Centers for Medicare & Medicaid Services

Executive Summary for Research Identifiable Data

Version 8/2013

For CMS Use Only

Privacy Board Approval Date: 4/29/2022
Part D Approval Date: Modified Executive Summary

RSCH-2016-50354

DUA User name and title (see Item 16 of DUA)	Chase Bunger, Research Contracts Manager
Requesting Organization[1] (see Item 1 of DUA)	The Curators of the University of Missouri (MU)
Type of Organization	Academic
Study PI (if different from DUA User)	Lemuel R. Waitman, PhD
Study Title	"Greater Plains Collaborative PCORnet Cohort Characterization for Breast Cancer, ALS, and Obesity"
Funding Source	Patient-Centered Outcomes Research Institute (PCORI)

Executive Summary | Dissemination and Reporting of Findings | Data Management Plan | Project Staff | Collaborator Checklist

Nov 2021 Amendment
May 2021 Amendment
January 2019 Amendment
May 2018 Amendment – expand cohort to include additional states & add additional years of data.

EXECUTIVE SUMMARY

1. Study Overview

PCORnet was created in 2014 by the Patient Centered Outcomes Research Institute (PCORI) to further the goals of the Learning Health System and help to answer questions that are important to patient, clinician, and health system stakeholders. PCORnet includes 11 clinical data research networks (CDRNs) and 18 patient-powered research networks (PPRNs). Each of the CDRNs is charged with aggregating longitudinal data on over 1 million individuals that is appropriate to inform all care that a patient receives, as well as creating 3 longitudinal cohorts — one rare disease, one common condition, and one to examine obesity. In addition, PCORI envisions PCORnet as a national resource to understand the development of health and disease in our country; we share this goal. In order to understand all care a patient receives, either the CDRN's clinical delivery system needs to provide all of the care for that patient or needs to supplement delivery system data with insurance claims data that identifies care received outside of the CDRN's clinical delivery system. Throughout phase 1 of the PCORnet contract (March 2014-September 2015) and continuing in phase 2 (October 2015-October 2018) **and continuing in phase 3 (November 2018-December 2024), there** are contractual milestones to understand the degree to which a CDRN manages complete and comprehensive data for the patient population and to test that complete and comprehensive data first with the three longitudinal cohorts defined in the phase 1 contract.

For the Greater Plains Collaborative (GPC) CDRN, our rare disease is amyotrophic lateral sclerosis (ALS), our common disease is breast cancer, and we are tracking obesity in the context of all patients with healthy and unhealthy weights as specified by the PCORnet obesity task force. This obesity cohort is shared with the other CDRN PCORnet colleagues. Within the context of GPC and PCORnet, our study has four overarching aims:

[1] Throughout this document, "organization" can be interpreted as the company, agency, or group or team within a company, depending on which makes more sense in context with the research study for which CMS data files are being used. For example, large companies may defer to a CMS data file inventory for just their team; whereas smaller companies may keep a single CMS data file inventory for the entire company.

(1) To understand the development, treatment, and progression of breast cancer
(2) To understand the development, treatment, and progression of amyotrophic lateral sclerosis (ALS)
(3) To understand the development, treatment, progression, and consequences of healthy and unhealthy (overweight and obesity) weight
(4) Serve as a greater national resource to understand the development, treatment, progression, and consequences of acute and chronic disease cared for within the United States healthcare system and in support of quality care for the conditions championed by the Patient Powered Research Networks in PCORnet as well as the other conditions studied by our peer CDRNs. We are including investigators from the ABOUT Hereditary Breast Cancer PPRN (http://pcornet.org/patient-powered-research-networks/pprn8-university-of-south-florida/) and the Vasculitis PPRN (http://pcornet.org/patient-powered-research-networks/pprn14-university-of-pennsylvania/).

To support these aims from a methodological basis, we seek to expand the data completeness of our patients' health care processes and outcomes and understand the information gain for our complete and comprehensive population by comparing correlations between the Medicare and Medicaid claims data with the data in our CDRN that includes the electronic health record and billing data from each of our component health systems, clinical registry data (e.g. hospital tumor registries), private payer claims data, and patient-reported outcomes, as available, and work with our GPC and CDRN investigators to answer specific cohort questions to achieve our overarching aims.

Our first methodological hypothesis is that when our health systems function as tertiary or quaternary care facilities, they will hold very comprehensive (detailed, multifactorial signals, structured and un-structured) indicators of health and care processes but as patients return to their primary care and especially rural environments, the Medicare and Medicaid claims data will provide valuable signals of follow up care processes and outcomes. In this context, we seek to measure the contribution of claims data to fill in a more "complete" picture of our patients' health.

Second, there is considerable variability in the mix of primary versus specialty care within our health systems and in comparison to the larger United States healthcare system composed of community hospitals and providers. The degree to which analyses based upon the cohorts contained in the GPC reflect the larger populations within our states is not well characterized. We will calculate the distributions of our three conditions and their treatment patterns for our cohorts within the GPC relative to their occurrence in our larger state regions.

Finally, during the tertiary care periods, understanding the correlation between EMR, billing and claims data is not well described for a diverse set of health systems encompassed by the GPC. Most of the risk adjustment and health services research has rested on standardized claims data from sources such as CMS. Our ability to use rich and more current clinical and administrative data to drive trial recruitment and observation for PCORnet will hinge on the ability to quality control these new data sources and quantify the relative support of EMR and billing systems-based computable phenotypes versus established claims-based models.

In closing, we will create a de-identified resource that merges CMS claims and GPC site data to characterize the increase in data completeness and comprehensiveness provided through claims integration. We will be conducting three technical and methodological analyses in support of our cohorts:

(1) Quantifying completeness of the health system-derived data repositories
(2) Evaluating the distributions of health and care processes for the patients within the GPC versus the larger Medicare and Medicaid populations in our region to understand how studies of the GPC population generalize to the broader populations in our states.
(3) Using CMS claims data to enhance quality control processes for aggregating health system-derived data and establishing correlations with CMS claims data for health system-derived data to support trial recruitment and observational studies, i.e. validating the use of EMR data for recruitment.

Collaboration Outline:

Collaborators with access to RIF

Organization	Description of Role in Project	Access to CMS RIF (cell size less than 11)	Submits Finder Files?	Reviewing aggregate results?	WG/ Area of Focus?
University of Missouri (Requesting Org)	Leading site. Provides centralized data governance, storage, integration and manage accesses to CMS data for all approved parties.	Yes	Yes	Yes	Obesity, ALS, Breast Cancer
University of Kansas Medical Center	Leading working groups with investigators performing data analysis to questions in Aim 2 and 3 for ALS and Obesity population.	Yes	Yes	Yes	ALS, Obesity
University of Iowa	Leading working groups with investigators performing data analysis to questions in Aim 1 for breast cancer population.	Yes	Yes	Yes	Breast Cancer
Medical College of Wisconsin	Leading working groups with investigators performing data analysis to questions in Aim 1 for breast cancer population.	Yes	Yes	Yes	Breast Cancer

Collaborators without access to RIF

Organization	Description of Role in Project	Submits Finder Files?	Reviewing aggregate results?
Marshfield Clinic	The site will contributed local EHR data for data linkage and integration. Their data is essential for investigators to address questions for all Aims.	Yes	Yes
University of Nebraska Medical Center	The site will contributed local EHR data for data linkage and integration. Their data is essential for investigators to address questions for all Aims.	Yes	Yes
University of Texas Health Science Center, San Antonio	The site will contributed local EHR data for data linkage and integration. Their data is essential for investigators to address questions for all Aims.	Yes	Yes
University of Texas Southwestern Medical Center	The site will contributed local EHR data for data linkage and integration. Their data is essential for investigators to address questions for all Aims.	Yes	Yes
~~Indiana University~~	~~The site will contributed local EHR data for data linkage and integration. Their~~	~~Yes~~	~~Yes~~

	~~data is essential for investigators to address questions for all Aims.~~		
Allina Health	The site will contributed local EHR data for data linkage and integration. Their data is essential for investigators to address questions for all Aims.	Yes	Yes
Intermountain Healthcare	The site will contributed local EHR data for data linkage and integration. Their data is essential for investigators to address questions for all Aims.	Yes	Yes
University of Utah	The site will contributed local EHR data for data linkage and integration. Their data is essential for investigators to address questions for all Aims.	Yes	Yes
University of Texas Houston	The site will contributed local EHR data for data linkage and integration. Their data is essential for investigators to address questions for all Aims.	Yes	Yes
Washington University	The site will contributed local EHR data for data linkage and integration. Their data is essential for investigators to address questions for all Aims.	Yes	Yes

2. How have you ensured that your data request includes the minimum amount of data necessary to achieve your research objectives?

2.1. **Please describe how this cohort will meet minimum data necessary.** *(Include estimated cohort size. Refer to your cost invoice.)*

This cohort will include only individuals residing in GPC regions supported by the GPC institutions NIH Clinical and Translational Science Award/Great Plains IDEA – Clinical and Translational Research, which includes the following states: Kansas, Missouri, Iowa, Wisconsin, Nebraska, Minnesota, Texas, Utah, North Dakota, South Dakota and Indiana. Note: Utah joined the GPC as a partner in 2018 and Nebraska's CTR now supports the state of North Dakota and South Dakota. The geographic region described contains approximately 16 million beneficiaries. Because people may move between health systems, it is necessary to cover the regions, rather than the individuals, covered by the health systems. Therefore to accomplish the second aim in comparing our three studied conditions and their treatment patterns against the larger population within the states, it is necessary to obtain claims data for all patients residing in the eleven states.

The reason for requesting the use of multiple years of data for the pre-defined patient cohort (i.e., ALS, breast cancer, obesity) is to be able to support cohort characterization over time within our health systems. We anticipate refreshing this resource annually to (1) adjust our quality control methods, (2) continue our cohort characterization efforts for ALS, breast cancer, and obesity, and (3) provide an efficient resource to support additional cohort characterization activities approved by CMS as separate subsequent reuse Data Use Agreements.

2.2. **List the CMS data files and years being requested at this time and provide justification for how each will be used in the analysis.** *If requesting reuse of data, include the DUA # to be reused. The list of files should match Item #5 of DUA.*

2.2.1. **Medicare (claims and enrollment) or Medicaid (claims and enrollment)**

We are requesting the following research identifiable files (RIF) from CMS and provided justification for each.

- **Outpatient & Carrier, cohort1 [2011-2017]**
  Outpatient and Carrier claims are necessary to capture healthcare events that occur outside of our health systems (capturing follow up care processes and outcomes) as well as validate events that occur within our health systems. For example, a patient may be diagnosed at the University of Kansas Cancer Center but return to Western Kansas for follow up care. In addition, for breast cancer patients not treated at university cancer centers, we will be able to establish comparative data regarding treatment patterns and outcomes. Furthermore, our cohort leaders have identified ambulatory procedures as key characteristics for characterizing our cohorts.

- **Home Health & Hospice, cohort1 [2011-2017]**
  The homehealth and hospice RIFs are necessary to capture healthcare services provided by home health and hospice providers, which is outside of our health systems. The files contribute towards completing the picture of patients' healthcare history and is especially valuable for understanding the care of our ALS patients.

- **DMERC, cohort1 [2011-2017]**
  The Durable Medical Equipment files are necessary to capture medical equipment provided by suppliers. The files contribute towards completing the picture of patients' healthcare history and supporting services and is especially valuable for understanding the care of our ALS patients

- **Master Beneficiary Summary File Base, (A/B/D) Segment, cohort1 [2011-2017]**
  The base (A/B/D) segment is necessary for evaluating patient characteristics across GPC versus the larger Medicare and Medicaid population in our regions.

- **MedPAR (SS/LS/SNF), cohort1 [2011-2017]**
  The MedPAR file is necessary to capture inpatient and skilled nursing facility (SNF) healthcare events that occur outside of our health systems (capturing follow up care processes and outcomes) as well as validate events that occur within our health systems. The inpatient hospital data is necessary for understanding the correlation between EMR, billing, and claims data during tertiary care periods. For example, we hope to study whether hospital encounters claimed correlate with what is documented in EMR. SNF event data is valuable for completing the picture of patients' follow-up care occurring outside of our health system.

- **MAX Personal Summary, cohort 2, [2011-2012]**
  The MAX Person Summary is a denominator file necessary for evaluating patient characteristics and summarized healthcare utilization of patients across GPC versus the larger Medicare and Medicaid population in our regions.

- **MAX Inpatient File, cohort2, [2011-2012]**
  The MAX Inpatient File is necessary to capture healthcare events that occur outside of our health systems (capturing follow up care processes and outcomes) as well as validate events that occur within our health systems.

- **MAX Prescription Drug, cohort 2, [2011-2012]**

The MAX Prescription Drug file is necessary to capture medication exposures that occur outside of our health systems (capturing follow up care processes and outcomes) as well as validate medication exposures that occur within our health systems.

- **MAX Other Therapies, cohort2, [2011-2012]**
  The MAX Other Therapies file is necessary to capture healthcare events that occur outside of our health systems (capturing follow up care processes and outcomes) as well as validate events that occur within our health systems.

- **MAX Long Term Care File, cohort2 [2011-2012]**
  The MAX Long Term Carefile is necessary to capture healthcare events that occur outside of our health systems (capturing follow up care processes and outcomes) as well as validate events that occur within our health systems. Medicaid accounts for nearly 50% of the nursing home care in our region.

- **Crosswalk (GPC ID to Bene ID; GPC ID to MSIS ID) [current], 2022**
  Crosswalk is necessary to link the Medicare and MAX files to GPC site data (EMR, registry and billing information, death status)

2.2.2. **Part D event data (if using in study)**

- **Part D Event data (available starting CY2006), unencrypted IDs., cohort 1 [2011-2017]**

We are requesting Medicare Part D Drug Event Data for 2011, 2012 and 2013; Medicare Part D Drug Event Data are necessary to capture medication exposures that occur outside of our health systems (capturing follow up care processes and outcomes) as well as validate medication exposures that occur within our health systems. Our cohort leaders have identified medications as key characteristics for characterizing our cohorts

2.2.3. **Part D characteristics files (if using in study)**

- **Part D Drug Characteristic File (append to PDE), cohort 1 [2011-2017]**

We are also requesting Medicare Part D Drug Characteristics File for 2011, 2012, and 2013. It allows us to capture more detailed information on treatment received by patients. We are not requesting Provider, Pharmacy, Plan or Formulary characteristics files.

2.2.4. **Assessment data (if using in study)**

Not applicable

2.3. **If this study will require further years of CMS data that are not yet available for request, please list those CMS data files and years that will be required for the *entire scope of your study*** *(Note: Approval of data files for years that are not yet available will NOT be granted at this time, the information included here will simply provide CMS with an overview of your study).*

We are requesting Medicare 2017 with the potential for subsequent year files.

2.4. **Please list any non-identifiable or non-CMS files you are planning to use in conjunction with the above files for your analysis.** (e.g. Provider of Services (POS) file, AMA Physician Master file, etc.)

1. We will link the CMS data files with our CDRN i2b2-enabled clinical data warehouses at each GPC site containing EMR, registry and billing information, death status, but also sociodemographic information on the census tract level which was derived from the American Community Survey and the 2010 Census

Summary File (https://assets.nhgis.org/original-data/acs/2012ACS_5.pdf and https://assets.nhgis.org/original-data/modern-census/2010sf1.pdf) . The final linked resource used in the analysis will be completely de-identified.

2. We will use the CMS National Plan and Provider Enumeration System (NPPES) for quality assurance of the claims provider versus the providers identified in the electronic medical record for the encounter.
3. We will use the Provider of Services file also to categorize the institutions providing care within and external to our participating GPC institutions.

3. You are requesting Research Identifiable Files (RIF). Why can't Limited Data Set (LDS) files be used for this study?

The objective of PCORnet is to link complete and comprehensive data for our regions and health systems. That requires linking CMS data to each site's EMR data. The linking requires RIF (as opposed to the LDS Files) because we need data files that are not available as in an LDS file, such as the MAX and PDE. We plan to enhance the privacy of our patient level data by hashing the patient identifiers so that the final product managed at MU on behalf of the GPC members is a fully de-identified repository to conduct our completeness/comprehensiveness analysis. This is illustrated in the figure on the following page.

While unique identification numbers or personal identifying information (PII) such as surnames, given names, date of birth, and address would make record linkage straightforward, distribution of such information is however restricted due to privacy concerns. Since patient's trust is with their home health systems and academic medical centers, we seek to avoid transmitting PII through the GPC network infrastructure. We would seek to work with ResDAC, the CMS Contractor, and other CDRNs such as PaTH (http://www.pcornet.org/clinical-data-research-networks/cdrn11-university-of-pittsburgh/) and CAPriCORN to implement a privacy-preserving approach for brokering file linkage between CDRNs and CMS to support PCORnet.

As shown in Figure 1, individual GPC sites will first send the NewWave-GDIT (Research Data Distribution Center under contract with CMS) encrypted beneficiary identifiers (e.g., SSN, HICs, name, DOB) along with a hashed IDs generated from the hashing algorithm (reference figure below). After receiving beneficiary information from each GPC sites, NewWave-GDIT will generate the finder files that contain mappings between the hashed IDs and CMS IDs and send the crosswalk files and CMS data back to MU (not individual GPC sites). Finally, MU will integrate the crosswalk and CMS data with individual site EMR data (limited data set) through linking the hashed IDs on their secure virtual private cloud. This will allow MU to achieve record linkage of patients's EMR and claims data without obtaining or retaining actual patient level PII from individual GPC sites. At last, the merged dataset will be de-identified and made available on a separate virtual private cloud to project team members for running analyses or using the i2b2 client.

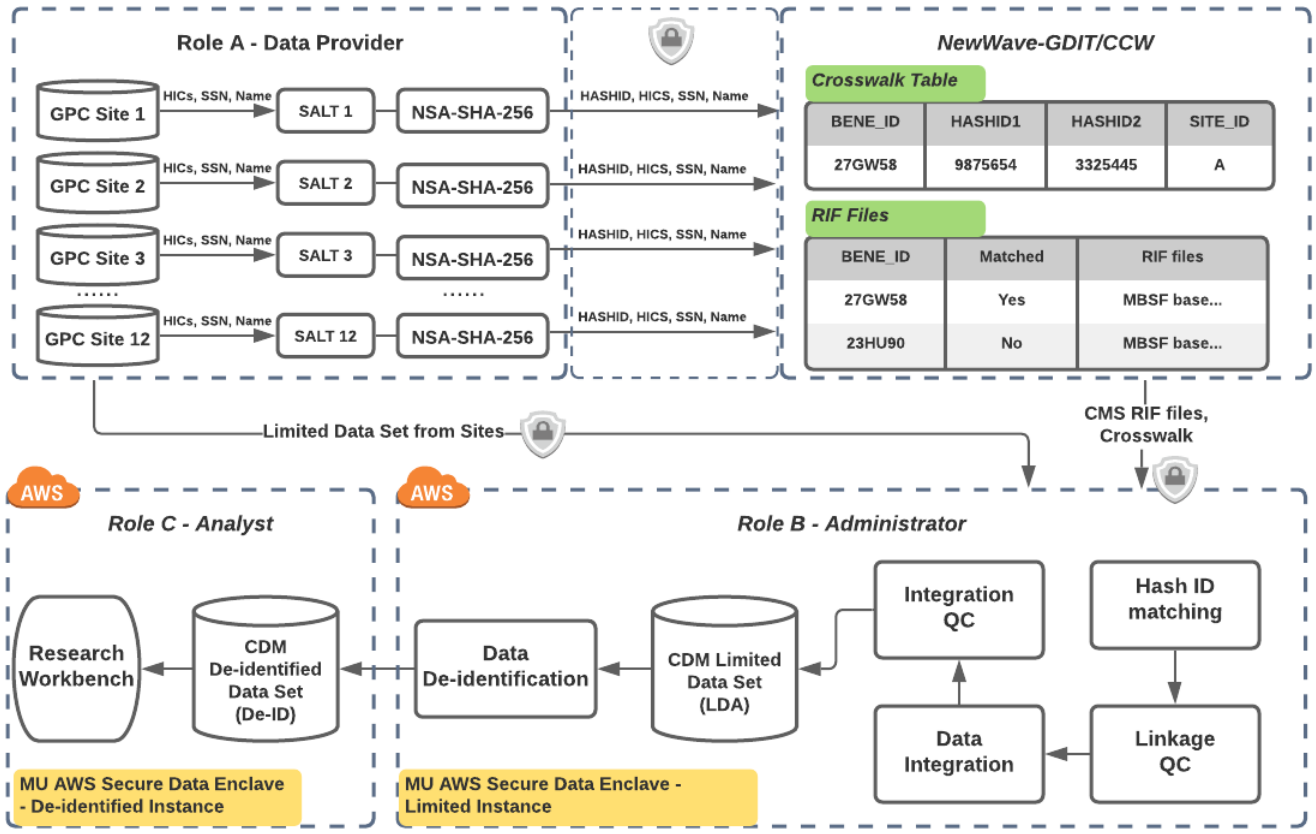

Figure 1 – Privacy-Preserving Data Linkage. NSA-SHA-256 stands for National Security Agency Security Hash Algorithm, which is set of cryptographic hash function approved by NIST for encrypting identifiers. SALT stands for a set of random bits added to each identifier before applying the SHA-256 hashing algorithm. QC stands for quality control.

4. Is it feasible to obtain individual level authorization from Medicare/Medicaid beneficiaries for your research? Explain.

With over one million patients' data needed for these analyses, there is a lack of the time and fiscal resources to obtain patient-level authorization for these data. We also seek to make the resulting data set fully de-identified to enhance patient privacy and need for individual level authorization.

5. If you intend on requesting the National Death Index segment of the Master Beneficiary Summary File, please complete the NDI Supplement.

☐ YES, I've included the NDI Supplement ☒ NO, I'm not requesting the NDI

6. **If this research project is funded by a commercial entity, the (primary) lead investigator attests that they will limit data sharing with the funding entity to aggregated analytic results and will retain the right to independently prepare publications of the study results.** ☒ I attest

	09/15/2015
Signature of (Primary) Lead Investigator	**Date**

DISSEMINATION AND REPORTING OF FINDINGS

From the CMS DUA, "The User agrees that any use of CMS data in the creation of any document (manuscript, table, chart, study, report, etc.) concerning the purpose specified in section 4 (regardless of whether the report or other writing expressly refers to such purpose, to CMS, or to the files specified in section 5 or any data derived from such files) must adhere to CMS' current cell size suppression policy. **This policy stipulates that no cell (e.g. admittances, discharges, patients, services) 10 or less may be displayed.** Also, no use of percentages or other mathematical formulas may be used if they result in the display of a cell 10 or less."
☒ I agree.

Please describe your plans for disseminating the findings from your analysis, including specific media through which you will report results.

The analytic methods and code developed to support PCORnet will be shared as open source materials on our GPC websites to facilitate adoption and dissemination with the PCORnet and potential investigators who might use the PCORnet resource. Our summarized analytic and quality control results (adhering to the cell size suppression policies) will be shared with the PCORnet coordinating center and PCORI.

Results from the study will be targeted for publication in peer-reviewed journals for each of the cohorts and the informatics methods employed. Prior to publication, the results may be presented at a national or international scientific meeting, such as the annual AMIA Joint Summits on Translational Science, Association for Clinical and Translational Science, San Antonio Breast Cancer Symposium, and the American Association of Neuromuscular & Electrodiagnostic Medicine Annual Meeting.

PROJECT STAFF

This section specifically identifies the project staff, organization, and the role of individuals in this project. The requestor and custodian should be named in this section at a minimum.

1. Name & Title of Requestor /User	Chase Bunger, Research Contracts Manager
Organization	The Curators of the University of Missouri
Role in this Study	Research Contracts Manager, Authorized legal signatory
Will this individual have access to raw data, analytic files, or output with cell sizes less than 11?	☒ NO. ☐ YES, this individual will be directly supervised by DUA signatory, [Name] . ☐ YES, this individual has signed the DUA or signature addendum.

2. Name & Title of Primary Custodian	Dr. Lemuel R. Waitman, Associate Dean of Informatics
Organization	The Curators of the University of Missouri
Role in this Study	PI, DMP SAQ Custodian (Roles A and C)
Will this individual have access to raw data, analytic files, or output with cell sizes less than 11?	☐ NO. ☐ YES, this individual will be directly supervised by DUA signatory, [Name] . ☒ YES, this individual has signed the DUA or signature addendum.

3. Name & Title	Shaun Ferguson
Organization	The Curators of the University of Missouri
Role in this Study	DevOps Engineer (Role B)
Will this individual have access to raw data, analytic files, or output with cell sizes less than 11?	☐ NO. ☐ YES, this individual will be directly supervised by DUA signatory, [Name] . ☒ YES, this individual has signed the DUA or signature addendum.

4. Name & Title	Xing Song, Assistant Professor
Organization	The Curators of the University of Missouri
Role in this Study	Research Lead, Data Scientist, QA Analyst (Role B)
Will this individual have access to raw data, analytic files, or output with cell sizes less than 11?	☐ NO. ☐ YES, this individual will be directly supervised by DUA signatory, [Name] . ☒ YES, this individual has signed the DUA or signature addendum.

5. Name & Title	Abu Saleh Mohammad Mosa, Assistant Professor
Organization	The Curators of the University of Missouri
Role in this Study	Bioinformatics Lead (Role C)
Will this individual have access to raw data, analytic files, or output with cell sizes less than 11?	☐ NO. ☐ YES, this individual will be directly supervised by DUA signatory, [Name] . ☒ YES, this individual has signed the DUA or signature addendum.

6. Name & Title	Mei Liu, PhD, Assistant Professor, Medical Informatics
Organization	University of Kansas Medical Center
Role in this Study	Site PI, Co-investigator, leading GPC implementation of linkage methods (Role C)
Will this individual have access to raw data, analytic files, or output with cell sizes less than 11?	☐ NO. ☐ YES, this individual will be directly supervised by DUA signatory, [Name] . ☒ YES, this individual has signed the DUA or signature addendum.

7. Name & Title	Andrew Roberts, Assistant Professor
Organization	University of Kansas Medical Center
Role in this Study	Breast Cancer Researcher (Role C)
Will this individual have access to raw data, analytic files, or output with cell sizes less than 11?	☐ NO. ☐ YES, this individual will be directly supervised by DUA signatory, [Name] . ☒ YES, this individual has signed the DUA or signature addendum.

8. Name & Title	Elizabeth Chrischilles, PhD, Professor, Chair, Public Health
Organization	University of Iowa
Role in this Study	Site PI, Epidemiologist and lead investigator on breast cancer (Role C)
Will this individual have access to raw data, analytic files, or output with cell sizes less than 11?	☐ NO. ☐ YES, this individual will be directly supervised by DUA signatory, [Name] . ☒ YES, this individual has signed the DUA or signature addendum.

9. Name & Title	Bradley McDowell, Director, Population Research Core
Organization	University of Iowa
Role in this Study	Research scientist (Role C)
Will this individual have access to raw data, analytic files, or output with cell sizes less than 11?	☐ NO. ☐ YES, this individual will be directly supervised by DUA signatory, [Name] . ☒ YES, this individual has signed the DUA or signature addendum.

10. Name & Title	Ryan Carnahan, Associate Professor
Organization	University of Iowa
Role in this Study	Breast Cancer Analytics (Role C)
Will this individual have access to raw data, analytic files, or output with cell sizes less than 11?	☐ NO. ☐ YES, this individual will be directly supervised by DUA signatory, [Name] . ☒ YES, this individual has signed the DUA or signature addendum.

11. Name & Title	Mary Schroeder, Assistant Professor
Organization	University of Iowa
Role in this Study	Breast Cancer Analytics (Role C)
Will this individual have access to raw data, analytic files, or output with cell sizes less than 11?	☐ NO. ☐ YES, this individual will be directly supervised by DUA signatory, [Name] . ☒ YES, this individual has signed the DUA or signature addendum.

12. Name & Title	Cole Chapman, Assistant Professor
Organization	University of Iowa
Role in this Study	Breast Cancer Analytics (Role C)
Will this individual have access to raw data, analytic files, or output with cell sizes less than 11?	☐ NO. ☐ YES, this individual will be directly supervised by DUA signatory, [Name] . ☒ YES, this individual has signed the DUA or signature addendum.

13. Name & Title	Joan Neuner, MD, MPH, Associate Professor, General Internal Medicine
Organization	Medical College of Wisconsin
Role in this Study	Site PI, Breast cancer researcher and co-investigator (Role C)
Will this individual have access to raw data, analytic files, or output with cell sizes less than 11?	☐ NO. ☐ YES, this individual will be directly supervised by DUA signatory, [Name] . ☒ YES, this individual has signed the DUA or signature addendum.

14. Name & Title	Kristen Osinski, Biomedical Informatics Business Analyst IV
Organization	Medical College of Wisconsin
Role in this Study	Breast Cancer Analytics (Role A, C)
Will this individual have access to raw data, analytic files, or output with cell sizes less than 11?	☐ NO. ☐ YES, this individual will be directly supervised by DUA signatory, [Name] . ☒ YES, this individual has signed the DUA or signature addendum.

15. Name & Title	Aaron Winn, Assistant Professor
Organization	Medical College of Wiscosin
Role in this Study	Breast Cancer Analytics (Role C)
Will this individual have access to raw data, analytic files, or output with cell sizes less than 11?	☐ NO. ☐ YES, this individual will be directly supervised by DUA signatory, [Name] . ☒ YES, this individual has signed the DUA or signature addendum.

16. Name & Title	James McClay, PhD, Professor
Organization	University of Nebraska Medical Center
Role in this Study	Site PI
Will this individual have access to raw data, analytic files, or output with cell sizes less than 11?	☒ NO. ☐ YES, this individual will be directly supervised by DUA signatory, [Name] . ☐ YES, this individual has signed the DUA or signature addendum.

17. Name & Title	Meredith Zozus, Division Chief, Clinical Research Informatics
Organization	University of Texas Health Science Center, San Antonio
Role in this Study	Site PI
Will this individual have access to raw data, analytic files, or output with cell sizes less than 11?	☒ NO. ☐ YES, this individual will be directly supervised by DUA signatory, [Name] . ☐ YES, this individual has signed the DUA or signature addendum.

18. Name & Title	Jeffrey VanWormer, Research Scientist, Center for Clinical Epidemiology & Population Health, Marshfield Clinic Research Institute
Organization	Marshfield Clinic Inc.
Role in this Study	Site PI
Will this individual have access to raw data, analytic files, or output with cell sizes less than 11?	☒ NO. ☐ YES, this individual will be directly supervised by DUA signatory, [Name] . ☐ YES, this individual has signed the DUA or signature addendum.

19. Name & Title	Lindsey Cowell, Associate Professor
Organization	University of Texas Southwestern Medical Center
Role in this Study	Site PI
Will this individual have access to raw data, analytic files, or output with cell sizes less than 11?	☒ NO. ☐ YES, this individual will be directly supervised by DUA signatory, [Name] . ☐ YES, this individual has signed the DUA or signature addendum.

20. Name & Title	Vino Raj, Director of Clinical Research Informatics and Analytics
Organization	Allina Health System
Role in this Study	Site PI
Will this individual have access to raw data, analytic files, or output with cell sizes less than 11?	☒ NO. ☐ YES, this individual will be directly supervised by DUA signatory, [Name] . ☐ YES, this individual has signed the DUA or signature addendum.

21. Name & Title	Benjamin D. Horn, Director, Cardiovascular and Genetic Epidemiology
Organization	Intermountain Healthcare, Inc.
Role in this Study	Site PI, Data Analytics (Role C)
Will this individual have access to raw data, analytic files, or output with cell sizes less than 11?	☒ NO. ☐ YES, this individual will be directly supervised by DUA signatory, [Name] . ☐ YES, this individual has signed the DUA or signature addendum.

22. Name & Title	Rachel Hess, Chief, Division of Health System Innovation and Research
Organization	University of Utah
Role in this Study	Site PI, Data Analytics (Role A, C)
Will this individual have access to raw data, analytic files, or output with cell sizes less than 11?	☒ NO. ☐ YES, this individual will be directly supervised by DUA signatory, [Name] . ☐ YES, this individual has signed the DUA or signature addendum.

23. Name & Title	Elmer V Bernstam
Organization	University of Texas Health
Role in this Study	Site PI
Will this individual have access to raw data, analytic files, or output with cell sizes less than 11?	☒ NO. ☐ YES, this individual will be directly supervised by DUA signatory, [Name] . ☐ YES, this individual has signed the DUA or signature addendum.

24. Name & Title	Albert Lai
Organization	Washington University at St. Louis
Role in this Study	Site PI
Will this individual have access to raw data, analytic files, or output with cell sizes less than 11?	☒ NO. ☐ YES, this individual will be directly supervised by DUA signatory, [Name] . ☐ YES, this individual has signed the DUA or signature addendum.

DATA MANAGEMENT PLAN[2]

DMP SAQ completed and approved by DPSP. See DMP SAQ Summary Report for the Curators of the University of Missouri Amazon AWS Server Instance environment, expiration 05/25/2022.

[2] Note that CMS is specifically asking for reference to written policies and procedures related to your organization's administrative, technical and physical safeguards. If policies and procedures have not been developed, please explain any ongoing activities your organization is taking to document and make them available to staff. Organizations selected for DPSP reviews will be asked to provide copies of written policies and procedures. Please note that an explanation of the process is not sufficient.