DATA MANAGEMENT AN SHARING PLAN

Element 1: Data Type

A. Types and amount of scientific data expected to be generated in the project:

I. **Quantitative data**

a. Clinical data (via electronic health records, EHR) from estimated 1,000 patients enrolled in Recipe4Health across 12 multi-site safety net clinic systems.
b. We will use propensity score matching to create a control group of non-participating patients from clinics that have not implemented Recipe4Health.
c. We will conduct surveys with Recipe4Health participants (at least 100) as well as non-participants (at least 100) to collect information on social determinants of health, health behaviors, and health outcomes.

II. **Qualitative data**

a. Focus groups and interviews the HERA and other key stakeholders (n=20) We plan to conduct 10 focus groups across Alameda County; 5 with Black adults and 5 with Latinx adults (in Spanish and English).
b. Secondary survey data such as the United States Department of Agriculture (USDA)'s agricultural census on agricultural land
c. Policy archives such as: 1) model ordinances regarding land use and zoning codes

After finalizing data sharing agreements between Stanford and clinic systems, EHR data will be encrypted and transferred from the healthcare systems via secure methods (i.e., Secure File Transfer Path (SFTP) and deposited in the secure server at Stanford School of Medicine and analyzed on encrypted computers approved by Stanford University for high risk data. The data will be de-identified to generate a global unique identifier for the NIMH Data Archive (NDA).

B. Scientific data that will be preserved and shared, and the rationale for doing so:

The quantitative data from this project will be preserved to enable sharing via NDA data of sufficient quality to validate and replicate research findings described in the Aims. NIMH requires data measured from human subjects to be shared using the NDA.

C. Metadata, other relevant data, and associated documentation:

In addition to the subject level data described above, all associated documentation (e.g., study protocols, data collection forms, and data dictionary) will be deposited in the NDA.

Element 2: Related Tools, Software and/or Code:

The data will be analyzed with SAS and R codes using standard procedures and packages, all of which are freely available. The description of statistical analysis will be published, and the codes can be publicly accessed if requested.

Element 3: Standards:

I. Quantitative data:

Participant age, sex, and ethnicity demographic data will be extracted from EHR or collected using the following instruments as defined in NDA:

1) Demographics Short Form (demsf01); **2)** Ethnic Group Questionnaire (ethgrp01)

In compliance with NOT-MH-20-067, the following data will be collected to facilitate aggregation of this data set with other data sets and new data dictionary will be defined in NDA for those validated instruments but not shown in current NIMH data archive:

1) Dietary questionnaire (dq01); **2)** UCLA 3-Item Loneliness Scale (prsi01); **3)** US Household Food Security Survey Module: Six-Item Short Form (new data dictionary); **4)** Generalized Anxiety Disorder Screener (new data dictionary); **5)** Patient Health Questionnaire-9 (new data dictionary); **6)** The standard 4-item set of Healthy Days core questions (CDC HRQOL–4) (new data dictionary); **7)** 2-item exercise as a vital sign (new data dictionary)

II. Qualitative data:

No consensus standards exist.

Element 4: Data Preservation, Access, and Associated Timelines

A. Repository where scientific data and metadata will be archived:

All data will be deposited to NDA starting 12 months after the award begins and will be deposited every six months thereafter following the usual NDA data submission dates.

B. How scientific data will be findable and identifiable:

Data will be findable through the NDA Collection (to be established when this application is funded). For all publications, an NDA study will be created. Each of those studies is assigned a digital object identifier (DOI). The data DOI will be referenced in the publications.

C. When and how long the scientific data will be made available:

The research community will have access to data when the award ends. As required by NDA, studies will also be created that contain the data used for every publication. Those studies will be shared when the pre-print is available. NDA studies have digital object identifiers (DOI) to aid in findability. We will include DOI in relevant publications. NDA will make decisions about how long to preserve the data.

Element 5: Access, Distribution, or Reuse Considerations

A. Factors affecting subsequent access, distribution, or reuse of scientific data:

All enrolled Recipe4Health participants will be consented for broad data sharing. The control group participants meet HIPPA waiver requirements (unfeasible/impractical to contact).

B. Whether access to scientific data will be controlled:

To request access of the data, researchers will use the standard processes at NDA, and the NDA Data Access Committee will decide which requests to grant. The standard NDA data access process allows access for one year and is renewable.

C. Protections for privacy, rights, and confidentiality of human research participants:

All the data will de-identified to generate a global unique identifier for the NIMH Data Archive before sharing and the NDA GUID tool allows researchers to aggregate data from the same research participant without different laboratories having to share personally identifiable information about that research participant. The NDA data dictionaries do not permit personally identifiable information to be shared. NDA maintains a Certificate of Confidentiality.

Element 6: Oversight of Data Management and Sharing:

The Office of Sponsored Research (OSR) at Stanford University has a data management and sharing plan compliance system as part of their process for submitting the annual NIH progress report. OSR is collecting information related to the number of research participants that are deposited each reporting year. OSR will also look for the NDA data DOIs from NDA Studies and will include that information in the annual progress report.

Validation Schedule (this section is required by NIMH)

Within 6 months of the Notice of Award date we will submit a Data Submission Agreement signed by the principal investigators and an institutional business official, as well as define and complete the Data Expected section of this project. Uploads of all initial demographic, clinical, and survey data will be completed using the second submission cycle deadline following the Notice of Award date. Subsequent data uploads will be harmonized, validated, and submitted biannually on the standard January 15th and July 15th submission deadlines.

We also plan to use the NDA validation tool as a quality control measure in the laboratory. The data manager in charge of submitting data to NDA will help researchers in the group validate their data once every month.