GUIDING FRAMEWORK FOR SETTING UP A SICKLE CELL DISEASE SURVEILLANCE SYSTEM

A. DEVELOP MULTI-STAKEHOLDER COLLABORATIONS

1. Form a Multidisciplinary Team

The multidisciplinary team, which may consist of public health professionals, healthcare providers, community organizations, and individuals with sickle cell disease (SCD) or their family members, reviews goals for the system, builds relationships with key partners, and develops plans for using the SCD surveillance data.

The goals of the multidisciplinary team are:

- Form an understanding of the needs and purpose for the SCD surveillance data system with a focus on how these data can be used and shared to improve the lives of people living with SCD;

- Identify the types of data sources that contain the information necessary for SCD surveillance; and

- Inform and engage with partners that will assist with access to data.

2. Engage Stakeholders

Working with stakeholders who will use the data to help improve policy, healthcare practices, and treatments for people with SCD is also important. Engaging with a variety of stakeholders helps to identify the needs and priorities of the SCD surveillance system early, assuring that the information collected and generated by the surveillance system will be useful.

3. Determine Data Sharing Needs

Much of the data included in an SCD surveillance system are subject to the terms and conditions of the Health Insurance Portability and Accountability Act (HIPAA). The groups that maintain these data operate under strict instructions to protect the health and privacy of the individuals whose data are shared. The sharing of this information may be authorized through official letters or agreements. If a pre-existing agreement for data sharing is not in place, another option, depending on the institution, may be to request a waiver of authorization (a legal document that allows an individual's health information to be used or disclosed to a third party) through an accredited Institutional Review Board (IRB).

4. Establish Data Use Agreements

Data Use Agreements (DUAs), which are established for receiving, linking, and using data, are critical to the success of SCD surveillance. It is best to develop DUAs early, as there may be considerable time needed to fully execute these agreements. In particular, DUAs involving the sharing of personal identifying information can take additional time in order to ensure HIPAA guidelines are followed. Most data stewards (the entities responsible for the data) or sources (the location where data comes from) will have their own procedures and rules for developing a DUA.

A DUA will list the steps the SCD surveillance team will need to take as well as specify how the data are to be used and a timeframe for sharing/use. A DUA may require training to ensure data are protected and used appropriately. It may be beneficial to include as part of the multidisciplinary team individuals who are knowledgeable about the legal aspects of DUAs, IRB processes, state and federal laws, changes in state policy, organization rules, and other crucial parts of the program.

DUAs typically require specific information about each data element or variable being requested. This can be very time consuming and requires a thorough understanding of the goals of the SCD surveillance system and the planned analyses and uses of the data.

> **A Data Use Agreement (DUA)** is a legal binding agreement between a lead organization (the organization establishing the surveillance system, in this case) and an external entity (ex. state public health agency, academic institution, health plans/payers, clinics/hospitals), when a request is made that includes the use of personal identifiable data that is covered by legal authority.

B. BUILD DATA INFRASTRUCTURE

1. Data Transfer

Transfer (the sending of information electronically from one location to another) and receipt of information from data sources takes careful planning in order to minimize delays. Data transfer methods will need to be described in detail in the DUA in order to receive IRB approval. These methods may include a careful description of both the data format and the way that the data will be securely transferred. Each data source will have unique data transfer methods. It is helpful to plan ahead for future years of SCD surveillance by determining when additional years of data can be expected and what steps are necessary to initiate future transfers.

2. Establish a Data Linkage Algorithm

Data linkage is a method of bringing together information from different sources about the same person or entity to create a new, richer dataset. Once data are in hand and staff can identify the variables that allow for data linkage, a procedure for linkage can be developed. It is important for the multidisciplinary team, which may include database management experts or members who understand informatics (the study of information technology), to develop a concrete plan that outlines:

- The identified data sources (e.g., newborn screening, vital records, clinical data, laboratory data, Medicaid data, hospital/emergency department discharge data);

- The variables needed from each data source that will be used to link datasets;

- The order of different linkage processes; and

- The creation and contents of the SCD surveillance system file.

Common identifiers that can be used for linkage may include:

- Social security number (SSN)
- Date of birth (DOB)
- Full name
- Sex
- Geographic location (zip code or county)
- Telephone numbers
- Diagnosis
- Dates of service
- Medical record number

Data linkage is typically an ongoing process; each linkage provides new information that may inform the next potential linkage. As a part of the plan, the team will need to select a data linkage software (e.g., R, SAS) or tool (e.g., Link+) to link the datasets. When developing data linkage plans, consider the costs of data source access, data linkage software, and staff time to complete the data linking.

3. Prepare Data for Linkage

Preparing data for linkage is a time-consuming task that involves standardizing data format, cleaning the data (the process of assessing data quality and, if needed, fixing or removing incorrect or incomplete data within a dataset) and de-duplicating the data (searching for and removing duplicate information). However, these tasks are critical for the multi-step process of linking together all the datasets that make up an SCD surveillance system.

4. Develop a Data Analysis Plan

A good data analysis plan is informed by the goals of the program. These goals define the questions the program is intended to answer and what will be done with the data and analytic results. Some research questions that have been explored in states with SCD surveillance programs include:

- How does access to care change over time among people living with SCD?
- How do different treatments for SCD impact the health of people living with SCD?
- How does geographic location or insurance status impact the health of people with SCD?
- How are demographic variables, such as socio-economic status (SES) or ethnicity, associated with treatment, care, and support services?

C. SHARE DATA FINDINGS

The most important phase in the SCD surveillance process is informing interested stakeholders and the SCD community of key findings resulting from the data collection. While developing a data analysis plan, it is important to think about how to make data findings available and meaningful to key stakeholders, the public, and other state health agencies. A plan for sharing findings from the SCD surveillance system and a timeline for regular sharing of data analysis progress may improve collaboration. This plan should be developed early on as part of the planning process and presented to key stakeholders for their input and buy-in, which will improve the team's ability to maintain the surveillance system.

The SCD surveillance system data can be used to

- Describe the SCD population of a state, including ages, geographical location, and where health care was received;

- Investigate questions related to health care, such as the use and impact of new SCD treatments; and

- Describe differences in access to care between children and adults with SCD.

Making data available to the public, healthcare providers, and key stakeholders helps to inform best practices, improve healthcare access and services, and develop new resources.

For more information

Visit: https://www.cdc.gov/ncbddd/hemoglobinopathies/building-surveillance-capacity.html

Email: https://wwwn.cdc.gov/DCS/ContactUs/Form

Call: 800-CDC-INFO

Follow us on Twitter: @CDC_NCBDDD