Agreement is executed, and data are accessed in the N3C platform and during publication processes. 4. The N3C will continue to engage NIH’s Tribal Health Research Office and Tribal Nations as issues for discussion arise. [4] It is important to note that the N3C does not have EHR data from the Indian Health Service, nor does it contain information about Tribal affiliation. [5] This Community Guiding Principles document outlines expectations of every N3C community member, which includes, institutions and individuals who use N3C resources, expertise, and data. The N3C community is committed to providing a welcoming and inspiring environment for all community members. These principles were established by the N3C Community in an unprecedented effort to encourage and enable a highly collaborative research environment to address one of the worst health crises of our time. As such, we expect these guiding principles to be honored by each community member. [6] Purpose: This document provides community-driven guidelines and approaches that all users within the N3C research community uphold, and it addresses attribution and publication principles regarding N3C Community dissemination of research. These publication and attribution principles apply to N3C analysis reports, data, resources, abstracts, presentations, preprints, manuscripts, and other publications arising from the use of content in N3C Data Enclave. This document will be reviewed and updated regularly as N3C evolves, but at least annually by the N3C Governance Workstream and modified or recertified as deemed necessary. [7]		
5.2.7	Other (specify)	Does not authorize/specify	Does not authorize/specify	
6	**PII Elements**			
6.1	PII elements collected	Combinations of PII elements to generate 18 tokens per record (last name, first name, DOB, gender, SSN, email, zip5/9, cell phone) [1]	Last name, first name, DOB, gender, SSN, email, zip5/9, cell phone number are collected	[1] Predecessor report (https://www.nichd.nih.gov/sites/default/files/inline-files/NICHD_ODSS_PPRL_for_Pediatric_COVID-19_Studies_Public_Final_Report_508.pdf)

Dataset 8 - N3C				
		Raw Language	**Interpretation**	**Source**
6.2	PII elements holder (i.e., party that holds the PII)	The data remains under the complete control of the organizations that provide data to N3C and is never accessible by or under the control of the linkage honest broker.	N3C data partners hold the PII.	[1] https://ncats.nih.gov/n3c/about/data-overview (Accessed: 4/19/23)
6.3	Use of common data model, if any, for data collection	COVID-19 Clinical Data Warehouse Data Dictionary Based on OMOP Common Data Model Specifications Version 5.3 [1]	N3C data uses OMOP.	[1] https://ncats.nih.gov/sites/default/files/OMOP_CDM_COVID.pdf (Accessed 4/19/23)
7	**Prior Data Linkages**			
7.1	**Dataset linked with other datasets**			
7.1.1	Name of other dataset linked to this dataset	Information not available/found	Information not available/found	
7.1.2	Other dataset type (clinical, EHR, survey, claims, SDOH, etc.)	N3C Class 2 External Data Linkage • Description: The N3C Data Enclave is a centralized, secure, national clinical data resource with powerful analytics capabilities that the research community can use to study COVID-19, including potential risk factors, protective factors and long-term health consequences. The N3C collects data derived from the EHRs of people who were tested for COVID-19 or who had related symptoms, as well as data from individuals infected with pathogens that can support comparative studies, such as SARS 1, MERS and H1N1. N3C facilitates the linking of external datasets with EHR data available in the N3C enclave using PPRL to generate a richer dataset that can answer new questions about COVID-19. The details below specifically pertain to PPRL linkage with two examples of external data linkages that fall under Class 2 and are currently performed within N3C—viral variant data and mortality data. • Data sources: EHR data in the N3C enclave (submitted by N3C data partners) and the following external datasets: o Viral variant data: Collected by N3C data partners but stored in NIH's National Center for Biotechnology Information (NCBI) Sequence Read Archive (SRA); currently only viral variant summary data provided by the N3C data partners are linked. When the viral variant sequence data use case is ready, it will be imported into N3C. o Mortality data: The mortality data sources include government mortality sources (death certificates and person-reporting), ObituaryData.com, and private obituary sites) [1]	1. Viral variant 2. Mortality data	[1] Predecessor report https://www.nichd.nih.gov/sites/default/files/inline-files/NICHD_ODSS_PPRL_for_Pediatric_COVID-19_Studies_Public_Final_Report_508.pdf

Dataset 8 - N3C				
		Raw Language	**Interpretation**	**Source**
7.1.3	Other dataset source(s)	N3C Class 2 External Data Linkage • Description: The N3C Data Enclave is a centralized, secure, national clinical data resource with powerful analytics capabilities that the research community can use to study COVID-19, including potential risk factors, protective factors and long-term health consequences. The N3C collects data derived from the EHRs of people who were tested for COVID-19 or who had related symptoms, as well as data from individuals infected with pathogens that can support comparative studies, such as SARS 1, MERS and H1N1. N3C facilitates the linking of external datasets with EHR data available in the N3C enclave using PPRL to generate a richer dataset that can answer new questions about COVID-19. The details below specifically pertain to PPRL linkage with two examples of external data linkages that fall under Class 2 and are currently performed within N3C—viral variant data and mortality data. • Data sources: EHR data in the N3C enclave (submitted by N3C data partners) and the following external datasets: o Viral variant data: Collected by N3C data partners but stored in NIH's National Center for Biotechnology Information (NCBI) Sequence Read Archive (SRA); currently only viral variant summary data provided by the N3C data partners are linked. When the viral variant sequence data use case is ready, it will be imported into N3C. o Mortality data: The mortality data sources include government mortality sources (death certificates and person-reporting), ObituaryData.com, and private obituary sites) [1]	1. N3C Data partners 2. Government mortality sources (death certificates and person-reporting), ObituaryData.com, and private obituary sites	[1] Predecessor report https://www.nichd.nih.gov/sites/default/files/inline-files/NICHD_ODSS_PPRL_for_Pediatric_COVID-19_Studies_Public_Final_Report_508.pdf
7.1.4	Linking methodology (PPRL or non-PPRL); linkage technology	N3C facilitates the linking of external datasets with EHR data available in the N3C enclave using PPRL to generate a richer dataset that can answer new questions about COVID-19. [1] From Table 12: PPRL Vendors and Tools "Datavant: Used by N3C and for NIH funded studies in PEDSnet" [1]	PPRL, Datavant	[1] Predecessor report https://www.nichd.nih.gov/sites/default/files/inline-files/NICHD_ODSS_PPRL_for_Pediatric_COVID-19_Studies_Public_Final_Report_508.pdf
7.1.5	PII elements used for the linkage	Combinations of PII elements to generate 18 tokens per record (last name, first name, DOB, gender, SSN, email, zip5/9, cell phone) [1]	Combinations of PII elements to generate 18 tokens per record (last name, first name, DOB, gender, SSN, email, zip5/9, cell phone)	[1] Predecessor report https://www.nichd.nih.gov/sites/default/files/inline-files/NICHD_ODSS_PPRL_for_Pediatric_COVID-19_Studies_Public_Final_Report_508.pdf
7.1.6	Entity resolver (data originator or data linker or third party)	In all three cases, entity resolution is performed by a linkage honest broker (LHB), Regenstrief, based on matching the hashed tokens generated by N3C EHR data partners and external enclaves/data sources (for Class 0 and 2) [1]	Regenstrief (Honest Broker)	[1] Predecessor report https://www.nichd.nih.gov/sites/default/files/inline-files/NICHD_ODSS_PPRL_for_Pediatric_COVID-19_Studies_Public_Final_Report_508.pdf

Dataset 8 - N3C				
		Raw Language	**Interpretation**	**Source**
7.1.7	Party performing the linkages	Entity resolution is performed by the LHB, and data linkage is done within N3C enclave by requesting researchers. [1]	Researchers	[1] Predecessor report https://www.nichd.nih.gov/sites/default/files/inline-files/NICHD_ODSS_PPRL_for_Pediatric_COVID-19_Studies_Public_Final_Report_508.pdf
7.1.8	Linkage quality assessment	A linkage quality assessment report is being prepared by N3C. In general, the linkage quality measures will depend on the linkage use case—for example: the linkage quality for EHR data linkage is stringent to support academic research but could be less stringent for recruitment use cases (what N3C calls “cohort discovery”) [1]	Linkage quality measures depend on the linkage use case	[1] Predecessor report https://www.nichd.nih.gov/sites/default/files/inline-files/NICHD_ODSS_PPRL_for_Pediatric_COVID-19_Studies_Public_Final_Report_508.pdf
7.1.9	Linked data sharing method (linkage maps or pre-linked dataset)	All three N3C PPRL implementations —internal EHR to EHR linkage and external Class 0 and Class 2 linkages —follow the linked database model where entity resolution occurs as the data are received by N3C and tokens are sent to the linkage honest broker from the data partners and the external sources. [1]	Linked dataset	[1] Predecessor report https://www.nichd.nih.gov/sites/default/files/inline-files/NICHD_ODSS_PPRL_for_Pediatric_COVID-19_Studies_Public_Final_Report_508.pdf

USE CASE 3 - GOVERNANCE INFORMATION				
Use Case 3: SARS-CoV-2 Vaccination and Asthma-Related School Absence - Does SARS-CoV-2 vaccination result in reduced asthma-related school absences at 3/6/12+ months post-vaccination?				
Dataset 9 - PEDSnet Limited Data Sets (LDS) with Exact Dates				
	Dataset Source	PCORnet		
	Dataset Source Agency	PCORI		
	Dataset Type (Clinical, EHR, Survey, SDOH, etc.)	EHR		
	Information Sources	Website, Information from predecessor report, CHOP/PEDSnet Interview (as part of the predecessor report), Master IRB Protocol provided by PEDSnet, Responses to PEDSnet Questions		
Dataset 9 - PEDSnet Limited Data Sets (LDS)				
		Raw Language	**Interpretation**	**Source**
1	**Data Collection**			
1.1	**Authorizations and Applicable Regulations/Policies**			
1.1.1	**Authorizations**		1. Assent 2. Consent 3. CHOP or local/site-level IRB approval 4. Master Reliance Agreement 5. Waiver of consent	
1.1.1.1	Assent	Please confirm our understanding that PEDSnet data are collected either under a waiver of consent (large observational studies) or under consent/assent (specific studies)? [1, 2] "This is substantially correct. Data from routine delivery of health care are collected under waiver; collection of study-specific data from patients and families is with consent/assent." [2]	Assent authorizes data collection.	[1] CHOP/PEDSnet Interview (as part of predecessor report) [2] Responses to PEDSnet Questions
1.1.1.2	Consent	Please confirm our understanding that PEDSnet data are collected either under a waiver of consent (large observational studies) or under consent/assent (specific studies)? [1, 2] "This is substantially correct. Data from routine delivery of health care are collected under waiver; collection of study-specific data from patients and families is with consent/assent." [2]	Consent authorizes data collection.	[1] CHOP/PEDSnet Interview (as part of predecessor report) [2] Responses to PEDSnet Questions

Dataset 9 - PEDSnet Limited Data Sets (LDS)				
		Raw Language	**Interpretation**	**Source**
1.1.1.3	IRB/equivalent Privacy Board determination	All PEDSnet institutional members are required to endorse the PEDSnet Single IRB policy by becoming a participating institution of the NCATS sponsored SMART IRB Agreement. Participation in these agreements allows an institution to choose on a case-by-case protocol basis whether to participate in a ceded review, as a relying or reviewing institution, or perform its own IRB review. Given the network's underlying principle of collaboration, PEDSnet expects that IRB reliance opt-out will be a rare occurrence, and would require an appropriate justification. [1] This project will use the PEDSnet Reciprocal Single IRB Master Reliance Agreement with CHOP acting as the IRB of record and the other 13 institutions relying on the CHOP IRB. [2] PEDSnet centers currently capture a wide variety of data as part of clinical operations obtained from primary care, Emergency Department (ED), specialty care, and inpatient settings. Participating sites will collect data for PEDSnet by extraction and transformation to the network data model of these data collected as part of standard clinical and business operations. The specific characteristics of each site's data extraction are described in the site-specific appendices to approved protocol #14-011242, for sites using the CHOP IRB as the IRB of record for these activities, or in separate protocols reviewed by the local IRB. [2] EHR data collection and de-identification at sites is governed by site-level IRB oversight. [3]	CHOP IRB or local/site-level IRB approval authorizes data collection.	[1] https://pedsnet.org/resources/governance/ (Accessed: 4/18/23) [2] Master IRB Protocol [3] Responses to PEDSnet Questions
1.1.1.4	Local/state/federal laws	Information not available/found	Information not available/found	
1.1.1.5	Institutional Certification	Information not available/found	Information not available/found	
1.1.1.6	Data originator agreement	From the launch the network, PEDSnet adapted and executed a Master Reliance Agreement among all its members, which set forth the respective authorities, roles, and responsibilities of each party when a ceded review arrangement is determined to be acceptable. [1] Does the Master Reliance Agreement executed among PEDSnet partners authorize EHR data collection by the partners? [1, 2] "There is a master protocol with reliance by all PEDSnet members that authorizes observational research using EHR data" [2]	Master Reliance Agreement authorizes data collection under one IRB.	[1] https://pedsnet.org/resources/irb/ (Accessed: 4/18/23) [2] Responses to PEDSnet Questions
1.1.1.7	Repository agreements/policies	Information not available/found	Information not available/found	

Dataset 9 - PEDSnet Limited Data Sets (LDS)				
		Raw Language	**Interpretation**	**Source**
1.1.1.8	Other (specify)	Please confirm our understanding that PEDSnet data are collected either under a waiver of consent (large observational studies) or under consent/assent (specific studies)? [1, 2] "This is substantially correct. Data from routine delivery of health care are collected under waiver; collection of study-specific data from patients and families is with consent/assent." [2]	Waiver of consent authorizes data collection for routine delivery of health care.	[1] CHOP/PEDSnet Interview (as part of predecessor report) [2] Responses to PEDSnet Questions
1.1.2	**Applicable Regulations/Policies**			
1.1.2.1	Local regulations/policies	Information not available/found	Information not available/found	
1.1.2.2	Tribal regulations/policies	Information not available/found	Information not available/found	
1.1.2.3	State regulations/policies	Information not available/found	Information not available/found	
1.1.2.4	Federal regulations/policies	Information not available/found	Information not available/found	
1.1.2.5	International regulations/policies	Information not available/found	Information not available/found	
1.1.2.6	Contractual obligations	Information not available/found	Information not available/found	
1.1.2.7	Repository policies	Information not available/found	Information not available/found	
1.2	**Governance for data linkage, sharing, access, and use based on data collection authorization or applicable regulations/policies (i.e., the origin of the governance)**			
1.2.1	Whether the data can be linked	Consent for data linking is embedded in the broad study consent (when obtained). [1]	Consent (when obtained) specifies that data can be linked.	[1] CHOP/PEDSnet Interview (as part of predecessor report)
1.2.2	With what other data can it be linked or can it not be linked (scope of linkage)	We also understand that consent for data linking is embedded in the broad study consent (when obtained). Are there any other authorizations for data linkages? [1, 2] Authorization for privacy-preserving record linkage is part of governing IRB protocols for research conducted under waiver of consent. [2]	Governing IRB protocols specify that data can be linked using PPRL for research conducted under a waiver of consent.	[1] CHOP/PEDSnet Interview (as part of predecessor report) [2] Responses to PEDSnet Questions
1.2.3	Whether data can be shared	Does not authorize/specify	Does not authorize/specify	
1.2.4	How data can be shared (de-identification status, disclosure review)	Does not authorize/specify	Does not authorize/specify	
1.2.5	How data can be accessed (access type, data use agreement, data access committee/group approval, IRB LOD, etc.)	Does not authorize/specify	Does not authorize/specify	
1.2.6	How data can be used (data use limitations)	Does not authorize/specify	Does not authorize/specify	
1.2.7	Other (specify)	Does not authorize/specify	Does not authorize/specify	
2	**Data Linkage**			
2.1	**Authorizations and Applicable Regulations/Policies**			
2.1.1	**Authorizations**		1. Consent 2. IRB 3. PEDSnet Steering Committee approval 4. Individual PEDSnet sites/study participation vote	
2.1.1.1	Assent	Information not available/found	Information not available/found	
2.1.1.2	Consent	Consent for data linking is embedded in the broad study consent (when obtained). [1]	Consent (when obtained) authorizes data linkage.	[1] CHOP/PEDSnet Interview (as part of predecessor report)

Dataset 9 - PEDSnet Limited Data Sets (LDS)				
		Raw Language	**Interpretation**	**Source**
2.1.1.3	IRB/equivalent Privacy Board determination	We also understand that consent for data linking is embedded in the broad study consent (when obtained). Are there any other authorizations for data linkages? [1, 2] Authorization for privacy-preserving record linkage is part of governing IRB protocols for research conducted under waiver of consent. [2]	IRB authorizes data linkage for research conducted under a waiver of consent.	[1] CHOP/PEDSnet Interview (as part of predecessor report) [2] Responses to PEDSnet Questions
2.1.1.4	Local/state/federal laws	Linkage is determined by the PEDSnet sites based on the particular use case/research study. When sites sign up to be part of the PEDSnet network, they must agree to hash their data; however, individual PEDSnet study sites can decide whether to participate in data linkages on a study-by-study basis once the PEDSnet Steering Committee approves the data linkage as part of a specific study research plan. [1]	PEDSnet Steering Committee approval authorizes data linkage as part of a specific study research plan.	[1] Predecessor report (https://www.nichd.nih.gov/sites/default/files/inline-files/NICHD_ODSS_PPRL_for_Pediatric_COVID-19_Studies_Public_Final_Report_508.pdf)
2.1.1.5	Institutional Certification	Information not available/found	Information not available/found	
2.1.1.6	Data originator agreement	Linkage is determined by the PEDSnet sites based on the particular use case/research study. When sites sign up to be part of the PEDSnet network, they must agree to hash their data; however, individual PEDSnet study sites can decide whether to participate in data linkages on a study-by-study basis once the PEDSnet Steering Committee approves the data linkage as part of a specific study research plan. [1] Is the decision to participate in data linking documented via the 'written affirmation from the study site PI' or how is it authorized/documented? Where the need for linkage is known at the inception of the study (preferred), it is part of the study participation vote. Where an opportunity for linkage arises during conduct of a study (e.g. new dataset becomes available for linkage), a separate authorization is obtained. [2]	Individual PEDSnet sites authorize data linkage on a study-by-study basis.	[1] Predecessor report (https://www.nichd.nih.gov/sites/default/files/inline-files/NICHD_ODSS_PPRL_for_Pediatric_COVID-19_Studies_Public_Final_Report_508.pdf) [2] Responses to PEDSnet Questions
2.1.1.7	Repository agreements/policies	Information not available/found	Information not available/found	
2.1.1.8	Other (specify)	Information not available/found	Information not available/found	
2.1.2	**Applicable Regulations/Policies**			
2.1.2.1	Local regulations/policies	Information not available/found	Information not available/found	
2.1.2.2	Tribal regulations/policies	Information not available/found	Information not available/found	
2.1.2.3	State regulations/policies	Information not available/found	Information not available/found	
2.1.2.4	Federal regulations/policies	Information not available/found	Information not available/found	
2.1.2.5	International regulations/policies	Information not available/found	Information not available/found	
2.1.2.6	Contractual obligations	Information not available/found	Information not available/found	
2.1.2.7	Repository policies	Information not available/found	Information not available/found	
2.2	**Governance for data linkage, sharing, access, and use based on data linkage authorization or applicable regulations/policies (i.e., the origin of the governance)**			

Dataset 9 - PEDSnet Limited Data Sets (LDS)				
		Raw Language	**Interpretation**	**Source**
2.2.1	Whether the data can be linked	Consent for data linking is embedded in the broad study consent (when obtained). [1] Linkage is determined by the PEDSnet sites based on the particular use case/research study. When sites sign up to be part of the PEDSnet network, they must agree to hash their data; however, individual PEDSnet study sites can decide whether to participate in data linkages on a study-by-study basis once the PEDSnet Steering Committee approves the data linkage as part of a specific study research plan. [2] Is the decision to participate in data linking documented via the 'written affirmation from the study site PI' or how is it authorized/documented? Where the need for linkage is known at the inception of the study (preferred), it is part of the study participation vote. Where an opportunity for linkage arises during conduct of a study (e.g. new dataset becomes available for linkage), a separate authorization is obtained. [3] We also understand that consent for data linking is embedded in the broad study consent (when obtained). Are there any other authorizations for data linkages? [1, 3] Authorization for privacy-preserving record linkage is part of governing IRB protocols for research conducted under waiver of consent. [3]	1. Consent (when obtained) specifies that data can be linked. 2. PEDSnet Steering Committee approval specifies that data can be linked according to the approved research plan. 3. Individual PEDSnet sites, through a study particpation vote, specify that the sites can participation in data linkage on a study-by-study basis. 4. IRB specifies that PEDSnet data can be linked for research conducted under a waiver of consent.	[1] CHOP/PEDSnet Interview (as part of predecessor report) [2] Predecessor report (https://www.nichd.nih.gov/sites/default/files/inline-files/NICHD_ODSS_PPRL_for_Pediatric_COVID-19_Studies_Public_Final_Report_508.pdf) [3] Responses to PEDSnet Questions
2.2.2	With what other data can it be linked or can it not be linked (scope of linkage)	We also understand that consent for data linking is embedded in the broad study consent (when obtained). Are there any other authorizations for data linkages? [1, 2] Authorization for privacy-preserving record linkage is part of governing IRB protocols for research conducted under waiver of consent. [2] Linkage is determined by the PEDSnet sites based on the particular use case/research study. When sites sign up to be part of the PEDSnet network, they must agree to hash their data; however, individual PEDSnet study sites can decide whether to participate in data linkages on a study-by-study basis once the PEDSnet Steering Committee approves the data linkage as part of a specific study research plan. [3]	1. IRB specifies that data can be linked using PPRL for research conducted under a waiver of consent. 2. Individual PEDSnets study sites specify the scope of data linkage on a study-by-study basis.	[1] CHOP/PEDSnet Interview (as part of predecessor report) [2] Responses to PEDSnet Questions [3] Predecessor report (https://www.nichd.nih.gov/sites/default/files/inline-files/NICHD_ODSS_PPRL_for_Pediatric_COVID-19_Studies_Public_Final_Report_508.pdf)
2.2.3	Whether data can be shared	Does not authorize/specify	Does not authorize/specify	
2.2.4	How data can be shared (de-identification status, disclosure review)	Does not authorize/specify	Does not authorize/specify	
2.2.5	How data can be accessed (access type, data use agreement, data access committee/group approval, IRB LOD, etc.)	Does not authorize/specify	Does not authorize/specify	
2.2.6	How data can be used (data use limitations)	Each site must authorize use of their data and resources in a particular study via a participation vote. [1]	PEDSnet site participation vote specifies data use for individual studies.	[1] Responses to PEDSnet Questions

Dataset 9 - PEDSnet Limited Data Sets (LDS)				
		Raw Language	**Interpretation**	**Source**
2.2.7	Other (specify)	Does not authorize/specify	Does not authorize/specify	
3	**Data Sharing**			
3.1	**Authorizations and Applicable Regulations/Policies**			
3.1.1	**Authorizations**		1. Data release vote 2. PEDSnet Participation and DUA 3. PEDSnet Master Data Use Agreement	
3.1.1.1	Assent	Information not available/found	Information not available/found	
3.1.1.2	Consent	Information not available/found	Information not available/found	
3.1.1.3	IRB/equivalent Privacy Board determination	Information not available/found	Information not available/found	
3.1.1.4	Local/state/federal laws	Information not available/found	Information not available/found	
3.1.1.5	Institutional Certification	Information not available/found	Information not available/found	
3.1.1.6	Data originator agreement	Institutional members will be required to sign: • Participation and Data Use Agreement, which an institutional official must sign [1] In addition, each institution must develop and submit a data governance process and procedures for reviewing and approving release of institutional data. [1] In addition, if the study involves release of a dataset beyond the PEDSnet data center, each site must authorize the release of their data in a dataset release vote. [2]	Two agreements authorize data sharing: 1. PEDSnet Participation and DUA signed by PEDSnet members authorizes data sharing between PEDSnet members. 2. PEDSnet sites data release vote authorizes data sharing (i.e., release).	[1] https://pedsnet.org/documents/303/PEDSnet-Policies-May-2020_.pdf [2] Responses to PEDSnet Questions
3.1.1.7	Repository agreements/policies	The PEDSnet Master DUA authorizes sharing of limited data sets for use in studies. [1]	PEDSnet Master Data Use Agreement authorizes data sharing.	[1] Responses to PEDSnet Questions
3.1.1.8	Other	Information not available/found	Information not available/found	
3.1.2	**Applicable Regulations/Policies**			
3.1.2.1	Local regulations/policies	Note that access to individual-level data, whether meeting the Privacy Rule Safe Harbor standard or the definition of a Limited Data Set, is governed by the policies for Limited Data Sets, due to the potential for pattern-based reidentification in highly detailed data. [1] Limited data sets with and without actual dates: This is the typical case. Information about standard privacy risk reduction measures, including date shifting, can be found in PEDSnet Policies on our web site. [1]	PEDSnet policy	[1] Responses to PEDSnet Questions
3.1.2.2	Tribal regulations/policies	Information not available/found	Information not available/found	

Dataset 9 - PEDSnet Limited Data Sets (LDS)				
		Raw Language	**Interpretation**	**Source**
3.1.2.3	State regulations/policies	All PEDSnet member systems are based in states that legislate specific protections for minor patients accessing reproductive and mental health care, and state and federal regulations affect the privacy of HIV-related data. [1] Removal of data relating to testing or care for HIV, pregnancy, and mental health, excluding educational performance. [2]	State regulations	[1] Responses to PEDSnet Questions [2] https://pedsnet.org/resources/governance/ (Accessed: 4/18/23)
3.1.2.4	Federal regulations/policies	All PEDSnet member systems are based in states that legislate specific protections for minor patients accessing reproductive and mental health care, and state and federal regulations affect the privacy of HIV-related data. [1] Access to individual-level data is governed by the policies described above, and appropriate regulatory oversight as defined in the Common Rule (or equivalent) and HIPAA. [1] Removal of data relating to testing or care for HIV, pregnancy, and mental health, excluding educational performance. [2] This study will be conducted in full accordance with all applicable Children's Hospital of Philadelphia Research Policies and Procedures and all applicable Federal and state laws and regulations including 45 CFR 46, and the HIPAA Privacy Rule. [3]	1. Federal regulations 2. 45 CFR 46 (Common Rule) 3. HIPAA Privacy Rule	[1] Responses to PEDSnet Questions [2] https://pedsnet.org/resources/governance/ (Access: 4/18/23) [3] Master IRB Protocol
3.1.2.5	International regulations/policies	Information not available/found	Information not available/found	
3.1.2.6	Contractual obligations	Information not available/found	Information not available/found	
3.1.2.7	Repository policies	Information not available/found	Information not available/found	
3.2	**Governance for data linkage, sharing, access, and use based on data sharing authorization or applicable regulations/policies (i.e., the origin of the governance)**			
3.2.1	Whether the data can be linked	Does not authorize/specify	Does not authorize/specify	
3.2.2	With what other data can it be linked or can it not be linked (scope of linkage)	Does not authorize/specify	Does not authorize/specify	

Dataset 9 - PEDSnet Limited Data Sets (LDS)				
		Raw Language	**Interpretation**	**Source**
3.2.3	Whether data can be shared	In addition, if the study involves release of a dataset beyond the PEDSnet data center, each site must authorize the release of their data in a dataset release vote. [1] The PEDSnet Master DUA authorizes sharing of limited data sets for use in studies. [1] Access to individual-level data is governed by the policies described above, and appropriate regulatory oversight as defined in the Common Rule (or equivalent) and HIPAA. [1] This study will be conducted in full accordance with all applicable Children's Hospital of Philadelphia Research Policies and Procedures and all applicable Federal and state laws and regulations including 45 CFR 46, and the HIPAA Privacy Rule. [2]	1. PEDSnet sites data release vote specifies whether data can be shared (i.e., release). 2. The PEDSnet Master DUA specifies whether data can be shared. 3. 45 CFR 46 (Common Rule) and HIPAA Privacy Rule specifies that de-identified data can be shared.	[1] Responses to PEDSnet Questions [2] Master IRB Protocol
3.2.4	How data can be shared (de-identification status, disclosure review)	All PEDSnet member systems are based in states that legislate specific protections for minor patients accessing reproductive and mental health care, and state and federal regulations affect the privacy of HIV-related data. Accounting for these needs is part of the PEDSnet study review process and the privacy risk reduction procedures described above. Institutional policies regarding prudential use of data are incorporated into the site's study participation review and dataset release processes. [1] Access to individual-level data is governed by the policies described above, and appropriate regulatory oversight as defined in the Common Rule (or equivalent) and HIPAA. Note that access to individual-level data, whether meeting the Privacy Rule Safe Harbor standard or the definition of a Limited Data Set, is governed by the policies for Limited Data Sets, due to the potential for pattern-based reidentification in highly detailed data. [1] DE-IDENTIFIED INDIVIDUAL LEVEL DATA (according to the Safe Harbor guidelines). Patient-level data sets may be released to investigators within or outside of the institution of origin with an IRB determination of non-human subjects research (NHSR). The receipt of NHSR determination is for documentation purposes. For the purposes of PEDSnet, to be considered de-identified, the data set must use the safe harbor method for De-identification of Protected Health Information. [2] Datasets containing any person-level records that are not synthetic, or have a k-anonymity of ≤10, (i.e. have any records that are identical for 10	1. State and federal regulations specify that HIV-related data and reproductive and mental health health care data for minors must be removed before sharing data. 2. PEDSnet regulations specify that individual level data must be de-identified using the Safe Harbor method of de-identification of PHI before sharing data. 3. PEDSnet policy specifies that a risk review is performed on the requested datasets as well as data transformations, such as date shifts, replacement labels for free text fields and geographic information, and removing HIV/pregnancy/ mental health data prior to data sharing.	[1] Responses to PEDSnet Questions [2] https://pedsnet.org/resources/data-governance/ (Accessed: 4/18/23) [3] https://pedsnet.org/resources/governance/ (Accessed: 4/18/23) [4] Predecessor report (https://www.nichd.nih.gov/sites/default/files/inline-files/NICHD_ODSS_PPRL_for_Pediatric_COVID-19_Studies_Public_Final_Report_508.pdf)

Dataset 9 - PEDSnet Limited Data Sets (LDS)				
		Raw Language	**Interpretation**	**Source**
		or fewer persons), are subject to the following, unless specifically waived during the PEDSnet review process: 6. Replacement of stable identifiers with study-specific identifiers. 7. Shifting of dates within a one-year window centered on the actual date. 8. Replacement of free-text fields with single-use labels. 9. Replacement of geographic information with single-use labels. [3] A risk review is done for each proposed study, but no separate deductive disclosure review of the linked data is performed. Data use agreement prohibits reidentification and re-use for other studies. All members must agree to the PEDSnet standard policies when joining the Network; additional terms of data use are stipulated on a case-by-case basis. [4]		
3.2.5	How data can be accessed (access type, data use agreement, data access committee/group approval, IRB LOD, etc.)	Does not authorize/specify	Does not authorize/specify	
3.2.6	How data can be used (data use limitations)	Does not authorize/specify	Does not authorize/specify	
3.2.7	Other (specify)	Does not authorize/specify	Does not authorize/specify	
4	**Data Access**			
4.1	**Authorizations and Applicable Regulations/Policies**			
4.1.1	**Authorizations**		1. Human Subjects Review (Letter of determination) 2. IRB approval for requester if HSR 3. IRB reliance agreement (NPRA MRA or SMART IRB MRA) if HSR 4. Legal Review (DUA and RUD) 5. Network Participation Approval (PEDSnet Executive Committee) 6. Institutional Participation Approval (Prospective Site PI Approval)	
4.1.1.1	Assent	Information not available/found	Information not available/found	
4.1.1.2	Consent	Information not available/found	Information not available/found	

Dataset 9 - PEDSnet Limited Data Sets (LDS)				
		Raw Language	**Interpretation**	**Source**
4.1.1.3	IRB/equivalent Privacy Board determination	Human Subjects Review: IRB Determination/Approval required by PEDSnet site -If NHSR/Exempt: no further review/MRA required -If HSR: IRB approval with IRB reliance for sites providing data (NPRA MRA or SMART IRB MRA) [1]	Human Subjects Review (IRB review or determination) authorizes data access in two possible paths: 1. Non-Human Subjects Research (NHSR) determination: no further review/MRA required 2. Human Subjects Research (HSR) determination: requester IRB approval with IRB reliance for site providing data (NPRA MRA or SMART IRB MRA is also required	[1] https://pedsnet.org/resources/governance/ (Accessed: 4/18/23)
4.1.1.4	Local/state/federal laws	Information not available/found	Information not available/found	
4.1.1.5	Institutional Certification	Information not available/found	Information not available/found	
4.1.1.6	Data originator agreement	Institutional Participation Approval: Prospective Site PI Approval for all data requests regardless of the de-identification status of the dataset being requested. [1]	Institutional participation approval (prospective site PI approval) authorizes data access.	[1] https://pedsnet.org/resources/governance/ (Accessed: 4/18/23)
4.1.1.7	Repository agreements/policies	Legal Review: PEDSnet DUA and RUD [1]	Legal review (comprised of PEDSnet Data use agreement and Responsible Use of Data Agreement) authorizes data access.	[1] https://pedsnet.org/resources/governance/ (Accessed: 4/18/23)
4.1.1.8	Other (specify)	Network Participation Approval: Executive Committee Approval [1]	Network Participation Approval (PEDSnet Executive Committee Approval) authorizes data access.	[1] https://pedsnet.org/resources/governance/ (Accessed: 4/18/23)
4.1.2	**Applicable Regulations/Policies**			
4.1.2.1	Local regulations/policies	Information not available/found	Information not available/found	
4.1.2.2	Tribal regulations/policies	Information not available/found	Information not available/found	
4.1.2.3	State regulations/policies	Information not available/found	Information not available/found	
4.1.2.4	Federal regulations/policies	Information not available/found	Information not available/found	
4.1.2.5	International regulations/policies	Information not available/found	Information not available/found	
4.1.2.6	Contractual obligations	Information not available/found	Information not available/found	
4.1.2.7	Repository policies	Legal Review: PEDSnet DUA and RUD [1]	PEDSnet policy	[1] https://pedsnet.org/resources/governance/ (Accessed: 4/18/23)
4.2	**Governance for data linkage, sharing, access, and use based on data access authorization or applicable regulations/policies (i.e., the origin of the governance)**			
4.2.1	Whether the data can be linked	Does not authorize/specify	Does not authorize/specify	
4.2.2	With what other data can it be linked or can it not be linked (scope of linkage)	Does not authorize/specify	Does not authorize/specify	
4.2.3	Whether data can be shared	Does not authorize/specify	Does not authorize/specify	
4.2.4	How data can be shared (de-identification status, disclosure review)	Does not authorize/specify	Does not authorize/specify	

Dataset 9 - PEDSnet Limited Data Sets (LDS)				
		Raw Language	**Interpretation**	**Source**
4.2.5	How data can be accessed (access type, data use agreement, data access committee/group approval, IRB LOD, etc.)	Investigators or sponsors that would like to conduct a study within PEDSnet will complete a request form and submit that to the Coordinating Center. The Coordinating Center will ensure that the request is complete. The Research Committee will vote on study concepts. Approval allows the Coordinating Center to assist the requestor in the development of study proposals. [1] Human Subjects Review: IRB Determination/Approval required by PEDSnet site -If NHSR/Exempt: no further review/MRA required -If HSR: IRB approval with IRB reliance for sites providing data (NPRA MRA or SMART IRB MRA) [1] Legal Review: PEDSnet DUA and RUD [1] Network Participation Approval: Executive Committee Approval [1] Institutional Participation Approval: Prospective Site PI Approval for all data requests regardless of the de-identification status of the dataset being requested.[1] Due to its size and the corresponding charges by cloud service providers, the main PEDSnet database is located in a physical datacenter. Project-related data is sometimes located in cloud enclaves to facilitate collaborative research. [2] In general, research using the database will be done within the secure PEDSnet database environment that is managed by the Coordinating	PEDSnet human subjects review, network participation review, institutional participation approval, and legal review specify that the requester: 1. Must submit request form for approval by the Research Committee 2. Must undergo IRB review/determination (Human Subjects Review) 3. If IRB determines the proposed study is NHSR, then no further review/MRA required 4. If IRB determines the proposed study HSR, the requester must provide IRB approval with IRB reliance for site providing data (NPRA MRA or SMART IRB MRA is also required) 5. Must sign DUA and RUD (Responsible Use of Data) (Legal Review) 6. Must receive prospective site PI approval (Institutional Participation Approval) 7. Must receive PEDSnet Executive Committee approval (Network Participation Approval) 8. Must access the data through a workspace within the PEDSnet cloud enclave--OR--to have the data transfered to their institution, the PEDSnet Study	[1] https://pedsnet.org/resources/governance/ (Accessed: 4/18/23) [2] Responses to PEDSnet Questions

Dataset 9 - PEDSnet Limited Data Sets (LDS)				
		Raw Language	**Interpretation**	**Source**
		Center. This approach generally precludes the need to share patient-level data outside the secure PEDSnet network environment. Investigators, data analysts, and statisticians who need to access the database will first apply to be an authorized user. For approved PEDSnet studies, a certified user either at the Coordinating Center or PEDSnet institution will set up a workspace within the PEDSnet database environment and transfer the minimum necessary data for the research project to the workspace. The workspace will support database and statistical applications allowing the team to conduct data analyses. [1] Investigators who would like to have a de-identified or limited (as defined by HIPAA) patient-level dataset transferred to their institution make this request at the time of seeking PEDSnet Study Approval. Each institution that supplies data for the dataset must affirm its approval during the Steering Committee voting process. At each site's discretion, either the PEDSnet Site PI or Site Informatics Lead may approve dataset release votes, obtaining approval from appropriate institutional officials as needed. Approved requests will be processed by the Coordinating Center, which will provide the minimum data necessary to answer study questions. The Coordinating Center will maintain procedures to reduce risk of individual patient reidentification from datasets released to investigators. Data provided by PEDSnet can be used only for the purposes specified and approved by the Steering Committee. [1]	Approval request should specify, pending approval from all PEDSnet institutions providing data for the request.	
4.2.6	How data can be used (data use limitations)	Does not authorize/specify	Does not authorize/specify	
4.2.7	Other (specify)	Does not authorize/specify	Does not authorize/specify	
5	**Data Use**			
5.1	**Authorizations and Applicable Regulations/Policies**			
5.1.1	**Authorizations**		1. PEDSnet Steering Committee approval 2. PEDSnet site PIs written affirmation 3. PEDSnet site participation vote	
5.1.1.1	Assent	Information not available/found	Information not available/found	
5.1.1.2	Consent	Information not available/found	Information not available/found	
5.1.1.3	IRB/equivalent Privacy Board determination	Information not available/found	Information not available/found	
5.1.1.4	Local/state/federal laws	Information not available/found	Information not available/found	
5.1.1.5	Institutional Certification	Information not available/found	Information not available/found	

Dataset 9 - PEDSnet Limited Data Sets (LDS)				
		Raw Language	**Interpretation**	**Source**
5.1.1.6	Data originator agreement	Participation by an Institution in a particular PEDSnet study is voluntary. A written affirmation is required by the Site Principal Investigator for all studies. However, the expectation is that participating institutions allow their data to be used for retrospective observational studies that do not require contact with human subjects unless there is a compelling reason to not participate. [1] Each site must authorize use of their data and resources in a particular study via a participation vote. [2]	Two agreements authorize data use: 1. PEDSnet Site PI written affirmation authorizes data use for other investigators 2. PEDSnet site participation vote authorizes data use for individual studies	[1] https://pedsnet.org/documents/303/PEDSnet-Policies-May-2020_.pdf [2] Responses to PEDSnet Questions
5.1.1.7	Repository agreements/policies	Legal Review: PEDSnet DUA and RUD [1]	Legal review (comprised of PEDSnet Data use agreement and Responsible Use of Data Agreement) authorizes data use.	[1] https://pedsnet.org/resources/governance/ (Accessed: 4/18/23)
5.1.1.8	Other (specify)	Data provided by PEDSnet can be used only for the purposes specified and approved by the Steering Committee. [1]	Steering Committee authorizes data use.	[1] https://pedsnet.org/resources/governance/ (Accessed: 4/19/23)
5.1.2	**Applicable Regulations/Policies**			
5.1.2.1	Local regulations/policies	Requests for patient level data will be assessed for research concept fitness and will undergo a formal vote by the Research Committee. PEDSnet will seek to prioritize studies that will inform or directly address clinical decision making among diagnostic or treatment alternatives available to parents, patients and providers or by health care delivery systems. [1] The expectation is that participating institutions allow their data to be used for retrospective observational studies that do not require contact with human subjects unless there is a compelling reason to not participate. [1] In order to support a learning health system, PEDSnet fosters the use of data from real-world clinical settings for research, quality measurement, and improvement of child health. Use of these data make it possible to reach conclusions that more accurately reflect actual health and medical care than simulated or idealized data. [1] PEDSnet data are not available for commercial sale; other uses are reviewed for their alignment with advancement of child health. When released for a study, data use is limited to the conduct of that study; blanket releases of datasets are not made. [2]	PEDSnet policy	[1] https://pedsnet.org/resources/data-governance/ [2] Responses to PEDSnet Questions
5.1.2.2	Tribal regulations/policies	Information not available/found	Information not available/found	
5.1.2.3	State regulations/policies	Information not available/found	Information not available/found	
5.1.2.4	Federal regulations/policies	Information not available/found	Information not available/found	
5.1.2.5	International regulations/policies	Information not available/found	Information not available/found	
5.1.2.6	Contractual obligations	Information not available/found	Information not available/found	
5.1.2.7	Repository policies	Legal Review: PEDSnet DUA and RUD [1]	PEDSnet policy	[1] https://pedsnet.org/resources/governance/ (Accessed: 4/18/23)
5.2	**Governance for data linkage, sharing, access, and use based on data access authorization or applicable regulations/policies (i.e., the origin of the governance)**			

Dataset 9 - PEDSnet Limited Data Sets (LDS)				
		Raw Language	**Interpretation**	**Source**
5.2.1	Whether the data can be linked	Linkage is determined by the PEDSnet sites based on the particular use case/research study. When sites sign up to be part of the PEDSnet network, they must agree to hash their data; however, individual PEDSnet study sites can decide whether to participate in data linkages on a study-by-study basis once the PEDSnet Steering Committee approves the data linkage as part of a specific study research plan. [1] Is the decision to participate in data linking documented via the 'written affirmation from the study site PI' or how is it authorized/documented? Where the need for linkage is known at the inception of the study (preferred), it is part of the study participation vote. Where an opportunity for linkage arises during conduct of a study (e.g. new dataset becomes available for linkage), a separate authorization is obtained. [2]	Individual PEDSnet sites, through a study particpation vote, specify participation in data linkage on a study-by-study basis	[1] Predecessor report (https://www.nichd.nih.gov/sites/default/files/inline-files/NICHD_ODSS_PPRL_for_Pediatric_COVID-19_Studies_Public_Final_Report_508.pdf) [2] Responses to PEDSnet Questions
5.2.2	With what other data can it be linked or can it not be linked (scope of linkage)	Does not authorize/specify	Does not authorize/specify	
5.2.3	Whether data can be shared	Does not authorize/specify	Does not authorize/specify	
5.2.4	How data can be shared (de-identification status, disclosure review)	Does not authorize/specify	Does not authorize/specify	
5.2.5	How data can be accessed (access type, data use agreement, data access committee/group approval, IRB LOD, etc.)	Legal Review: PEDSnet DUA and RUD [1] Network Participation Approval: Executive Committee Approval [1] Institutional Participation Approval: Prospective Site PI Approval for all data requests regardless of the de-identification status of the dataset being requested.[1] Due to its size and the corresponding charges by cloud service providers, the main PEDSnet database is located in a physical datacenter. Project-related data is sometimes located in cloud enclaves to facilitate collaborative research. [2] In general, research using the database will be done within the secure PEDSnet database environment that is managed by the Coordinating Center. This approach generally precludes the need to share patient-level data outside the secure PEDSnet network environment. Investigators, data analysts, and statisticians who need to access the database will first apply to be an authorized user. For approved PEDSnet studies, a certified user either at the Coordinating Center or PEDSnet institution will set up a workspace within the PEDSnet database environment and transfer the minimum necessary data for the research project to the workspace. The workspace will support database and statistical applications allowing the team to conduct data analyses. [1] Investigators who would like to have a de-identified or limited (as defined by HIPAA) patient-level dataset transferred to their institution make this request at the time of seeking PEDSnet Study Approval. Each institution that supplies data for the dataset must affirm its approval during the	PEDSnet policy specifies that the requester: 1. Must sign DUA and RUD (Responsible Use of Data) (Legal Review) 2. Must access the data through a workspace within the PEDSnet cloud enclave--OR--to have the data transfered to their institution, the PEDSnet Study Approval request should specify, pending approval from all PEDSnet institutions providing data for the request	[1] https://pedsnet.org/resources/governance/ (Accessed: 4/18/23) [2] Responses to PEDSnet Questions

Dataset 9 - PEDSnet Limited Data Sets (LDS)				
		Raw Language	**Interpretation**	**Source**
		Steering Committee voting process. At each site's discretion, either the PEDSnet Site PI or Site Informatics Lead may approve dataset release votes, obtaining approval from appropriate institutional officials as needed. Approved requests will be processed by the Coordinating Center, which will provide the minimum data necessary to answer study questions. The Coordinating Center will maintain procedures to reduce risk of individual patient reidentification from datasets released to investigators. Data provided by PEDSnet can be used only for the purposes specified and approved by the Steering Committee. [1]		
5.2.6	How data can be used (data use limitations)	Requests for patient level data will be assessed for research concept fitness and will undergo a formal vote by the Research Committee. PEDSnet will seek to prioritize studies that will inform or directly address clinical decision making among diagnostic or treatment alternatives available to parents, patients and providers or by health care delivery systems. [1] The expectation is that participating institutions allow their data to be used for retrospective observational studies that do not require contact with human subjects unless there is a compelling reason to not participate. [1] In order to support a learning health system, PEDSnet fosters the use of data from real-world clinical settings for research, quality measurement, and improvement of child health. Use of these data make it possible to reach conclusions that more accurately reflect actual health and medical care than simulated or idealized data. [1] PEDSnet data are not available for commercial sale; other uses are reviewed for their alignment with advancement of child health. When released for a study, data use is limited to the conduct of that study; blanket releases of datasets are not made. [2] Data provided by PEDSnet can be used only for the purposes specified and approved by the Steering Committee. [1]	PEDSnet policy specifies two data use limitations: 1. Data can only be used for the purposes specified and approved by the Steering Committee. Namely, using data from real-world clinical settings for research, quality measurement, and improvement/advancement of child health, particularly studies that inform or directly address clinical decision making, including retrospective observational studies 2. Cannot be used for commercial sale	[1] https://pedsnet.org/resources/data-governance/ (Accessed: 4/19/23) [2] Responses to PEDSnet Questions
5.2.7	Other (specify)	Does not authorize/specify	Does not authorize/specify	
6	**PII Elements**			

Dataset 9 - PEDSnet Limited Data Sets (LDS)				
		Raw Language	**Interpretation**	**Source**
6.1	PII elements collected	All sites collect a variety of PII, including demographic, familial, and insurance data, for the purposes of health care treatment, payment, and operations. PII other than dates and geocodes at the census block group level are removed from PEDSnet data, with the goal of producing a limited dataset for core use and best-effort deidentification for study specific activities (*e.g.* text for NLP, chart reviews). [1] PEDSnet uses the following PII elements to generate five different hashed codes, which are all salted: 1. Last name, first name, sex, DOB: Gives best match with over 93% accuracy. 2. Last name, first three letters of first name, sex, DOB: 87% accuracy • The more the name is truncated, less accurate it is. 3. Last name, first initial, sex and DOB: overall accuracy is 62%. 4. Last name, first name, DOB and zip3: does not perform as well as sex and DOB. 5. Soundex of last name, Soundex of first name, sex, and DOB (in children populations, PEDSnet would use trigraphs instead of Soundex). [2]	PEDSnet collects first name, last name, sex, and date of birth.	[1] Responses to PEDSnet Questions [2] CHOP/PEDSnet Interview (as part of predecessor report)
6.2	PII elements holder (i.e., party that holds the PII)	Institutions contributing data to the PEDSnet Database will retain direct patient identifiers within each institution and will not share this information with the Coordinating Center except in defined study contexts. Patients will be assigned a site-level Patient Identifier that has no internal meaning. Institutions will retain the mapping between the site-level Patient Identifier and local identifiers (such as a medical record number) to enable reidentification at the local institution. 6.7.2 Reidentification All studies requiring reidentification will have Institutional Review Board oversight. Patient reidentification will be done by providing institutions with the site-level patient identifiers of interest, and the institutions will perform the reidentification. 6.7.3 Network-wide Identifiers The Coordinating Center will maintain a unique network-wide PEDSnet Patient Identifier, which will not be disclosed to sites outside defined study context, in order to maintain an honest broker role. [1]	PEDSnet sites hold the PII.	[1] https://pedsnet.org/resources/governance/ (Accessed: 4/18/23)

Dataset 9 - PEDSnet Limited Data Sets (LDS)				
		Raw Language	**Interpretation**	**Source**
6.3	Use of common data model, if any, for data collection	PEDSnet Common Data Model PEDSnet chose from the onset of the network to establish a pediatric-specific Common Data Model (PEDSnet CDM) for the storage of PEDSnet data. The use of an internal CDM allows PEDSnet to quickly add data domains or data elements needed by pediatric investigators. Two examples are age-normalized anthropomorphic measurements (height, weight, BMI percentiles based on CDC routines) and census block geocoding for location-based data queries linked to geocoded environmental data sets. The PEDSnet CDM is based on the Observational Health Data Sciences and Informatics collaborative's OMOP common data model. This CDM focuses strongly on terminology standardization, resulting in use of common standard terminologies such as SNOMED-CT, RxNorm, CPT, and LOINC for both clinical and demographic facts. The OMOP model was expanded to include the PCORnet and pediatric specific data standards, as developed by PEDSnet. The PEDSnet CDM subsumes all PCORnet elements, but addresses important data elements not yet addressed in the PCORnet-wide CDM. PCORnet Common Data Model PCORnet created the PCORnet Common Data Model, summarized below for networks in the PCORnet initiative, to facilitate the sharing of information across the wider PCORnet network. The PCORnet Common Data Model is based on the Mini-Sentinel Common Data Model, and is used for PCORnet sponsored studies and data queries. As a PCORnet network, PEDSnet participates in these queries when approved by PEDSnet governing bodies. The PEDSnet DCC has developed automated transformations from the PEDSnet CDM to the PCORnet CDM, demonstrating the feasibility of interconversion between CDM representations of similar content, and supporting participation in many collaborative research communities. [1] The PEDSnet Coordinating Center will support research done using the PCORnet Common Data Model, and will maintain a translation between the PEDSnet and the PCORnet Common Data Models. As per PCORnet policy, the PCORnet Common Data Model will be modified on an annual basis. [2]	PEDSnet CDM is based on OMOP and includes all PCORnet CDMs in addition to important data elements not yet addressed in PCORnet.	[1] https://pedsnet.org/data/common-data-model/ (Accessed: 4/18/23) [2] https://pedsnet.org/resources/governance/ (Accessed: 4/18/23)
7	**Prior Data Linkages**			
7.1	**Dataset linked with other datasets**			
7.1.1	Name of other dataset linked to this dataset	Information not available/found	Information not available/found	
7.1.2	Other dataset type (clinical, EHR, survey, claims, SDOH, etc.)	PEDSnet has used PPRL for linkage to large claims and pharmacy benefits datasets, and is currently employing linkage to mortality data as part of the RECOVER EHR ini[ti]ative. We have also pursued clear text linkage in smaller study and registry contexts where participant consent was available. Finally, clear text linkage at sites is routinely used for purposes such as identifying charts to review. [1]	1. Claims 2. Mortality	[1] Responses to PEDSnet Questions
7.1.3	Other dataset source(s)	Information not available/found	Information not available/found	

Dataset 9 - PEDSnet Limited Data Sets (LDS)				
		Raw Language	**Interpretation**	**Source**
7.1.4	Linking methodology (PPRL or non-PPRL); linkage technology	PEDSnet has used PPRL for linkage to large claims and pharmacy benefits datasets, and is currently employing linkage to mortality data as part of the RECOVER EHR iniative. We have also pursued clear text linkage in smaller study and registry contexts where participant consent was available. Finally, clear text linkage at sites is routinely used for purposes such as identifying charts to review. [1]	1. PPRL 2. PPRL 3. Non-PPRL (clear text linkage)	[1] Responses to PEDSnet Questions
7.1.5	PII elements used for the linkage	Information not available/found	Information not available/found	
7.1.6	Entity resolver (data originator or data linker or third party)	Information not available/found	Information not available/found	
7.1.7	Party performing the linkages	Information not available/found	Information not available/found	
7.1.8	Linkage quality assessment	Information not available/found	Information not available/found	
7.1.9	Linked data sharing method (linkage maps or pre-linked dataset)	Information not available/found	Information not available/found	

USE CASE 3 - GOVERNANCE INFORMATION				
Use Case 3: SARS-CoV-2 Vaccination and Asthma-Related School Absence - Does SARS-CoV-2 vaccination result in reduced asthma-related school absences at 3/6/12+ months post-vaccination?				
Dataset 10 - RADx-UP Return to School Hawaii Study (Empowering schools as community assets to mitigate the adverse impacts of COVID-19)				
	Dataset Source	RADx Data Hub		
	Dataset Source Agency	NIH		
	Dataset Type (Clinical, EHR, Survey, SDOH, etc.)	Survey, socioeconomic		
	Information Sources	Website, Interview Meeting Summary/Recording, RADx Data Hub documentation provided by ODSS, RADx consent documentation sent by RADx UP researcher, Response to RADx Data Hub Questions		
Dataset 10 - RADx-UP				
		Raw Language	**Interpretation**	**Source**
1	**Data Collection**			
1.1	**Authorizations and Applicable Regulations/Policies**			
1.1.1	**Authorizations**		1. Assent from children 2. Consent from parents 3. AHARO Health Centers/Comprehensive Health Center IRB 4. Hawaii DOE Data Governance and Analytics Branch	
1.1.1.1	Assent	First Visit: If you agree to participate, we will give you an appointment time and ask you to complete the following: 1. Survey: Before you come, we will ask that you take a survey either online or through video teleconferencing. a. We will ask you for basic information such as your name, date of birth, address, contact information, race, ethnicity, gender, language, health insurance status. Your name and contact information will enable us to provide results to you. They are also required by the Hawaii Department of Health for all those taking a COVID-19 test. b. We will ask you for information about COVID-19, including information about any symptoms, such as fever or headache, and test results. We will ask about your medical history and if you have or have not had vaccines and why. c. We will ask you for information about your health, education, knowledge, and attitude about Covid-19 infection, prevention, testing, vaccination, employment status, income, alcohol consumption, tobacco use. 2. SARS-CoV-2 Antigen Test. This test will tell us if you are currently infected with the virus that causes Covid-19. Either a trained professional will swab their nose with a special cotton swab or we will train you to do your own nasal sampling. Once the sample is done, we will place it on the antigen test card. Results will be available in about an hour, sometimes sooner. [1]	Assent from children authorizes data collection.	[1] Assent Form

Dataset 10 - RADx-UP				
		Raw Language	**Interpretation**	**Source**
1.1.1.2	Consent	The Hawaii consent form was modified after reviewing with community members. [1] First Visit: If you agree for your child to participate, we will give them an appointment and ask to complete the following information. We estimate that it will take your child about 30 minutes to complete all the steps. 1. Survey: Before your child comes, we will ask that they take a survey either online or in person. a. We will ask for your child's name, date of birth, address, contact information, race, ethnicity, gender, language. Their name and preferred contact method will enable us to provide results to you. These are also required by the Hawaii Department of Health for anyone taking a COVID-19 test. b. We will ask questions about COVID-19, including any symptoms, such as fever or headache. We will ask about their medical history and if your child has or not had the COVID vaccines. c. We will ask questions about your child's overall health, school grade, and a few questions about their thoughts about Covid-19 infection, prevention, testing, and vaccination. 2. SARS-CoV-2 Antigen Test: This test will tell us if your child is currently infected with the virus that causes Covid-19. Either a trained professional will swab their nose with a special cotton swab or we will train your child to do their own nasal sampling. We will provide instructions on how to take the sample and will be present to answer questions. Once the sample is done, we will place it on the antigen test card. Results will be available in about 30 minutes, often sooner. Retesting: Sometimes we will need to retest a person on the same day to make sure the results are accurate. a. All positive tests will be confirmed with a PCR nasal swab test to make sure the results are correct. This is a very good method to detect infection but it usually takes 1-2 days to process. b. If your child's first test is NEGATIVE but have had serious exposure to Covid-19, or have developed symptoms, we may also ask for a second test. c. To make sure everyone is safe during this waiting period, we will ask that your child leaves school and quarantine at home. We will call you as soon as the results are available. [2]	Consent from parents authorizes data collection.	[1] RADx UP Meeting 1 [2] Parental Consent Form for Students
1.1.1.3	IRB/equivalent Privacy Board determination	The Hawaii study is under the IRB of the main community health center [AHARO Health Centers] that has oversight over the study. The community Comprehensive Health Center IRB is a more strict IRB than the typical University of Hawaii IRB. [1, 2]	AHARO Health Centers/Comprehensive Health Center IRB authorizes data collection.	[1] RADx UP Meeting 1 [2] RADx UP Meeting 2
1.1.1.4	Local/state/federal laws	The Hawaii Department of Education (DOE) Data Governance and Analytics Branch department had to approve since the study was based at schools. The only data we gathered directly was testing data for the students, teachers, staff who underwent school-based testing. [1, 2]	Hawaii DOE Data Governance and Analytics Branch authorizes data collection.	[1] RADx UP Meeting 1 [2] RADx UP Meeting 2
1.1.1.5	Institutional Certification	Information not available/found	Information not available/found	
1.1.1.6	Data originator agreement	Information not available/found	Information not available/found	
1.1.1.7	Repository agreements/policies	Information not available/found	Information not available/found	
1.1.1.8	Other (specify)	Information not available/found	Information not available/found	
1.1.2	**Applicable Regulations/Policies**			
1.1.2.1	Local regulations/policies	Information not available/found	Information not available/found	
1.1.2.2	Tribal regulations/policies	Information not available/found	Information not available/found	
1.1.2.3	State regulations/policies	Information not available/found	Information not available/found	
1.1.2.4	Federal regulations/policies	Information not available/found	Information not available/found	
1.1.2.5	International regulations/policies	Information not available/found	Information not available/found	
1.1.2.6	Contractual obligations	Information not available/found	Information not available/found	
1.1.2.7	Repository policies	Information not available/found	Information not available/found	
1.2	**Governance for data linkage, sharing, access, and use based on data collection authorization or applicable regulations/policies (i.e., the origin of the governance)**			

Dataset 10 - RADx-UP				
		Raw Language	**Interpretation**	**Source**
1.2.1	Whether the data can be linked	1. The project is funded by the National Institutes of Health (NIH). It is part of the NIH RADx-UP, a health research program to learn more about Covid-19 disease. If you join RADx-UP, we will gather some data (information) about your child. We will combine the data from all who join, to understand how to help more people at risk for or with COVID-19. [1] 2. The Hawaii study consent form contains language that the data will be used to participate in the overall cohort for RADx-UP but does not explicitly mention linking to other datasets. “Your data can be used for other research studies” language leaves the option for the data to be used for linkage. [2] 3. Linkage is possible but any linkage at an individual level must be approved by the Comprehensive Health Center IRB. Anything beyond [general research purposes] (such as linking data and working with identifiers) must be approved by the community IRB. [2]	1. Parental informed consent and assent do not specify linkage. Raw language referring to "other research studies" is interpreted by the study PI as leaving the option open for data linkage. 2. AHARO Health Centers/Comprehensive Health Center IRB specifies that data can be linked at an individual level only if the IRB approves the linkage.	[1] Parental Consent Form for Students [2] RADx UP Meeting 2
1.2.2	With what other data can it be linked or can it not be linked (scope of linkage)	Linkage is possible but any linkage at an individual level must be approved by the Comprehensive Health Center IRB. Anything beyond [general research purposes] (such as linking data and working with identifiers) must be approved by the community IRB. [1]	AHARO Health Centers/Comprehensive Health Center IRB specifies that any data linkages at an individual level or outside of general research purposes must be approved by the AHARO Health Centers IRB.	[1] RADx UP Meeting 2
1.2.3	Whether data can be shared	The project is funded by the National Institutes of Health (NIH). It is part of the NIH RADx-UP, a health research program to learn more about Covid-19 disease. If you join RADx-UP, we will gather some data (information) about your child. We will combine the data from all who join, to understand how to help more people at risk for or with COVID-19. [1] When the data is shared with other researchers, they will not have information that can identify your child. [1]	Parental informed consent specifies that data will be shared with other researchers.	[1] Parental Consent Form for Students
1.2.4	How data can be shared (de-identification status, disclosure review)	Does not authorize/specify	Does not authorize/specify	
1.2.5	How data can be accessed (access type, data use agreement, data access committee/group approval, IRB LOD, etc.)	Does not authorize/specify	Does not authorize/specify	
1.2.6	How data can be used (data use limitations)	Does not authorize/specify	Does not authorize/specify	
1.2.7	Other (specify)	Does not authorize/specify	Does not authorize/specify	
2	**Data Linkage**			
2.1	**Authorizations and Applicable Regulations/Policies**			
2.1.1	**Authorizations**		1. AHARO Health Centers/Comprehensive Health Center IRB	
2.1.1.1	Assent	Does not authorize/specify The project is funded by the National Institutes of Health (NIH). It is part of the NIH RADx-UP, a health research program to learn more about Covid-19 disease. If you join RADx-UP, we will gather some data (information) about your child. We will combine the data from all who join, to understand how to help more people at risk for or with COVID-19. [1] The Hawaii study consent form contains language that the data will be used to participate in the overall cohort for RADx-UP but does not explicitly mention linking to other datasets. “Your data can be used for other research studies” language leaves the option for the data to be used for linkage. [2]	Does not authorize/ specify	[1] Assent Form [2] RADx UP Meeting 2

Dataset 10 - RADx-UP				
		Raw Language	**Interpretation**	**Source**
2.1.1.2	Consent	Does not authorize/specify The project is funded by the National Institutes of Health (NIH). It is part of the NIH RADx-UP, a health research program to learn more about Covid-19 disease. If you join RADx-UP, we will gather some data (information) about your child. We will combine the data from all who join, to understand how to help more people at risk for or with COVID-19. [1] The Hawaii study consent form contains language that the data will be used to participate in the overall cohort for RADx-UP but does not explicitly mention linking to other datasets. “Your data can be used for other research studies” language leaves the option for the data to be used for linkage. [2]	Does not authorize/ specify	[1] Parental Consent Form for Students [2] RADx UP Meeting 2
2.1.1.3	IRB/equivalent Privacy Board determination	Linkage is possible but any linkage at an individual level must be approved by the Comprehensive Health Center IRB [AHARO Health Center]. Anything beyond [general research purposes] (such as linking data and working with identifiers) must be approved by the community IRB. [1]	AHARO Health Centers/Comprehensive Health Center IRB authorizes data linkage.	[1] RADx UP Meeting 2
2.1.1.4	Local/state/federal laws	Information not available/found	Information not available/found	
2.1.1.5	Institutional Certification	Information not available/found	Information not available/found	
2.1.1.6	Data originator agreement	Information not available/found	Information not available/found	
2.1.1.7	Repository agreements/policies	Information not available/found	Information not available/found	
2.1.1.8	Other (specify)	Information not available/found	Information not available/found	
2.1.2	**Applicable Regulations/Policies**			
2.1.2.1	Local regulations/policies	Information not available/found	Information not available/found	
2.1.2.2	Tribal regulations/policies	Information not available/found	Information not available/found	
2.1.2.3	State regulations/policies	Information not available/found	Information not available/found	
2.1.2.4	Federal regulations/policies	Information not available/found	Information not available/found	
2.1.2.5	International regulations/policies	Information not available/found	Information not available/found	
2.1.2.6	Contractual obligations	Information not available/found	Information not available/found	
2.1.2.7	Repository policies	Information not available/found	Information not available/found	
2.2	**Governance for data linkage, sharing, access, and use based on data linkage authorization or applicable regulations/policies (i.e., the origin of the governance)**			
2.2.1	Whether the data can be linked	1. The project is funded by the National Institutes of Health (NIH). It is part of the NIH RADx-UP, a health research program to learn more about Covid-19 disease. If you join RADx-UP, we will gather some data (information) about your child. We will combine the data from all who join, to understand how to help more people at risk for or with COVID-19. [1] 2. The Hawaii study consent form contains language that the data will be used to participate in the overall cohort for RADx-UP but does not explicitly mention linking to other datasets. “Your data can be used for other research studies” language leaves the option for the data to be used for linkage. [2] 3. Linkage is possible but any linkage at an individual level must be approved by the Comprehensive Health Center IRB. Anything beyond [general research purposes] (such as linking data and working with identifiers) must be approved by the community IRB. [2]	1. Parental informed consent and assent do not specify linkage. Raw language refering to "other research studies" is interpreted by the study PI as leaving the option open for data linkage. 2. AHARO Health Centers/Comprehensive Health Center IRB specifices that data can be linked at an individual level only if the IRB approves the linkage.	[1] Parental Consent Form for Students [2] RADx UP Meeting 2
2.2.2	With what other data can it be linked or can it not be linked (scope of linkage)	Linkage is possible but any linkage at an individual level must be approved by the Comprehensive Health Center IRB. Anything beyond [general research purposes] (such as linking data and working with identifiers) must be approved by the community IRB. [1]	AHARO Health Centers/Comprehensive Health Center IRB specifies that any data linkages at an individual level or outside of general research purposes must be approved by the AHARO Health Centers IRB.	[1] RADx UP Meeting 2
2.2.3	Whether data can be shared	The project is funded by the National Institutes of Health (NIH). It is part of the NIH RADx-UP, a health research program to learn more about Covid-19 disease. If you join RADx-UP, we will gather some data (information) about your child. We will combine the data from all who join, to understand how to help more people at risk for or with COVID-19. [1] When the data is shared with other researchers, they will not have information that can identify your child. [1]	Parental informed consent specifies that data will be shared with other researchers.	[1] Parental Consent Form for Students

Dataset 10 - RADx-UP				
		Raw Language	**Interpretation**	**Source**
2.2.4	How data can be shared (de-identification status, disclosure review)	Does not authorize/specify	Does not authorize/specify	
2.2.5	How data can be accessed (access type, data use agreement, data access committee/group approval, IRB LOD, etc.)	Does not authorize/specify	Does not authorize/specify	
2.2.6	How data can be used (data use limitations)	Does not authorize/specify	Does not authorize/specify	
2.2.7	Other (specify)	Does not authorize/specify	Does not authorize/specify	
3	**Data Sharing**			
3.1	**Authorizations and Applicable Regulations/Policies**			
3.1.1	**Authorizations**		1. Assent from children 2. Consent from parents 3. FERPA 4. Data originator agreement (Agreement For Disclosure And Transfer Of Confidential Information And Protected Health Information) 5. Registration of studies in dbGaP 6. RADx Institutional Certification	
3.1.1.1	Assent	It is part of the NIH RADx-UP, a health research program to learn more about Covid-19 disease. If you join RADx-UP, we will gather some data (information) about you. We will combine these with data from other people who join RADx-UP. We will study the data from all who join to understand how to help more people at risk for or with COVID-19. [1] When the data is shared with other researchers, they will not have information that can identify you. [1]	Assent from children authorizes data sharing.	[1] Assent Form
3.1.1.2	Consent	The project is funded by the National Institutes of Health (NIH). It is part of the NIH RADx-UP, a health research program to learn more about Covid-19 disease. If you join RADx-UP, we will gather some data (information) about your child. We will combine the data from all who join, to understand how to help more people at risk for or with COVID-19. [1] When the data is shared with other researchers, they will not have information that can identify your child. [1]	Consent from parents authorizes data sharing.	[1] Parental Consent Form for Students
3.1.1.3	IRB/equivalent Privacy Board determination	Information not available/found	Information not available/found	
3.1.1.4	Local/state/federal laws	The data was collected indirectly from schools, so FERPA applies to the data collected. [1]	FERPA authorizes data sharing.	[1] RADx UP Meeting 2

Dataset 10 - RADx-UP				
		Raw Language	**Interpretation**	**Source**
3.1.1.5	Institutional Certification	The _____ hereby assures that submission of data from the study entitled _____ to an NIH-designated data repository meets the following expectations: • The data submission is consistent, as appropriate, with applicable national, tribal, and state laws and regulations as well as relevant institutional policies. • Any limitations on the research use of the data, as expressed in the informed consent documents, are delineated in the table on page 3. • The identities of research participants will not be disclosed to NIH-designated data repositories. • An Institutional Review Board (IRB), and/or Privacy Board, and/or equivalent body, and a relevant senior-level institutional staff (e.g., Dean, Vice President/Provost for Research, Chief Science Officer) as applicable, has reviewed the investigator's proposal for data submission and assures that: o The protocol for the collection of genomic and phenotypic data is consistent with 45 CFR Part 46;2 o Data submission and subsequent data sharing for research purposes are consistent with the informed consent of study participants from whom the data were obtained; o Consideration was given to risks to individual participants and their families associated with data submitted to NIH-designated data repositories and subsequent sharing, including unrestricted access to genomic summary results; o To the extent relevant and possible,consideration was given to risks to groups or populations associated with submitting data to NIH-designated data repositories and subsequent sharing, including unrestricted access to genomic summary results; and o The investigator's plan for de-identifying datasets is consistent with the HHS Regulations for the Protection of Human Subjects** [1]	RADx Institutional Certification authorizes data sharing.	[1] Rapid Acceleration for Diagnostics Program (RADx) Institutional Certification: https://radx-up.org/wp-content/uploads/2022/05/RADx-INSTITUTIONAL-CERTIFICATION-FINAL-02022021-1.pdf (Accessed 4/17/23)
3.1.1.6	Data originator agreement	Discloser wishes to share with Duke, or Duke will be given access to (i) research subjects' information that constitutes Protected Health Information ("PHI"), as defined in the Health Insurance Portability and Accountability Act of 1996, as amended ("HIPAA"), including, but not limited to, research subjects' contact information and (ii) other information that Discloser considers to be confidential ("Discloser Confidential Information") to enable Duke to: (a) obtain research subjects' written informed consent/HIPAA authorization ("Subject Authorization"), RADx-UP Common Data Elements, related questionnaires, surveys and forms for performing data analyses and for collecting follow up data from research subjects; (b) better understand COVID-19 testing patterns among underserved and vulnerable populations; (c) strengthen the understanding of the impact of relevant data on disparities in infection rates, disease progression, and outcomes; (d) develop strategies to reduce disparities in COVID-19 testing; and (e) fulfill Duke's obligation as the CDCC under the Project to provide de-identified Project data and the results of its analyses to the Awarding Agency (collectively, "Duke Purpose") [1]	Agreement For Disclosure And Transfer Of Confidential Information And Protected Health Information authorizes data sharing.	[1] Agreement For Disclosure And Transfer Of Confidential Information And Protected Health Information: https://radx-up.org/research/cdes/ (Accessed 4/17/23)
3.1.1.7	Repository agreements/policies	Information not available/found	Information not available/found	
3.1.1.8	Other	Information not available/found	Information not available/found	
3.1.2	**Applicable Regulations/Policies**			
3.1.2.1	Local regulations/policies	Information not available/found	Information not available/found	
3.1.2.2	Tribal regulations/policies	Information not available/found	Information not available/found	
3.1.2.3	State regulations/policies	Information not available/found	Information not available/found	
3.1.2.4	Federal regulations/policies	Information not available/found	Information not available/found	
3.1.2.5	International regulations/policies	Information not available/found	Information not available/found	
3.1.2.6	Contractual obligations	Information not available/found	Information not available/found	
3.1.2.7	Repository policies	Every study loaded into the Data Hub must be registered in the dbGaP system first. The study registration process starts when the Principal Investigator (PI) emails the dbGaP Registration Form to the Genomic Program Administrator (GPA) [1] approver and the appropriate C(DCC) Administrator. Using the information from the form, the GPA enters the study registration information into dbGaP. Once the study has been successfully registered in dbGaP, the GPA will send the dbGaP Registration Form (PDF file) to the Super User. [2]	RADx Data Hub policy	[1] RADx Data Hub Email Communication [2] https://radx-hub.nih.gov/docs/dcc/create-study-page.html (Accessed 4/17/23)
3.2	**Governance for data linkage, sharing, access, and use based on data sharing authorization or applicable regulations/policies (i.e., the origin of the governance)**			
3.2.1	Whether the data can be linked	1. The project is funded by the National Institutes of Health (NIH). It is part of the NIH RADx-UP, a health research program to learn more about Covid-19 disease. If you join RADx-UP, we will gather some data (information) about your child. We will combine the data from all who join, to understand how to help more people at risk for or with COVID-19. [1] 2. The Hawaii study consent form contains language that the data will be used to participate in the overall cohort for RADx-UP but does not explicitly mention linking to other datasets. "Your data can be used for other research studies" language leaves the option for the data to be used for linkage. [2]	Parental informed consent and assent do not specify linkage. Raw language refering to "other research studies" is interpreted by the study PI as leaving the option open for data linkage.	[1] Parental Consent Form for Students [2] RADx UP Meeting 2

Dataset 10 - RADx-UP				
		Raw Language	**Interpretation**	**Source**
3.2.2	With what other data can it be linked or can it not be linked (scope of linkage)	Does not authorize/specify	Does not authorize/specify	
3.2.3	Whether data can be shared	When the data is shared with other researchers, they will not have information that can identify your child. [1] The data was collected indirectly from schools, so FERPA applies to the data collected. [2] Fulfill Duke's obligation as the CDCC under the Project to provide de-identified Project data and the results of its analyses to the Awarding Agency (collectively, "Duke Purpose") [3] Every study loaded into the Data Hub must be registered in the dbGaP system first. [4] The _____ hereby assures that submission of data from the study entitled _____ to an NIH-designated data repository meets the following expectations: • The data submission is consistent, as appropriate, with applicable national, tribal, and state laws and regulations as well as relevant institutional policies. • Any limitations on the research use of the data, as expressed in the informed consent documents, are delineated in the table on page 3. • The identities of research participants will not be disclosed to NIH-designated data repositories. • An Institutional Review Board (IRB), and/or Privacy Board, and/or equivalent body, and a relevant senior-level institutional staff (e.g., Dean, Vice President/Provost for Research, Chief Science Officer) as applicable, has reviewed the investigator's proposal for data submission and assures that: o The protocol for the collection of genomic and phenotypic data is consistent with 45 CFR Part 46;2 o Data submission and subsequent data sharing for research purposes are consistent with the informed consent of study participants from whom the data were obtained; o Consideration was given to risks to individual participants and their families associated with data submitted to NIH-designated data repositories and subsequent sharing, including unrestricted access to genomic summary results; o To the extent relevant and possible,consideration was given to risks to groups or populations associated with submitting data to NIH-designated data repositories and subsequent sharing, including unrestricted access to genomic summary results; and o The investigator's plan for de-identifying datasets is consistent with the HHS Regulations for the Protection of Human Subjects** [5]	The following specify that data can be shared: 1. Parental informed consent language 2. FERPA 3. Agreement For Disclosure And Transfer Of Confidential Information And Protected Health Information (study data, including PHI, to be sent to Duke (CDCC) and for Duke to provide de-identified project data for the awarding agency.) 4. Study registration in dbGaP 5. RADx Institutional Certification (in the RADx Data Hub by the CDCC/Duke)	[1] Parental Consent Form for Students [2] RADx UP Meeting 2 [3] Agreement For Disclosure And Transfer Of Confidential Information And Protected Health Information: https://radx-up.org/research/cdes/ (Accessed 4/17/23) [4] https://radx-hub.nih.gov/docs/dcc/create-study-page.html (Accessed 4/17/23) [5] Institutional Certification: https://radx-up.org/wp-content/uploads/2022/05/RADx-INSTITUTIONAL-CERTIFICATION-FINAL-02022021-1.pdf (Accessed 4/17/23)

Dataset 10 - RADx-UP				
		Raw Language	**Interpretation**	**Source**
3.2.4	How data can be shared (de-identification status, disclosure review)	Please confirm our understanding that all data shared in RADx Data Hub are de-identified. "Correct" [1] Every study loaded into the Data Hub must be registered in the dbGaP system first [2] The identities of research participants will not be disclosed to NIH-designated data repositories. [3] Zip codes. The RADxSM DCC will send awardees software to de-identify zip code data consistent with Section 164.514(a) of the Health Information Portability and Accountability Act (HIPAA) Privacy Rule to ensure the coordination centers receive HIPAA de-identified zip code data. All geographic subdivisions smaller than a state, including street address, city, county, precinct, ZIP code, and their equivalent geocodes, except for the initial three digits of the ZIP code if, according to the current publicly available data from the Bureau of the Census: (1) The geographic unit formed by combining all ZIP codes with the same three initial digits contains more than 20,000 people; and (2) The initial three digits of a ZIP code for all such geographic units containing 20,000 or fewer people is changed to 000 Date shifting. The RADxSM DCC will send the awardees software to de-identify dates by using a randomly selected, constant # of days (e.g., add 4 days to all dates) to ensure that date data submitted to the coordination centers are de-identified; awardees will keep a record of these data interval constants for each participant used for all data submissions. The process for randomly choosing the number of offset days will be discussed among the RADx coordination centers to ensure consistency across RADxSM programs. Preliminary discussions suggested the time interval should be 10 days (e.g., if the interval is -5 to 5 days from day 0, the original date, one research participant may have all their dates shifted 3 days forward while another research participant may have their dates shifted 5 days back). Age. For age, if a research participant is less than one year old, the submitted data will be listed as “0”; if a participant is 21 years old or above, the submitted age will be listed within +/- 2 years of the chronological age. In cases where the research participant is 90 years old or above, the submitted age element will be aggregated into a single number “90”. The RADxSM Data Hub will alert Awardees about these two age groups. Note: For projects that are generating data with no HIPAA-de-identifiers, awardees’ projects may require modifications to their IRB protocols (including consent form modifications), and require modifications to , Data Use and/or Data Transfer agreements, depending on the identifiers contained in the data. Additionally, a system for research participant notification of data breaches may be required depending on the identifier. [4]	1. RADx policy specifies that the study be registered in dbGaP prior to sharing through RADx Data Hub. 2. The RADx Institutional Certification specifies that all data shared in an NIH designated repository must be de-identified. 3. The RADx DCC works with study teams to de-identify zip codes, shift dates, and adjust ages into categories for specific ages.	[1] RADx Data Hub Email Communication [2] https://radx-hub.nih.gov/docs/dcc/create-study-page.html (Accessed 4/17/23) [3] Rapid Acceleration for Diagnostics Program (RADx) Institutional Certification: https://radx-up.org/wp-content/uploads/2022/05/RADx-INSTITUTIONAL-CERTIFICATION-FINAL-02022021-1.pdf (Accessed 4/17/23) [4] Proposal for RADxSM (C)DCC Data Sharing with the RADx Data Hub
3.2.5	How data can be accessed (access type, data use agreement, data access committee/group approval, IRB LOD, etc.)	All data in the RADx Data Hub are considered controlled access. [1] The individual-level data are to be made available through (check one) -controlled-access -unrestricted access [2]	RADx Institutional Certification specifies that all individual-level data are controlled access.	[1] Responses to RADx Data Hub Questions [2] Rapid Acceleration for Diagnostics Program (RADx) Institutional Certification: https://radx-up.org/wp-content/uploads/2022/05/RADx-INSTITUTIONAL-CERTIFICATION-FINAL-02022021-1.pdf (Accessed 4/17/23)

Dataset 10 - RADx-UP				
		Raw Language	**Interpretation**	**Source**
3.2.6	How data can be used (data use limitations)	NIH expects the submitting institution(s) to select one of the three standard Data Use Limitations (DULs) for appropriate secondary use, or, if necessary, create a customized DUL. DULs are developed based on the original informed consent of the participant(s). -General Research Use (GRU): Use of the data is s limited only by the terms of the Data Use Certification: these data will be added to the dbGaP Collection. -Health/Medical/Biomedical (HMB): Use of the data is limited to health/medical/biomedical purposes, does not include the study of population origins or ancestry -Disease-specific [list disease] (DS): Use of the data must be related to the specified disease -Other Additional modifiers to the standard DULs (e.g., Not-for-profit Use Only) can be indicated, if appropriate. Use of the modifiers should have a basis in the informed consent from the participants or in special knowledge of the preferences of the original study population. -IRB Approval Required (IRB): Requestor must provide documentation of local IRB approval -Publication Required (PUB): Requestor agrees to make results of studies using the data available to the larger scientific community. -Collaboration Required (COL): Requestor must provide a letter of collaboration with the primary study investigator(s). -Not-for-profit Use Only (NPU): Use of the data is limited to not-for-profit organizations -Methods (MDS): Use of the data includes methods development research (e.g., development and testing of software or algorithms) -Genetic Studies Only (GSO): Use of the data is limited to genetic studies only. [1] The data that was shared with Duke can be used for general research purposes. [2]	RADx Institutional Certification specifies that the data can be used for general research purposes.	[1] Rapid Acceleration for Diagnostics Program (RADx) Institutional Certification: https://radx-up.org/wp-content/uploads/2022/05/RADx-INSTITUTIONAL-CERTIFICATION-FINAL-02022021-1.pdf (Accessed 4/17/23) [2] RADx UP Meeting 2
3.2.7	Other (specify)	Does not authorize/specify	Does not authorize/specify	
4	**Data Access**			
4.1	**Authorizations and Applicable Regulations/Policies**			
4.1.1	**Authorizations**		Information not available/found	
4.1.1.1	Assent	Does not authorize/specify	Does not authorize/specify	
4.1.1.2	Consent	Does not authorize/specify	Does not authorize/specify	
4.1.1.3	IRB/equivalent Privacy Board determination	Information not available/found -- not all RADx-UP studies require IRB to access. Since the Return to School Hawaii study has not been submitted to the RADx Data Hub, information is not available on whether and IRB/equivalen Privacy Board determination is required for data access. Additional modifiers to the standard DULs (e.g., Not-for-profit Use Only) can be indicated, if appropriate. Use of the modifiers should have a basis in the informed consent from the participants or in special knowledge of the preferences of the original study population. Data Use Limitation Modifiers (Optional) -IRB Approval Required (IRB): Requestor must provide documentation of local IRB approval [1]	Information not available/found	[1] Rapid Acceleration for Diagnostics Program (RADx) Institutional Certification: https://radx-up.org/wp-content/uploads/2022/05/RADx-INSTITUTIONAL-CERTIFICATION-FINAL-02022021-1.pdf (Accessed 4/17/23)
4.1.1.4	Local/state/federal laws	Information not available/found	Information not available/found	
4.1.1.5	Institutional Certification	Information not available/found	Information not available/found	
4.1.1.6	Data originator agreement	Information not available/found	Information not available/found	

Dataset 10 - RADx-UP				
		Raw Language	**Interpretation**	**Source**
4.1.1.7	Repository agreements/policies	Investigators and their institutions are responsible for safeguarding the accessed datasets. They must also adhere to the terms of access described in the Data Use Certification (DUC) Agreement and the Genomic Data User Code of Conduct that were signed as part of the data access request. [1] RADxSM Data User Code of Conduct: Key principles and practices agreed to by all research investigators requesting access to RADxSM data. The elements within RADxSM Data User Code of Conduct reflect the terms of access in the Data Use Certification Agreement. Failure to abide by the RADxSM Data User Code of Conduct may result in revocation of an investigator's access to any and all approved datasets. [2] Must researchers sign the Code of Conduct to access RADx-UP data in the RADx Data Hub and if so, when is it signed? [A]t the time of the Data Access Request in submitted into the dbGaP system and attested to by the institutional signing official. [3]	Three repository agreements authorize data access: 1. RADx Data Use Certification (DUC) Agreement 2. Genomic Data Use Coder of Conduct 3. RADxSM Data User Code of Conduct	[1] https://sharing.nih.gov/accessing-data/accessing-genomic-data/how-to-request-and-access-datasets-from-dbgap (Accessed 4/17/23) [2] RADx Data Use Certification Agreement [3] Responses to RADx Data Hub Questions
4.1.1.8	Other (specify)	Information not available/found	Information not available/found	
4.1.2	**Applicable Regulations/Policies**			
4.1.2.1	Local regulations/policies	Information not available/found	Information not available/found	
4.1.2.2	Tribal regulations/policies	Information not available/found	Information not available/found	
4.1.2.3	State regulations/policies	Information not available/found	Information not available/found	
4.1.2.4	Federal regulations/policies	Information not available/found	Information not available/found	
4.1.2.5	International regulations/policies	Information not available/found	Information not available/found	
4.1.2.6	Contractual obligations	Information not available/found	Information not available/found	
4.1.2.7	Repository policies	Investigators and their institutions are responsible for safeguarding the accessed datasets. They must also adhere to the terms of access described in the Data Use Certification (DUC) Agreement and the Genomic Data User Code of Conduct that were signed as part of the data access request. [1] RADxSM Data User Code of Conduct: Key principles and practices agreed to by all research investigators requesting access to RADxSM data. The elements within RADxSM Data User Code of Conduct reflect the terms of access in the Data Use Certification Agreement. Failure to abide by the RADxSM Data User Code of Conduct may result in revocation of an investigator's access to any and all approved datasets. [2] Must researchers sign the Code of Conduct to access RADx-UP data in the RADx Data Hub and if so, when is it signed? [A]t the time of the Data Access Request in submitted into the dbGaP system and attested to by the institutional signing official. [3]	RADx Data Hub policy	[1] https://sharing.nih.gov/accessing-data/accessing-genomic-data/how-to-request-and-access-datasets-from-dbgap (Accessed 4/17/23) [2] RADx Data Use Certification Agreement [3] Responses to RADx Data Hub Questions
4.2	**Governance for data linkage, sharing, access, and use based on data access authorization or applicable regulations/policies (i.e., the origin of the governance)**			
4.2.1	Whether the data can be linked	Does not authorize/specify	Does not authorize/specify	
4.2.2	With what other data can it be linked or can it not be linked (scope of linkage)	Does not authorize/specify	Does not authorize/specify	
4.2.3	Whether data can be shared	Does not authorize/specify	Does not authorize/specify	
4.2.4	How data can be shared (de-identification status, disclosure review)	Does not authorize/specify	Does not authorize/specify	

Dataset 10 - RADx-UP				
		Raw Language	**Interpretation**	**Source**
4.2.5	How data can be accessed (access type, data use agreement, data access committee/group approval, IRB LOD, etc.)	1) To create a RADx Data Hub account users must have an eRA commons or Login.gov account. Any email may be used to create a Login.gov account. All users can view and use tutorial data, but study data access is controlled on a study-by-study basis. To view or use a particular study's data, you must have dbGaP access to that study and log in with your eRA commons account. [1] 2) Invesitgator Submits Data Access Request in dbGaP (Those eligible to apply for data access must be a permanent employee of their institution and either: At a level equivalent to a tenure-track professor, or Senior scientist with responsibilities that may include laboratory or research program administration and oversight.) 3) Signing Official Reviews, Approves, and Co-signs Request Once a data access request is submitted in dbGaP, it is automatically routed to the investigator's institutional Signing Official (SO) for review, approval, and co-signature. 4) Data Access Committee Reviews Request Once the SO approves and signs the data access request, the system forwards the data access request to the appropriate NIH Data Access Committee(s) (DAC) for review. The DACs review all access requests for human genomic datasets distributed through dbGaP to ensure the data access request complies with the NIH GDS policy, any IC or program-specific requirements, and any limitations or restrictions on data use. Each dataset has an associated DAC. Based on their review, DACs approve or reject data access requests, or return data access requests for revision. The DAC may ask for more information from the investigator requesting data if the proposed research is inconsistent with any applicable data use limitations, if the research intent is unclear, or if there is concern about potential harm (e.g., stigmatization) to groups or populations. 5) Access Data Investigators and their institutions are responsible for safeguarding the accessed datasets. They must also adhere to the terms of access described in the Data Use Certification (DUC) Agreement and the Genomic Data User Code of Conduct that were signed as part of the data access request. Once approved, investigators will be able to access the data for one year. Prior to the expiration of the one-year access period, investigators must submit a project renewal or close-out report, describing the progress made on the approved research project. [2] All data in the RADx Data Hub are considered controlled access. [3]	The dbGaP data access request process specifies RADx Data Hub data access. The requirements/steps specify that the user/eligible investigator: 1) Must have an eRA commons or Login.gov account 2) Must submit a Data Access Request (DAR). Data Use Certification (DUC) Agreement, the Genomic Data User Code of Conduct, and the RADx SM Data User Code of Conduct are signed as part of the DAR process 3) Must ensure that the Signing Official from investigator's institution reviews, approves, and co-signs the request 4) Must receive approval from Data Access Committee 5) Must access the controlled access data through RADx Data Hub Jupyter Notebooks	[1] https://www.radxdatahub.info/faqs (Accessed: 4/17/23) [2] https://sharing.nih.gov/accessing-data/accessing-genomic-data/using-genomic-data-responsibly (Accessed: 4/17/23) [3] Responses to RADx Data Hub Questions
4.2.6	How data can be used (data use limitations)	Does not authorize/specify	Does not authorize/specify	
4.2.7	Other (specify)	Does not authorize/specify	Does not authorize/specify	
5	**Data Use**			
5.1	**Authorizations and Applicable Regulations/Policies**			
5.1.1	**Authorizations**		1. Institutional Certification	
5.1.1.1	Assent	Does not authorize/specify	Does not authorize/specify	
5.1.1.2	Consent	Does not authorize/specify	Does not authorize/specify	
5.1.1.3	IRB/equivalent Privacy Board determination	Does not authorize/specify	Does not authorize/specify	
5.1.1.4	Local/state/federal laws	Does not authorize/specify	Does not authorize/specify	

Dataset 10 - RADx-UP				
		Raw Language	**Interpretation**	**Source**
5.1.1.5	Institutional Certification	NIH expects the submitting institution(s) to select one of the three standard Data Use Limitations (DULs) for appropriate secondary use, or, if necessary, create a customized DUL. DULs are developed based on the original informed consent of the participant(s). -General Research Use (GRU): Use of the data is limited only by the terms of the Data Use Certification: these data will be added to the dbGaP Collection. -Health/Medical/Biomedical (HMB): Use of the data is limited to health/medical/biomedical purposes, does not include the study of population origins or ancestry -Disease-specific [list disease] (DS): Use of the data must be related to the specified disease -Other Additional modifiers to the standard DULs (e.g., Not-for-profit Use Only) can be indicated, if appropriate. Use of the modifiers should have a basis in the informed consent from the participants or in special knowledge of the preferences of the original study population. -IRB Approval Required (IRB): Requestor must provide documentation of local IRB approval -Publication Required (PUB): Requestor agrees to make results of studies using the data available to the larger scientific community. -Collaboration Required (COL): Requestor must provide a letter of collaboration with the primary study investigator(s). -Not-for-profit Use Only (NPU): Use of the data is limited to not-for-profit organizations -Methods (MDS): Use of the data includes methods development research (e.g., development and testing of software or algorithms) -Genetic Studies Only (GSO): Use of the data is limited to genetic studies only. [1] The data that was shared with Duke can be used for general research purposes. [2]	The RADx Institutional Certification authorizes data use.	[1] Rapid Acceleration for Diagnostics Program (RADx) Institutional Certification: https://radx-up.org/wp-content/uploads/2022/05/RADx-INSTITUTIONAL-CERTIFICATION-FINAL-02022021-1.pdf (Accessed 4/17/23) [2] RADx UP Meeting 2
5.1.1.6	Data originator agreement	Does not authorize/specify	Does not authorize/specify	
5.1.1.7	Repository agreements/policies	This Data Use Certification Agreement outlines the terms of use for requested RADx datasets maintained in the RADx Data Hub. [1]	RADx Data Use Certification Agreement authorizes data use.	[1] RADx Data Use Certification Agreement
5.1.1.8	Other (specify)	Information not available/found	Information not available/found	
5.1.2	**Applicable Regulations/Policies**			
5.1.2.1	Local regulations/policies	Information not available/found	Information not available/found	
5.1.2.2	Tribal regulations/policies	Information not available/found	Information not available/found	
5.1.2.3	State regulations/policies	Information not available/found	Information not available/found	
5.1.2.4	Federal regulations/policies	Information not available/found	Information not available/found	
5.1.2.5	International regulations/policies	Information not available/found	Information not available/found	
5.1.2.6	Contractual obligations	Information not available/found	Information not available/found	
5.1.2.7	Repository policies	Investigators and their institutions are responsible for safeguarding the accessed datasets. They must also adhere to the terms of access described in the Data Use Certification (DUC) Agreement and the Genomic Data User Code of Conduct that were signed as part of the data access request. [1] RADxSM Data User Code of Conduct: Key principles and practices agreed to by all research investigators requesting access to RADxSM data. The elements within RADxSM Data User Code of Conduct reflect the terms of access in the Data Use Certification Agreement. Failure to abide by the RADxSM Data User Code of Conduct may result in revocation of an investigator's access to any and all approved datasets. [2] Must researchers sign the Code of Conduct to access RADx-UP data in the RADx Data Hub and if so, when is it signed? [A]t the time of the Data Access Request in submitted into the dbGaP system and attested to by the institutional signing official. [3]	RADx Data Hub policy	[1] https://sharing.nih.gov/accessing-data/accessing-genomic-data/how-to-request-and-access-datasets-from-dbgap (Accessed 4/17/23) [2] RADx Data Use Certification Agreement [3] Responses to RADx Data Hub Questions
5.2	**Governance for data linkage, sharing, access, and use based on data access authorization or applicable regulations/policies (i.e., the origin of the governance)**			
5.2.1	Whether the data can be linked	Does not authorize/specify	Does not authorize/specify	
5.2.2	With what other data can it be linked or can it not be linked (scope of linkage)	Does not authorize/specify	Does not authorize/specify	

Dataset 10 - RADx-UP				
		Raw Language	**Interpretation**	**Source**
5.2.3	Whether data can be shared	The _____ hereby assures that submission of data from the study entitled _____ to an NIH-designated data repository meets the following expectations: • The data submission is consistent, as appropriate, with applicable national, tribal, and state laws and regulations as well as relevant institutional policies. • Any limitations on the research use of the data, as expressed in the informed consent documents, are delineated in the table on page 3. • The identities of research participants will not be disclosed to NIH-designated data repositories. • An Institutional Review Board (IRB), and/or Privacy Board, and/or equivalent body, and a relevant senior-level institutional staff (e.g., Dean, Vice President/Provost for Research, Chief Science Officer) as applicable, has reviewed the investigator's proposal for data submission and assures that: o The protocol for the collection of genomic and phenotypic data is consistent with 45 CFR Part 46;2 o Data submission and subsequent data sharing for research purposes are consistent with the informed consent of study participants from whom the data were obtained; o Consideration was given to risks to individual participants and their families associated with data submitted to NIH-designated data repositories and subsequent sharing, including unrestricted access to genomic summary results; o To the extent relevant and possible,consideration was given to risks to groups or populations associated with submitting data to NIH-designated data repositories and subsequent sharing, including unrestricted access to genomic summary results; and o The investigator's plan for de-identifying datasets is consistent with the HHS Regulations for the Protection of Human Subjects** [1]	RADx Institutional Certification specifies that the CDCC (Duke in the case of RADx UP) submits data to the RADx Data Hub.	[1] Rapid Acceleration for Diagnostics Program (RADx) Institutional Certification: https://radx-up.org/wp-content/uploads/2022/05/RADx-INSTITUTIONAL-CERTIFICATION-FINAL-02022021-1.pdf (Accessed 4/17/23)
5.2.4	How data can be shared (de-identification status, disclosure review)	Please confirm our understanding that all data shared in RADx Data Hub are de-identified. "Correct" [1] The _____ hereby assures that submission of data from the study entitled _____ to an NIH-designated data repository meets the following expectations: • The data submission is consistent, as appropriate, with applicable national, tribal, and state laws and regulations as well as relevant institutional policies. • Any limitations on the research use of the data, as expressed in the informed consent documents, are delineated in the table on page 3. • The identities of research participants will not be disclosed to NIH-designated data repositories. • An Institutional Review Board (IRB), and/or Privacy Board, and/or equivalent body, and a relevant senior-level institutional staff (e.g., Dean, Vice President/Provost for Research, Chief Science Officer) as applicable, has reviewed the investigator's proposal for data submission and assures that: o The protocol for the collection of genomic and phenotypic data is consistent with 45 CFR Part 46;2 o Data submission and subsequent data sharing for research purposes are consistent with the informed consent of study participants from whom the data were obtained; o Consideration was given to risks to individual participants and their families associated with data submitted to NIH-designated data repositories and subsequent sharing, including unrestricted access to genomic summary results; o To the extent relevant and possible,consideration was given to risks to groups or populations associated with submitting data to NIH-designated data repositories and subsequent sharing, including unrestricted access to genomic summary results; and o The investigator's plan for de-identifying datasets is consistent with the HHS Regulations for the Protection of Human Subjects**[2]	1. RADx Institutional Certification specifies that all data shared in the RADx Data Hub must be-identified. 2. RADx Institutional Certification specifies that all data be shared to an NIH-designated repository [RADx Data Hub].	[1] Responses to RADx Data Hub Questions [2] Rapid Acceleration for Diagnostics Program (RADx) Institutional Certification: https://radx-up.org/wp-content/uploads/2022/05/RADx-INSTITUTIONAL-CERTIFICATION-FINAL-02022021-1.pdf (Accessed 4/17/23)
5.2.5	How data can be accessed (access type, data use agreement, data access committee/group approval, IRB LOD, etc.)	All data in the RADx Data Hub are considered controlled access. [1] The individual-level data are to be made available through (check one) -controlled-access -unrestricted access [2]	RADx Institutional Certification specifies that all individual-level data are controlled access.	[1] Responses to RADx Data Hub Questions [2] Rapid Acceleration for Diagnostics Program (RADx) Institutional Certification: https://radx-up.org/wp-content/uploads/2022/05/RADx-INSTITUTIONAL-CERTIFICATION-FINAL-02022021-1.pdf (Accessed 4/17/23)

Dataset 10 - RADx-UP				
		Raw Language	**Interpretation**	**Source**
5.2.6	How data can be used (data use limitations)	NIH expects the submitting institution(s) to select one of the three standard Data Use Limitations (DULs) for appropriate secondary use, or, if necessary, create a customized DUL. DULs are developed based on the original informed consent of the participant(s). -General Research Use (GRU): Use of the data is limited only by the terms of the Data Use Certification: these data will be added to the dbGaP Collection. -Health/Medical/Biomedical (HMB): Use of the data is limited to health/medical/biomedical purposes, does not include the study of population origins or ancestry -Disease-specific [list disease] (DS): Use of the data must be related to the specified disease -Other Additional modifiers to the standard DULs (e.g., Not-for-profit Use Only) can be indicated, if appropriate. Use of the modifiers should have a basis in the informed consent from the participants or in special knowledge of the preferences of the original study population. -IRB Approval Required (IRB): Requestor must provide documentation of local IRB approval -Publication Required (PUB): Requestor agrees to make results of studies using the data available to the larger scientific community. -Collaboration Required (COL): Requestor must provide a letter of collaboration with the primary study investigator(s). -Not-for-profit Use Only (NPU): Use of the data is limited to not-for-profit organizations -Methods (MDS): Use of the data includes methods development research (e.g., development and testing of software or algorithms) -Genetic Studies Only (GSO): Use of the data is limited to genetic studies only. [1] "The RADxSM Data Hub- Addendum to this Agreement outlines additional terms and information which are specific to each requested dataset such as: • Data Use Limitation(s) • Sponsoring NIH Institute or Center • Responsible Data Access Committee • Study Description • Suggested Acknowledgement Statement" [2] The data that was shared with Duke can be used for general research purposes. [3]	RADx Institutional Certification specifies that the use of data is for general research purposes.	[1] Rapid Acceleration for Diagnostics Program (RADx) Institutional Certification: https://radx-up.org/wp-content/uploads/2022/05/RADx-INSTITUTIONAL-CERTIFICATION-FINAL-02022021-1.pdf (Accessed 4/17/23) [2] RADx Data Use Certification Agreement [3] RADx UP Meeting 2
5.2.7	Other (specify)	Does not authorize/specify	Does not authorize/specify	
6	**PII Elements**			
6.1	PII elements collected	Data privacy is super important in communities that we're working in. With large numbers of native Hawaiian pacific islanders, the study has been careful about identifiers. The Hawaii study does not collect SSN but does collect DOB and names. [1] The Hawaii study also collects the email address of both the parent and the child. Duke does not have identifiers, only zip codes and sex. The study sites have the rest of the data elements such as name, DOB, etc. [2]	The study collects date of birth, names, parent email address, child email address, zip code, and sez.	[1] RADx UP Meeting 1 [2] RADx UP Meeting 2
6.2	PII elements holder (i.e., party that holds the PII)	Duke does not have identifiers, only zip codes and sex. The study sites have the rest of the data elements such as name, DOB, etc. There are indigenous (Native Hawaiian/Pacific Islander) considerations and the IRB only approved for using de-identified data. [1]	Duke holds zip code and sex. The study sites hold the rest of the PII (name, date of birth, and email addresses).	[1] RADx UP Meeting 2
6.3	Use of common data model, if any, for data collection	In order to ensure consistency in how RADx-UP projects collect data for the RADx Data Hub and simplify the analysis of that data, the NIH defined a set of Common Data Elements (CDEs). The NIH RADx-UP CDEs provide a standard set of study questions that RADx-UP projects are required to use in their COVID-19 testing studies. [1] The Hawaii study uses the format/common data elements that RADx-UP set up. [2]	RADx-UP studies use common data elements as defined by RADx-UP	[1] https://radx-up.org/research/cdes/ (Accessed 4/17/23) [2] RADx UP Meeting 2
7	**Prior Data Linkages**			
7.1	**Dataset linked with other datasets**			[1] RADx UP Meeting 2

Dataset 10 - RADx-UP				
		Raw Language	**Interpretation**	**Source**
7.1.1	Name of other dataset linked to this dataset	PI has worked on individual level location-based secondary analyses such as neighborhood socioeconomic status using Census blocks with geocodes and public datasets from the CDC. [1]	Neighborhood socioeconomic status from Census blocks, public datasets from the CDC	[1] RADx UP Meeting 2
7.1.2	Other dataset type (clinical, EHR, survey, claims, SDOH, etc.)	PI has worked on individual level location-based secondary analyses such as neighborhood socioeconomic status using Census blocks with geocodes and public datasets from the CDC. [1]	Social determinants of health, COVID data	[1] RADx UP Meeting 2
7.1.3	Other dataset source(s)	PI has worked on individual level location-based secondary analyses such as neighborhood socioeconomic status using Census blocks with geocodes and public datasets from the CDC. [1]	CDC	[1] RADx UP Meeting 2
7.1.4	Linking methodology (PPRL or non-PPRL); linkage technology	PI has worked on individual level location-based secondary analyses such as neighborhood socioeconomic status using Census blocks with geocodes and public datasets from the CDC. [1]	Non-PPRL; linked with geocodes	[1] RADx UP Meeting 2
7.1.5	PII elements used for the linkage	PI has worked on individual level location-based secondary analyses such as neighborhood socioeconomic status using Census blocks with geocodes and public datasets from the CDC. [1]	Zip code	[1] RADx UP Meeting 2
7.1.6	Entity resolver (data originator or data linker or third party)	PI has worked on individual level location-based secondary analyses such as neighborhood socioeconomic status using Census blocks with geocodes and public datasets from the CDC. [1]	Data originator	[1] RADx UP Meeting 2
7.1.7	Party performing the linkages	PI has worked on individual level location-based secondary analyses such as neighborhood socioeconomic status using Census blocks with geocodes and public datasets from the CDC. [1]	Study team	[1] RADx UP Meeting 2
7.1.8	Linkage quality assessment	Information not available/found	Information not available/found	
7.1.9	Linked data sharing method (linkage maps or pre-linked dataset)	Information not available/found	Information not available/found	

USE CASE 3 - GOVERNANCE INFORMATION				
Use Case 3: SARS-CoV-2 Vaccination and Asthma-Related School Absence - Does SARS-CoV-2 vaccination result in reduced asthma-related school absences at 3/6/12+ months post-vaccination?				
Dataset 11 - EPA Daily Air Quality Data				
	Dataset Source	Air Quality System		
	Dataset Source Agency	Environmental Protection Agency (EPA)		
	Dataset Type (Clinical, EHR, Survey, SDOH, etc.)	Air quality data		
	Information Sources	Website		
Dataset 11 - EPA Daily Air Quality Data				
		Raw Language	**Interpretation**	**Source**
1	**Data Collection**			
1.1	**Authorizations and Applicable Regulations/Policies**			
1.1.1	**Authorizations**		1. Clean Air Act	
1.1.1.1	Assent	N/A	N/A	
1.1.1.2	Consent	N/A	N/A	
1.1.1.3	IRB/equivalent Privacy Board determination	N/A	N/A	
1.1.1.4	Local/state/federal laws	The Clean Air Act requires that state, local, and tribal air pollution control agencies monitor the air for ambient levels of certain pollutants. This data is useful for health and policy research relating to air pollution and its control (Fann et al. 2015). The requirements for the monitoring program are codified in 40 CFR Part 58. In addition to the required monitoring, many agencies perform additional and/or voluntary monitoring of substances and meteorological parameters. The monitoring program is designed to meet three objectives (40 CFR Part 58 Appendix D.1): Provide air pollution data to the public in a timely manner; Support compliance with ambient air quality standards and emissions strategy development; and Support for air pollution research studies. This data is reported to the United State Environmental Protection Agency (EPA). The monitoring agencies are required to report the measured data, along with metadata about the site and monitoring equipment and associated quality assurance data to the US EPA's Air Quality System (AQS). AQS and its predecessors have been accepting and storing this data for more than 50 years and currently contains more than 3 billion measurements. This document describes (1) the methods by which the data can be obtained, (2) the general nature of of the AQS data set, and (3) some background material about the monitoring program that may help users select and interpret data. [1]	Clean Air Act authorizes data collection by state/local/tribal air pollution control agencies for reporting to the EPA.	[1] https://aqs.epa.gov/aqsweb/documents/about_aqs_data.html (Accessed: 4/18/23)
1.1.1.5	Institutional Certification	N/A	N/A	
1.1.1.6	Data originator agreement	Information not available/found	Information not available/found	
1.1.1.7	Repository agreements/policies	N/A	N/A	
1.1.1.8	Other (specify)	Information not available/found	Information not available/found	
1.1.2	**Applicable Regulations/Policies**			
1.1.2.1	Local regulations/policies	Information not available/found	Information not available/found	
1.1.2.2	Tribal regulations/policies	Information not available/found	Information not available/found	
1.1.2.3	State regulations/policies	Information not available/found	Information not available/found	
1.1.2.4	Federal regulations/policies	Information not available/found	Information not available/found	
1.1.2.5	International regulations/policies	Information not available/found	Information not available/found	
1.1.2.6	Contractual obligations	Information not available/found	Information not available/found	
1.1.2.7	Repository policies	N/A	N/A	
1.2	**Governance for data linkage, sharing, access, and use based on data collection authorization or applicable regulations/policies (i.e., the origin of the governance)**			
1.2.1	Whether the data can be linked	Does not authorize/specify	Does not authorize/specify	
1.2.2	With what other data can it be linked or can it not be linked (scope of linkage)	Does not authorize/specify	Does not authorize/specify	
1.2.3	Whether data can be shared	Does not authorize/specify	Does not authorize/specify	

Dataset 11 - EPA Daily Air Quality Data				
		Raw Language	**Interpretation**	**Source**
1.2.4	How data can be shared (de-identification status, disclosure review)	The Clean Air Act requires that state, local, and tribal air pollution control agencies monitor the air for ambient levels of certain pollutants. This data is useful for health and policy research relating to air pollution and its control (Fann et al. 2015). The requirements for the monitoring program are codified in 40 CFR Part 58. In addition to the required monitoring, many agencies perform additional and/or voluntary monitoring of substances and meteorological parameters. The monitoring program is designed to meet three objectives (40 CFR Part 58 Appendix D.1): Provide air pollution data to the public in a timely manner; Support compliance with ambient air quality standards and emissions strategy development; and Support for air pollution research studies. This data is reported to the United State Environmental Protection Agency (EPA). The monitoring agencies are required to report the measured data, along with metadata about the site and monitoring equipment and associated quality assurance data to the US EPA's Air Quality System (AQS). AQS and its predecessors have been accepting and storing this data for more than 50 years and currently contains more than 3 billion measurements. This document describes (1) the methods by which the data can be obtained, (2) the general nature of of the AQS data set, and (3) some background material about the monitoring program that may help users select and interpret data. [1]	Clean Air Act specifies that ambient air data be shared through EPA's Air Quality System (AQS).	[1] https://aqs.epa.gov/aqsweb/documents/about_aqs_data.html (Accessed: 4/18/23)
1.2.5	How data can be accessed (access type, data use agreement, data access committee/group approval, IRB LOD, etc.)	Does not authorize/specify	Does not authorize/specify	
1.2.6	How data can be used (data use limitations)	Does not authorize/specify	Does not authorize/specify	
1.2.7	Other (specify)	Does not authorize/specify	Does not authorize/specify	
2	**Data Linkage**			
2.1	**Authorizations and Applicable Regulations/Policies**			
2.1.1	**Authorizations**		None found	
2.1.1.1	Assent	N/A	N/A	
2.1.1.2	Consent	N/A	N/A	
2.1.1.3	IRB/equivalent Privacy Board determination	N/A	N/A	
2.1.1.4	Local/state/federal laws	Information not available/found	Information not available/found	
2.1.1.5	Institutional Certification	N/A	N/A	
2.1.1.6	Data originator agreement	Information not available/found	Information not available/found	
2.1.1.7	Repository agreements/policies	The ambient monitoring data in EPA's Air Quality System (AQS) are public domain. You are welcome to download the data and use freely, without submitting a request.	AQS being in the public domain authorizes data linkage.	[1] https://www.epa.gov/outdoor-air-quality-data/do-i-need-request-permission-use-monitoring-data-and-graphics-airdata
2.1.1.8	Other (specify)	Information not available/found	Information not available/found	
2.1.2	**Applicable Regulations/Policies**			
2.1.2.1	Local regulations/policies	Information not available/found	Information not available/found	
2.1.2.2	Tribal regulations/policies	Information not available/found	Information not available/found	
2.1.2.3	State regulations/policies	Information not available/found	Information not available/found	
2.1.2.4	Federal regulations/policies	Information not available/found	Information not available/found	
2.1.2.5	International regulations/policies	Information not available/found	Information not available/found	
2.1.2.6	Contractual obligations	Information not available/found	Information not available/found	
2.1.2.7	Repository policies	The ambient monitoring data in EPA's Air Quality System (AQS) are public domain. You are welcome to download the data and use freely, without submitting a request.	AQS policy	[1] https://www.epa.gov/outdoor-air-quality-data/do-i-need-request-permission-use-monitoring-data-and-graphics-airdata
2.2	**Governance for data linkage, sharing, access, and use based on data linkage authorization or applicable regulations/policies (i.e., the origin of the governance)**			
2.2.1	Whether the data can be linked	N/A	N/A	
2.2.2	With what other data can it be linked or can it not be linked (scope of linkage)	N/A	N/A	
2.2.3	Whether data can be shared	N/A	N/A	
2.2.4	How data can be shared (de-identification status, disclosure review)	N/A	N/A	
2.2.5	How data can be accessed (access type, data use agreement, data access committee/group approval, IRB LOD, etc.)	N/A	N/A	

Dataset 11 - EPA Daily Air Quality Data				
		Raw Language	**Interpretation**	**Source**
2.2.6	How data can be used (data use limitations)	N/A	N/A	
2.2.7	Other (specify)	N/A	N/A	
3	**Data Sharing**			
3.1	**Authorizations and Applicable Regulations/Policies**			
3.1.1	**Authorizations**		1. Clean Air Act	
3.1.1.1	Assent	N/A	N/A	
3.1.1.2	Consent	N/A	N/A	
3.1.1.3	IRB/equivalent Privacy Board determination	N/A	N/A	
3.1.1.4	Local/state/federal laws	The Clean Air Act requires that state, local, and tribal air pollution control agencies monitor the air for ambient levels of certain pollutants. This data is useful for health and policy research relating to air pollution and its control (Fann et al. 2015). The requirements for the monitoring program are codified in 40 CFR Part 58. In addition to the required monitoring, many agencies perform additional and/or voluntary monitoring of substances and meteorological parameters. The monitoring program is designed to meet three objectives (40 CFR Part 58 Appendix D.1): Provide air pollution data to the public in a timely manner; Support compliance with ambient air quality standards and emissions strategy development; and Support for air pollution research studies. This data is reported to the United State Environmental Protection Agency (EPA). The monitoring agencies are required to report the measured data, along with metadata about the site and monitoring equipment and associated quality assurance data to the US EPA's Air Quality System (AQS). AQS and its predecessors have been accepting and storing this data for more than 50 years and currently contains more than 3 billion measurements. This document describes (1) the methods by which the data can be obtained, (2) the general nature of of the AQS data set, and (3) some background material about the monitoring program that may help users select and interpret data.	Clean Air Act authorizes data sharing.	[1] https://aqs.epa.gov/aqsweb/documents/about_aqs_data.html (Accessed: 4/18/23)
3.1.1.5	Institutional Certification	N/A	N/A	
3.1.1.6	Data originator agreement	Information not available/found	Information not available/found	
3.1.1.7	Repository agreements/policies	Information not available/found	Information not available/found	
3.1.1.8	Other	Information not available/found	Information not available/found	
3.1.2	**Applicable Regulations/Policies**			
3.1.2.1	Local regulations/policies	Information not available/found	Information not available/found	
3.1.2.2	Tribal regulations/policies	Information not available/found	Information not available/found	
3.1.2.3	State regulations/policies	Information not available/found	Information not available/found	
3.1.2.4	Federal regulations/policies	Information not available/found	Information not available/found	
3.1.2.5	International regulations/policies	Information not available/found	Information not available/found	
3.1.2.6	Contractual obligations	Information not available/found	Information not available/found	
3.1.2.7	Repository policies	Information not available/found	Information not available/found	
3.2	**Governance for data linkage, sharing, access, and use based on data sharing authorization or applicable regulations/policies (i.e., the origin of the governance)**			
3.2.1	Whether the data can be linked	Does not authorize/specify	Does not authorize/specify	
3.2.2	With what other data can it be linked or can it not be linked (scope of linkage)	Does not authorize/specify	Does not authorize/specify	

Dataset 11 - EPA Daily Air Quality Data				
		Raw Language	**Interpretation**	**Source**
3.2.3	Whether data can be shared	The Clean Air Act requires that state, local, and tribal air pollution control agencies monitor the air for ambient levels of certain pollutants. This data is useful for health and policy research relating to air pollution and its control (Fann et al. 2015). The requirements for the monitoring program are codified in 40 CFR Part 58. In addition to the required monitoring, many agencies perform additional and/or voluntary monitoring of substances and meteorological parameters. The monitoring program is designed to meet three objectives (40 CFR Part 58 Appendix D.1): Provide air pollution data to the public in a timely manner; Support compliance with ambient air quality standards and emissions strategy development; and Support for air pollution research studies. This data is reported to the United State Environmental Protection Agency (EPA). The monitoring agencies are required to report the measured data, along with metadata about the site and monitoring equipment and associated quality assurance data to the US EPA's Air Quality System (AQS). AQS and its predecessors have been accepting and storing this data for more than 50 years and currently contains more than 3 billion measurements. This document describes (1) the methods by which the data can be obtained, (2) the general nature of of the AQS data set, and (3) some background material about the monitoring program that may help users select and interpret data. [1]	Clean Air Act specifies that ambient air data can be shared.	[1] https://aqs.epa.gov/aqsweb/documents/about_aqs_data.html (Accessed: 4/18/23)
3.2.4	How data can be shared (de-identification status, disclosure review)	The Clean Air Act requires that state, local, and tribal air pollution control agencies monitor the air for ambient levels of certain pollutants. This data is useful for health and policy research relating to air pollution and its control [1]. Site Address The street address giving an approximate location of the site. Zip Code The postal zip code in which the monitoring site resides. CBSA The name of the core based statistical area (metropolitan area) where the monitoring site is located. CBSA Code The code of the core based statistical area (metropolitan area) where the monitoring site is located. County The name of the county where the monitoring site is located. County Code The FIPS County Code where the monitor resides. State Code The FIPS code of the state in which the monitor resides. [2]	Clean Air Act specifies that full geographic identifiers including site address, zip code, CBSA, county, and state are shared.	[1] https://aqs.epa.gov/aqsweb/documents/about_aqs_data.html (Accessed: 4/18/23) [2] https://aqs.epa.gov/aqsweb/documents/AQS_Data_Dictionary.html (Accessed: 8/22/23)
3.2.5	How data can be accessed (access type, data use agreement, data access committee/group approval, IRB LOD, etc.)	Does not authorize/specify	Does not authorize/specify	
3.2.6	How data can be used (data use limitations)	Does not authorize/specify	Does not authorize/specify	
3.2.7	Other (specify)	Does not authorize/specify	Does not authorize/specify	
4	**Data Access**			
4.1	**Authorizations and Applicable Regulations/Policies**			
4.1.1	**Authorizations**		1. AQS in public domain	
4.1.1.1	Assent	N/A	N/A	
4.1.1.2	Consent	N/A	N/A	
4.1.1.3	IRB/equivalent Privacy Board determination	N/A	N/A	
4.1.1.4	Local/state/federal laws	Information not available/found	Information not available/found	
4.1.1.5	Institutional Certification	N/A	N/A	
4.1.1.6	Data originator agreement	Information not available/found	Information not available/found	

Dataset 11 - EPA Daily Air Quality Data				
		Raw Language	**Interpretation**	**Source**
4.1.1.7	Repository agreements/policies	The ambient monitoring data in EPA's Air Quality System (AQS) are public domain. You are welcome to download the data and use freely, without submitting a request. [1]	AQS being in the public domain authorizes data access.	[1] https://www.epa.gov/outdoor-air-quality-data/do-i-need-request-permission-use-monitoring-data-and-graphics-airdata
4.1.1.8	Other (specify)	Information not available/found	Information not available/found	
4.1.2	**Applicable Regulations/Policies**			
4.1.2.1	Local regulations/policies	Information not available/found	Information not available/found	
4.1.2.2	Tribal regulations/policies	Information not available/found	Information not available/found	
4.1.2.3	State regulations/policies	Information not available/found	Information not available/found	
4.1.2.4	Federal regulations/policies	Information not available/found	Information not available/found	
4.1.2.5	International regulations/policies	Information not available/found	Information not available/found	
4.1.2.6	Contractual obligations	Information not available/found	Information not available/found	
4.1.2.7	Repository policies	The ambient monitoring data in EPA's Air Quality System (AQS) are public domain. You are welcome to download the data and use freely, without submitting a request. [1]	AQS policy	[1] https://www.epa.gov/outdoor-air-quality-data/do-i-need-request-permission-use-monitoring-data-and-graphics-airdata
4.2	**Governance for data linkage, sharing, access, and use based on data access authorization or applicable regulations/policies (i.e., the origin of the governance)**			
4.2.1	Whether the data can be linked	Does not authorize/specify	Information not available/found	
4.2.2	With what other data can it be linked or can it not be linked (scope of linkage)	Does not authorize/specify	Does not authorize/specify	
4.2.3	Whether data can be shared	Does not authorize/specify	Does not authorize/specify	
4.2.4	How data can be shared (de-identification status, disclosure review)	The Clean Air Act requires that state, local, and tribal air pollution control agencies monitor the air for ambient levels of certain pollutants. This data is useful for health and policy research relating to air pollution and its control [1]. Site Address The street address giving an approximate location of the site. Zip Code The postal zip code in which the monitoring site resides. CBSA The name of the core based statistical area (metropolitan area) where the monitoring site is located. CBSA Code The code of the core based statistical area (metropolitan area) where the monitoring site is located. County The name of the county where the monitoring site is located. County Code The FIPS County Code where the monitor resides. State Code The FIPS code of the state in which the monitor resides. [2]	AQS policy specifies that full geographic identifiers including site address, zip code, CBSA, county, and state are shared.	[1] https://aqs.epa.gov/aqsweb/documents/about_aqs_data.html (Accessed: 4/18/23) [2] https://aqs.epa.gov/aqsweb/documents/AQS_Data_Dictionary.html (Accessed: 8/22/23)
4.2.5	How data can be accessed (access type, data use agreement, data access committee/group approval, IRB LOD, etc.)	The ambient monitoring data in EPA's Air Quality System (AQS) are public domain. You are welcome to download the data and use freely, without submitting a request.	AQS being in the public domain specifies that the data is open access.	[1] https://www.epa.gov/outdoor-air-quality-data/do-i-need-request-permission-use-monitoring-data-and-graphics-airdata
4.2.6	How data can be used (data use limitations)	Does not authorize/specify	Does not authorize/specify	
4.2.7	Other (specify)	Does not authorize/specify	Does not authorize/specify	
5	**Data Use**			
5.1	**Authorizations and Applicable Regulations/Policies**			
5.1.1	**Authorizations**		1. AQS in public domain	
5.1.1.1	Assent	N/A	N/A	
5.1.1.2	Consent	N/A	N/A	
5.1.1.3	IRB/equivalent Privacy Board determination	N/A	N/A	
5.1.1.4	Local/state/federal laws	Information not available/found	Information not available/found	
5.1.1.5	Institutional Certification	N/A	N/A	

Dataset 11 - EPA Daily Air Quality Data				
		Raw Language	**Interpretation**	**Source**
5.1.1.6	Data originator agreement	N/A	N/A	
5.1.1.7	Repository agreements/policies	The ambient monitoring data in EPA's Air Quality System (AQS) are public domain. You are welcome to download the data and use freely, without submitting a request. [1]	AQS being in the public domain authorizes data use.	[1] https://www.epa.gov/outdoor-air-quality-data/do-i-need-request-permission-use-monitoring-data-and-graphics-airdata
5.1.1.8	Other (specify)	Information not available/found	Information not available/found	
5.1.2	**Applicable Regulations/Policies**			
5.1.2.1	Local regulations/policies	The ambient monitoring data in EPA's Air Quality System (AQS) are public domain. You are welcome to download the data and use freely, without submitting a request. [1]	No data use limitations since the data is open access	[1] https://www.epa.gov/outdoor-air-quality-data/do-i-need-request-permission-use-monitoring-data-and-graphics-airdata
5.1.2.2	Tribal regulations/policies	Information not available/found	Information not available/found	
5.1.2.3	State regulations/policies	Information not available/found	Information not available/found	
5.1.2.4	Federal regulations/policies	Information not available/found	Information not available/found	
5.1.2.5	International regulations/policies	Information not available/found	Information not available/found	
5.1.2.6	Contractual obligations	Information not available/found	Information not available/found	
5.1.2.7	Repository policies	The ambient monitoring data in EPA's Air Quality System (AQS) are public domain. You are welcome to download the data and use freely, without submitting a request. [1]	AQS policy	[1] https://www.epa.gov/outdoor-air-quality-data/do-i-need-request-permission-use-monitoring-data-and-graphics-airdata
5.2	**Governance for data linkage, sharing, access, and use based on data access authorization or applicable regulations/policies (i.e., the origin of the governance)**			
5.2.1	Whether the data can be linked	Does not authorize/specify	Does not authorize/specify	
5.2.2	With what other data can it be linked or can it not be linked (scope of linkage)	Does not authorize/specify	Does not authorize/specify	
5.2.3	Whether data can be shared	Does not authorize/specify	Does not authorize/specify	
5.2.4	How data can be shared (de-identification status, disclosure review)	Does not authorize/specify	Does not authorize/specify	
5.2.5	How data can be accessed (access type, data use agreement, data access committee/group approval, IRB LOD, etc.)	Does not authorize/specify	Does not authorize/specify	
5.2.6	How data can be used (data use limitations)	N/A - no data use limitations The ambient monitoring data in EPA's Air Quality System (AQS) are public domain. You are welcome to download the data and use freely, without submitting a request. [1]	N/A	[1] https://www.epa.gov/outdoor-air-quality-data/do-i-need-request-permission-use-monitoring-data-and-graphics-airdata
5.2.7	Other (specify)	Does not authorize/specify	Does not authorize/specify	
6	**PII Elements**			
6.1	PII elements collected	Address Address where the monitoring site is located. City Name of the city, town, village or other municipality in which the site is located. Blank if the site is not located within such a jurisdiction, or if no value was provided. County Name of the county (or equivalent jurisdiction) in which a site is located. State Postal abbreviation for the state or territory in which a site is located. EPA Region EPA Region number in which the site is located. There are ten EPA regions. [1]	Address, city, county, state, and EPA Region are available for each monitor.	[1] https://www.epa.gov/outdoor-air-quality-data/about-air-data-reports#aqidaily

Dataset 11 - EPA Daily Air Quality Data				
		Raw Language	**Interpretation**	**Source**
6.2	PII elements holder (i.e., party that holds the PII)	Address Address where the monitoring site is located. City Name of the city, town, village or other municipality in which the site is located. Blank if the site is not located within such a jurisdiction, or if no value was provided. County Name of the county (or equivalent jurisdiction) in which a site is located. State Postal abbreviation for the state or territory in which a site is located. EPA Region EPA Region number in which the site is located. There are ten EPA regions. [1]	AQS holds the location data.	[1] https://www.epa.gov/outdoor-air-quality-data/about-air-data-reports#aqidaily
6.3	Use of common data model, if any, for data collection	Information not available/found	Information not available/found	Information not available/found
7	**Prior Data Linkages**			
7.1	**Dataset linked with other datasets**			
7.1.1	Name of other dataset linked to this dataset	Information not available/found	Information not available/found	
7.1.2	Other dataset type (clinical, EHR, survey, claims, SDOH, etc.)	Information not available/found	Information not available/found	
7.1.3	Other dataset source(s)	Information not available/found	Information not available/found	
7.1.4	Linking methodology (PPRL or non-PPRL); linkage technology	Information not available/found	Information not available/found	
7.1.5	PII elements used for the linkage	Information not available/found	Information not available/found	
7.1.6	Entity resolver (data originator or data linker or third party)	Information not available/found	Information not available/found	
7.1.7	Party performing the linkages	Information not available/found	Information not available/found	
7.1.8	Linkage quality assessment	Information not available/found	Information not available/found	
7.1.9	Linked data sharing method (linkage maps or pre-linked dataset)	Information not available/found	Information not available/found	