

Attachment 1: Detailed Description of Data Collection for the EHR Project

Electronic Health Record Sample and Data Extraction

To provide a robust source of EHR data, AIR engaged Mayo Clinic as its health system partner for this study. Mayo Clinic supports a broad primary care patient service area with over 380,000 primary care patients across southern Minnesota. All Mayo Clinic locations use a single, integrated EHR and revenue cycle management system, which is an ideal single-source platform for the patient demographic and clinical data needs of this study. We will select a sample of southern Minnesota patients with at least one primary care visit in the past 3 years and extract patient data from Mayo Clinic's Epic Systems ONC-certified EHR. With an alpha of 0.05 and power of 80%, our projected sample size is 2,500 randomly selected patients. Assuming a 90% eligibility rate and 40% response rate (thus, 900 completed surveys), this will result in a margin of error of no more than 4 percentage points on prevalence estimates from the EHR sample and a minimum detectable difference of no more than 5.9 percentage points between prevalence estimates from the EHR sample and BRFSS sample.

Following inclusion/exclusion specifications provided by AIR and IMPAQ, Mayo Clinic will define the eligible population, draw a random sample of 2,500 patients, assign each a study-specific ID, and produce two data files. The first file will include contact information for all sampled patients and will be delivered to SNG Research for survey administration. The second file will include requested clinical information and available demographic information and will be delivered to AIR for analysis. Following survey data collection, SNG will deliver a de-identified data file to AIR which will be linked to the EHR data using the study-specific ID.

For the EHR clinical data extraction, we will identify key concepts from the BRFSS survey that can be translated to EHR data elements. IMPAQ International EHR informatics experts will develop a clinical data extract specification that maps BRFSS concepts to EHR data fields related to patient diagnosis and treatment. These EHR data fields include information such as chronic conditions, cancer screenings, treatments, smoking and alcohol intake, pregnancy status, and vaccinations. The code sets developed to describe each data concept in the specification will adhere to the EHR data certification standards and will use coded medical vocabulary adhering to the CMS Quality Data model—such as Systematized Nomenclature of Medicine—Clinical Terms (SNOMED-CT) for clinical documentation and International Classification of Disease (ICD-10) for diagnoses—to support efficient, automated data extraction.

Survey Data Collection

We will administer the BRFSS survey using computer-assisted telephone interviewing (CATI). AIR partner SNG Research will conduct the survey. Mayo Clinic will provide SNG Research with contact information, such as name, phone number, and address from the EHR sample identified in Task 1, and each member's study ID to enable data sharing and linkage after data collection is complete. Given current trends in ownership and use of cell phones and landlines (Blumberg & Luke, 2017), we anticipate that telephone numbers will be a mix of both. To achieve a 40% response rate, we will use a multipronged approach that includes (1) sending sample members advance notice about the study to increase trust and credibility; (2) using web-based sources to obtain updated telephone numbers when needed; (3) scheduling calls at different times (days and evenings, weekdays and weekends); (4) leaving voicemails with the reason for the call; and (5) scheduling a callback for a time that is convenient for the respondent.