Multi-State EHR-Based Network for Disease Surveillance (MENDS) Governance Principles, Policies, and Processes

This document was prepared by the Public Health Informatics Institute, a program of the Taskforce for Global Health, to describe governance policies and procedures for the MENDS network.

Version History

Version and Date		Author(s)	Description and Notes
1.0	August 2020	E. Kraus	Initial draft reflecting input from six data partners, NACDD, and CDC Heart Disease and Stroke team prior to MENDS Governance Committee review.
1.2	October 2020	E. Kraus	Incorporated feedback from data partners, simplified and removed unnecessary content, and made terminology consistent. Approved by Governance Committee on 10.20.2020.
2.0	June 2021	E. Kraus and K. Hohman	Updated to incorporate data use and user rules, added software governance content, and retitled document to accommodate expanded scope. Approved by Governance Committee on 6.15.2021.
2.5	March 2023	E. Kraus	Updated to reflect partner site changes, addition of the national visualization tool, and 2021–2023 network operations. Circulated for comment in May 2023.
2.8	June 2023	E. Kraus	Updated to reflect comments from partner sites.
3.0	July 2023	E. Kraus	Voted on by the Governance Committee on 7.18.2023.
3.1	March 2024	L. Saintus and S. Shaw	Updated with preliminary edits and suggestions made to the document; reviewed by NACDD.
3.2	June 2024	L. Saintus	Updated to reflect comments from partner sites and updated Appendix 3.
4.0	July 2024	L. Saintus	Voted on and approved by the Governance Committee on 7.16.2024.
4.1	June 2025	S. Shaw, K. Hohman and A. Martinez	Updated with minor edits to correct partner names, add references, and revise appendices; reviewed by NACDD.
4.3	July 2025	S. Shaw	Updated with minor edits, circulated for comment in July 2025, and discussed at MENDS Governance Committee meeting on 7.15.25.

Table of Contents

MENDS
MULTI-STATE EHR-BASED NETWORK FOR DISEASE SURVEILLANCE

Section 1: Background

MENDS (Multi-State EHR-Based Network for Disease Surveillance) is a project begun in 2018 and initially funded by the Centers for Disease Control and Prevention's (CDC) Division for Heart Disease and Stroke Prevention. The project aims to improve chronic disease surveillance by using data from electronic health records (EHRs). MENDS is a distributed EHR-based surveillance network for monitoring chronic disease to inform public health planning and operations. MENDS uses open-source software, namely Electronic Medical Record Support for Public Health (ESP), PopMedNet™, RiskScape, and the interactive Visualization of EHR-Based Surveillance Tool (iVEST™), a national visualization tool. Detailed descriptions of MENDS components, implementation and governance have been published and are available for reference.[1,2]

Participation in the MENDS technology and partnership is voluntary, and establishing a governance infrastructure is critical for the health and functionality of the network. In MENDS, the governance infrastructure includes a decision-making body, policies, processes, and agreements that guide the network in collection, management, use, exchange, and release of public health surveillance information. The goal of MENDS governance is to establish relationships and infrastructure for a sustained data sharing partnership by achieving the following objectives:

- Codifying how MENDS data are accessed, assembled, used, and disseminated
- Facilitating the use of MENDS data to enable better public health decision-making
- Tracking and aligning national and partner site regulatory requirements and bylaws with MENDS governance infrastructure
- Enabling standards-based data and information technology (IT) best practices within MENDS infrastructure
- Supporting participation in MENDS to improve data quality and coverage for national chronic disease surveillance.

American Health Information Management Association (AHIMA) governance principles[3] for healthcare have been adapted to guide the MENDS network:

- *Transparency*: Participants will be fully apprised of the governance process and have opportunities to provide input.
- *Integrity*: The information managed (i.e., collected, generated, and provided by MENDS) can reasonably be expected to be authentic and reliable.
- *Protection*: The information managed has appropriate levels of privacy protection and protection from breaches and/or loss.
- *Compliance*: The information managed complies with applicable laws, regulations, standards, and organizational policies.

[1] Hohman, K.H., et al. (2023). "Leveraging Electronic Health Record Data for Timely Chronic Disease Surveillance: The Multi-State EHR-Based Network for Disease Surveillance." *Journal of Public Health Management and Practice*, *29*(2), 162–73

[2] Kraus, E.M., et al. (2024). "Fostering Governance and Information Partnerships for Chronic Disease Surveillance: The Multi-State EHR-Based Network for Disease Surveillance." *Journal of Public Health Management and Practice 30*(2), 244-54.

[3] Empel S. The way forward. AIHMA develops information governance principles to lead healthcare toward better data management. J AHIMA. 2014 Oct;85(10):30-2; quiz 34. PMID: 25682641.

- *Retention*: Legal, regulatory, fiscal, operational, and historical requirements are considered as information is maintained.
- *Disposition*: Secure and appropriate disposition for information no longer required is maintained.

1.1 MENDS Network

As of July 2025, MENDS has five partner sites. Each partner site has one data contributor (i.e., organization contributing data to MENDS) and most also have at least one data user. While not all sites currently have an active data user, the goal is for each to establish and maintain a collaborative relationship between contributors and users to support meaningful data use over time. Data contributors represent one or multiple health systems (data owners) and contribute data on their behalf. Each partner site is represented on the Governance Committee, i.e., the MENDS decision-making body. The committee is composed of individuals representing data contributors (i.e., those organizations that contribute clinical data for use by MENDS (Figure 1.1)) and data users (i.e., generally state or local health departments that will use MENDS data for chronic disease surveillance). MENDS project team members from the National Association of Chronic Disease Directors (NACDD), CDC, and other partners may also attend the MENDS Governance Committee in a non-voting capacity.[4]

Figure 1.1: Map of MENDS Data Contributors

[4] List of MENDS data users and other MENDS partners. https://chronicdisease.org/cphl/technical-assistance-hub/data-modernization/mends/

The content of the following pages reflects the network operations, active work (e.g., implementation of the Fast Healthcare Interoperability Resources (FHIR®) pilot) and guidance from members of the MENDS project team, the Governance Committee, emerging standards and industry best practices, and the needs of MENDS partner sites and public health. MENDS supports innovation and prioritizes adoption of emerging standards or technology to improve network efficiency or the experience of network participants. Networks such as MENDS evolve over time, and governance materials are adjusted to reflect that evolution.

1.2 Document Structure

This document addresses data, software, information, and partnership governance issues within the two main domains of data and software governance. Governance topics within each domain are listed below. See Appendix 1: Glossary of Terms for a definition of MENDS terms.

Data Governance	Software Governance
Data Use	Electronic Medical Record Support for Public Health (ESP)
Data Timeliness	Software Installation and Training
Data Quality and Validation	Source Data and ESP Data
Data Suppression and Privacy	PopMedNet
Representativeness	PopMedNet Queries
Advanced Analytics	RiskScape
Reproducibility and Acceptability	National Visualization Tool/iVEST
	Security
	Software Maintenance and Enhancement

Each governance topic has a devoted section that includes an objective, description of MENDS functionality, Governance Committee responsibilities, partner sites requirements, and project team responsibilities. In each section, references to relevant governance processes or guidelines are listed in the appendices.

Section 2: Data Governance

This section on information governance covers the following topics:

- Data use
- Data timeliness
- Data quality and validation
- Data suppression and privacy
- Representativeness
- Advanced analytics
- Reproducibility and acceptability

2.1 Data Use

Objective: Ensure local, state, and national public health surveillance information produced by MENDS can be used by project participants and, when appropriate, is made publicly available with the approval of appropriate partner site(s) and/or the Governance Committee.
Description of MENDS Functionality
• There are several types of MENDS data users: members of the MENDS project team and individuals from partner site data contributors or data user organizations. • Appendix 2: Data Use Guidelines defines MENDS data types and guidelines for the use of each. • MENDS data products contain aggregate (crude and/or weighted) data. Examples of data products include MENDS Coverage Report, RiskScape visualizations, Hypertension Prevalence and Control Reports, and other ad hoc reports (see Software Governance). • Authorized users can access partner site MENDS crude data via PopMedNet queries and RiskScape (see Software Governance). • The public is not considered a data user (in accordance with the MENDS definition) but may be given access to view de-identified aggregate and/or weighted data products as approved by the contributing partner site(s) and in accordance with Appendix 6: Data Product Review and Dissemination Guidelines. • MENDS uses algorithms and indicators to generate surveillance data. An algorithm is a defined set of data elements and logical expressions used to identify one or more chronic disease case definitions (sometimes referred to as e-phenotypes). Indicators are the inclusion and exclusion criteria for the numerator and denominator to express a measure. • ESP has more than 25 algorithms[5], each of which can be used to derive one or multiple public health measures. Six of these algorithms have been prioritized by MENDS and are in different phases of testing and use: hypertension, cholesterol/statin use, diabetes, obesity, smoking, and asthma. Some indicators include multiple measures. • No patient- or record-level data are made available to users through RiskScape, PopMedNet, iVEST, or MENDS data products; only aggregate surveillance information (i.e., counts and rates) is accessible to MENDS data users. • MENDS data products reflect national or local estimates. National and some local data products are created by the MENDS Coordinating Center, and partner sites may create additional local data products. • MENDS data use is limited to public health practice—research uses are not currently permitted. MENDS data use may be expanded in the future to include novel use cases or research. These may include patient-level data if approved by the partner sites, MENDS project team, and the Governance Committee.

[5] ESP Algorithms. https://espnet.atlassian.net/wiki/spaces/EP/pages/93585410/ESP+Algorithms

Governance Committee Responsibilities	Partner Site Requirements
• Monitor and provide guidance on use of MENDS data • Identify and discuss opportunities to use MENDS data • Disseminate MENDS data products to interested parties	• Use MENDS data and data products • Adhere to data use and user guidelines • Maintain the relationship between data contributor(s) and data user(s) within the partner site
MENDS Project Team Responsibilities	
• Create and maintain guidelines for MENDS data use • Support the MENDS Coordinating Center in creating MENDS data products • Support partner sites in creating local MENDS data products	

2.2 Data Timeliness

Objective: Ensure MENDS surveillance can be conducted on data that are as near to real time as possible	
Description of MENDS Functionality	
• Timeliness of MENDS data is a project priority and is achieved by keeping data in partner site ESP data marts as recent as possible through frequent refreshes. Providing timely data is one of MENDS' key strengths compared with traditional surveillance data sources. • Data contributors maintain an ESP database that is refreshed at least quarterly (preferably monthly) and are responsible for conducting refreshes of the ESP database and notifying the project team when refreshes will occur, are completed, or if an error occurs during a data refresh. • Data contributors share information about data latency with partner site data users as appropriate. • RiskScape displays the date of the most recent data refresh.	
Governance Committee Responsibilities	**Partner Site Governance Requirements**
• Discuss ongoing efforts to improve data timeliness	• Provide refreshed data for ESP data mart at least quarterly • Provide data refresh information to data users as appropriate
MENDS Project Team Responsibilities	
• Monitor data refreshes across partner sites • Support partner site efforts to progress toward more frequent refreshes	

2.3 Data Quality and Validation

Objective: Assure MENDS generates the highest quality surveillance information

Description of MENDS Functionality

- MENDS prioritizes examining the quality of EHR data and the validity of resulting surveillance information.
- MENDS specifies data quality and validation activities based on its surveillance use case.
- Appendix 5: Data Quality and Validation Process defines the process that MENDS follows for data quality and validation.
- Participating in data quality and validation activities is mandatory for partner sites.
- Data quality and validation activities occur in stages, beginning with overall data quality assessment and data characterization.
- The MENDS Coordinating Center manages assessment of data quality to examine data completeness, conformance to the ESP data model, consistency between source data and the data mart, and accuracy of mappings. (See Appendix 4: MENDS Electronic medical record Support for Public Health (ESP) Data Model.)
- Data quality issues that directly affect MENDS indicators and cannot be reconciled are adjudicated on an individual basis.

Governance Committee Responsibilities	Partner Site Governance Requirements
• Discuss network-wide data quality considerations • Provide guidance and feedback on data quality activities • Recommend improvements	• Participate in data quality and validation activities • Investigate and reconcile data quality issues as identified

MENDS Project Team Responsibilities

- Track data quality and validation activities across partner sites
- Work with partner sites to address any identified data quality issues

2.4 Data Suppression and Privacy

Objective: Protect the identity of data contributors and data owners and the privacy of individuals whose data are leveraged by MENDS

Description of MENDS Functionality

Data Suppression

- Data suppression refers to the methods or restrictions applied to limit the disclosure of information and reduce the number of estimates with unacceptable levels of statistical reliability.
- MENDS adheres to the Center for Medicaid and Medicare Services (CMS) small cell suppression policy[6], which defines small cells as cells with a value from 1 to 10. CMS stipulates that no cell containing a value of 1–10 can be reported. MENDS requires that cell counts or map units with a value of 1–10 in the numerator or denominator be labeled as "10 or less" and that use of such suppressed values for prevalence estimates be addressed using proper statistical procedures.
- A value of zero does not violate the minimum cell size policy and can be reported.
- Upon request by a partner site, RiskScape can be configured to suppress small cells and exclude suppressed cell counts from totals so that small cell counts cannot be calculated in reverse. Notably, configurations may not vary by jurisdiction within a partner site.
- Upon request by a partner site, PopMedNet queries can be customized to accommodate partner site-specific small cell suppression needs.

Privacy

- MENDS data can include Protected Health Information (PHI) as defined by the Health Information Portability and Accountability Act (HIPAA); thus, MENDS is subject to HIPAA privacy rules.
- RiskScape and PopMedNet may report counts of patients and prevalence estimates aggregated at the five-digit ZIP Code and census tract level that are considered a limited dataset under HIPAA.
- For national data products, MENDS uses both HIPAA-recommended methods for de-identification—Safe Harbor and Expert Determination[7]—as appropriate (See Appendix 6: Data Product Review and Dissemination Guidelines).
- The MENDS Coordinating Center performs expert determination for national data products but does not perform de-identification reviews for partner site-generated data products.

Governance Committee Responsibilities	Partner Site Governance Requirements

[6] CMS small cell suppression policy: https://resdac.org/articles/cms-cell-size-suppression-policy

[7] Guidance Regarding Methods for De-identification of Protected Health Information in Accordance with the Health Insurance Portability and Accountability Act (HIPAA) Privacy Rule: https://www.hhs.gov/hipaa/for-professionals/special-topics/de-identification/index.html

• Provide input and guidance on data suppression and privacy issues	• Adhere to MENDS guidelines and local regulations for data suppression and privacy • Ensure compliance with data suppression and privacy rules among data users • Ensure that locally generated partner site data products are de-identified
MENDS Project Team Responsibilities	
• Protect the privacy of individuals represented in MENDS data products • Perform the expert determination process for MENDS data products, when applicable • Adhere to MENDS data suppression guidelines in MENDS data products • Ensure that MENDS software adhere to MENDS and partner-site small cell suppression and privacy guidelines and regulations • Notify partner sites about privacy or suppression breaches within 2 business days	

2.5 Representativeness

Objective: Generate surveillance information that prioritizes representativeness of the underlying population based on all available socio-demographic and geographic data	
Description of MENDS Functionality	
• Representativeness may be an issue of demographic factors, geographies, social determinants (e.g., education, insurance status, or income), or patterns of healthcare utilization. • MENDS data products must include a description of the representativeness of the source data and, when appropriate, the methods used to minimize potential bias in creating the estimates. • RiskScape aggregates and presents crude counts of patients for each indicator. • Patient counts are visible on all RiskScape reports. • The MENDS project team conducts a coverage analysis annually by querying and aggregating overall and stratified counts of patients across partner sites to review and publish national population coverage information. National population coverage information is de-identified and may be used in publicly released MENDS informational materials.	
Governance Committee Responsibilities	**Partner Site Governance Requirements**
• Discuss network population coverage and related representativeness and health topics	• Monitor the MENDS population coverage for jurisdictions with data users • Maintain partner site profile with coverage information for data users
MENDS Project Team Responsibilities	
• Monitor population coverage within and across partner sites to highlight strengths and limitations in representativeness • Annually conduct coverage analysis and produce a network coverage report	

2.6 Advanced Analytics

Objective: Maximize the proper use of MENDS data for local and national public health surveillance, planning, and evaluation by applying statistical tools and complementary data

Description of MENDS Functionality

- Advanced analytics includes using weighting and modeling to generate adjusted estimates that are representative of the underlying population and surveillance estimates for geographic areas and sociodemographic subgroups.
- Advanced analytic activities are conducted by a weighting and estimation vendor that works in collaboration with partner sites, the Governance Committee, and the MENDS project team.
- The use of advanced analytics is restricted to public health practice use cases—MENDS does not use advanced analytics to conduct research.
- Advanced analytic activities include aspects of traditional weighting methods, as well as model-based estimation.
- Appendix 3: MENDS Weighting and Modeling Summary describes the statistical approach used on MENDS data, and a Modeling Operation Procedure (MOP) document offers a detailed description of weighting and modeling methods.
- Advanced analytics can be conducted within a partner site, with multiple sites, or across all partner sites.
- Partner site participation in advanced analytics is required.
- Tools to generate adjusted estimates are available to partner sites upon request.
- Weighted estimates are based on de-identified data as defined by HIPAA because they do not reflect observed data.
- Because adjusted estimates based on insufficient data may be unstable, MENDS uses thresholds of acceptable error to suppress unstable estimates.

Governance Committee Responsibilities	Partner Site Governance Requirements
• Discuss advanced analytic activities and outputs • Monitor partner site participation in advanced analytics	• Develop and execute a data use agreement with the weighting and estimation vendor • Provide input to advanced analytic projects

MENDS Project Team Responsibilities

- Manage advanced analytic activities
- Coordinate review of data products from advanced analytic activities
- Ensure validity of statistical methods used in advanced analytic activities

2.7 Reproducibility and Acceptability

Objective: Ensure surveillance estimates meet acceptable standards and best practices of data quality and presentation

Description of MENDS Functionality

- Reproducibility is a challenge because MENDS data change frequently as new sites are added, data are collected, data quality is improved, data contributors join, and existing partner sites gain or lose patients.
- Partner sites maintain the most current and highest quality data. As data are updated, previously obtained records may be updated or deleted, and results can change slightly over time.
- RiskScape and PopMedNet facilitate reproducibility by allowing users to repeat an analysis or using parameters, e.g., dates, to study historical time periods.
- RiskScape includes details on report dates and population size on all data visualizations.
- PopMedNet query results are archived with the date that the results were created.
- MENDS ensures acceptability of national data products by conducting a transparent and inclusive review process with partner sites (See Appendix 6: Data Product Review and Dissemination Guidelines).
- Partner sites are guaranteed an opportunity to review and comment on national data products that use their respective site's data.

Governance Committee Responsibilities	Partner Site Governance Requirements
• Discuss and provide feedback on the MENDS data product review process • Discuss strategies for achieving greater acceptability	• Review MENDS data products • Implement an internal review process for locally generated data products • In locally generated data products, adhere to MENDS guidelines and local publication guidelines

MENDS Project Team Responsibilities

- Prepare MENDS data products
- Conduct a transparent and inclusive review process for every MENDS data product

Section 3: Software Governance

This section on software governance addresses the following topics:

- ESP
- Software installation and training
- Source data and ESP data
- PopMedNet
- PopMedNet queries
- RiskScape
- National Visualization Tool/iVEST
- Security
- Software maintenance and enhancements

MENDS contracts with an IT vendor to support software implementation, use, and maintenance.

3.1 Electronic Medical Record Support for Public Health (ESP)

Objective: Use ESP to analyze clinical data to identify conditions of interest	
Description of MENDS Functionality	
• ESP is an automated software application that analyzes EHR data to identify and report conditions of interest to public health. • ESP contains two components: (1) a system to construct, load, and prepare an ESP data mart for use, and (2) functions to apply specific disease algorithms to the ESP data mart. More information about ESP is available from esphealth.org. • ESP transforms and aggregates EHR data into a format to be queried by PopMedNet and visualized by RiskScape and iVEST. • Disease algorithms use diagnostic codes, vital signs, laboratory tests, and prescribing information to identify cases and conditions, which are implemented through a software plugin. When ESP is installed at a partner site, selected ESP plugins are installed (see MENDS Plugins and Indicators on Basecamp). • ESP stores geographic information at the patient level, including the five-digit ZIP Code and census tract levels. • Plugins are periodically updated; when plugins are modified, the MENDS IT vendor updates the documentation and plugin at each partner site. • The ESP data model does not store information about specific healthcare organizations; thus, MENDS offers no reporting or comparison by healthcare organization.	
Governance Committee Responsibilities	**Partner Site Requirements**
• Discuss~~es~~ ESP issues and recommend improvements as needed	• Install and maintain a current ESP instance • Identify ESP issues and recommend improvements as needed
MENDS Project Team Responsibilities	
• Track updates made to ESP by other ESP-based efforts • Contribute MENDS-generated ESP enhancements to the public ESP codebase • Discuss relevant ESP changes with MENDS partner sites and other interested parties for consideration	

3.2 Software Installation and Training

Objective: Achieve complete installation, testing, and training of ESP and PopMedNet across all MENDS partner sites
Description of MENDS Functionality
• Software installation is a joint effort of data contributors and the MENDS IT vendor. • Data contributors are required to provide a dedicated MENDS server configured, based on ESP server specifications, for software installation. Appendix 7: Technical Requirements provides more information and defines the technical expectations of a data contributor. • The MENDS IT vendor is responsible for supporting data contributor software installation, creating and updating software materials, and providing training to the data contributor. • Either the data contributor or the MENDS IT vendor must complete the following: install ESP and PopMedNet, connect RiskScape and PopMedNet to the ESP data model, and conduct unit testing and user acceptance testing. • The MENDS IT vendor requires remote access and elevated privileges on the ESP server to install and maintain the system. • The national PopMedNet web portal application is installed and maintained within an Amazon Web Services (AWS) hosting environment, accessible only by the MENDS IT vendor. • Each data partner has the option to run its own dedicated PopMedNet web portal application and may choose to host this application using the MENDS IT vendor's AWS hosting environment or self-host with the MENDS IT vendor's support. • RiskScape can be installed and maintained either within the data partner's data center or within the MENDS IT vendor's AWS hosting environment. • Once sufficient access is granted to the MENDS IT vendor to install, administer, and maintain the ESP system, software implementation should take approximately 3 months, and completion is a requirement for the data contributor. • Partner sites are trained on ESP, PopMedNet, and RiskScape for both data contributor and data user functionality, as appropriate. Training materials are provided to ensure that partner sites can provide additional training to new users.

3.3 Source Data and the MENDS Common Data Model

Objective: Build and populate a data mart based on the MENDS common data model specification
Description of MENDS Functionality

- See Appendix 4: MENDS Electronic medical record Support for Public Health (ESP) Data Model for common data model specification.
- Source data include structured patient-level electronic health data (e.g., outpatient and inpatient EHRs, demographic information, insurance enrollment information) and the data stored in the ESP database.
- MENDS uses the ESP data model as its common data model for an infrastructure to organize, store, and analyze patient- and event-level data for use by MENDS software. The ESP data model uses data standards (i.e., Health Level 7 (HL7) messaging standards and terminology vocabularies).
- Each partner site is required to provide source data in accordance with the MENDS data model specification. The organization providing data and building the ESP data mart is known as the data contributor.
- Data contributors may submit source data from any EHR or from a data aggregator such as a health information exchange (HIE) or shared data warehouse. Source data may be contributed in the form of delimited text files, FHIR bulk data client data, or Consolidated Clinical Document Architecture (CCDA) documents.
- MENDS does not require personal identifiers such as name and contact information. However, a five-digit ZIP Code is required, and census tract is optional.
- The MENDS IT vendor is contracted to develop, implement, and maintain a procedure to extract, transform, and load (ETL) source data into the ESP data mart at each data contributor. Data contributors have the option to develop their own ETL, but this approach requires a high level of effort; when this is the case, ETL maintenance is the responsibility of the data contributor.
- The ESP data must be refreshed and backed up after each refresh (see 2.2 Data Timeliness) to push new data into ESP. Each ESP refresh cycle includes three data processing milestones: data load complete, indicator processing complete, and RiskScape reporting complete.
- Data contributors acting on behalf of multiple healthcare organizations are responsible for managing withdrawals or additions of healthcare organizations' data in the ESP data mart.
- Changes to the MENDS common data model must be approved by the Governance Committee.

Governance Committee Responsibilities	Partner Site Requirements
• Raise source data and ESP data model issues for discussion • Provide recommendations to improve MENDS functionality	• Collaborate with the MENDS IT vendor to build and maintain an ETL process/or data exchange for ongoing data-provisioning of the ESP data mart • Provide data for the ESP data mart from January 1, 2017, to present, assuming data are available • Monitor network functionality and ESP data refreshes and notify MENDS project team of EHR data issues • Conduct regular backups of ESP data mart and make backup data available when necessary

MENDS Project Team Responsibilities

- Support data contributors during implementation
- Monitor implementation progress and maintenance of partner site data

3.4 PopMedNet

Objective: Maintain PopMedNet connectivity between the data contributor and data user(s) and between the data contributor and the MENDS Coordinating Center

Description of MENDS Functionality

- PopMedNet is an open-source distributed querying tool used by MENDS to query partner site data. More information is available in the PopMedNet User's Guide (PMN) and the PMN Tip Sheet.
- PopMedNet has two software systems: (1) local and national web portal applications for data users to send query requests and compile returned results; and (2) a data mart client connected directly to the ESP data mart for providing review and approval functionality for query results.
- For MENDS data contributors, a single data mart client is installed and connected to two PopMedNet portals: a local portal for exchanging results with partner site data users and a national portal for exchanging results with the MENDS Coordinating Center.
- MENDS requires the use of the ESP-enabled fork of PopMedNet, which is maintained by the MENDS IT vendor.
- The PopMedNet data mart client may reside on an individual workstation or a remote desktop.
- Access to PopMedNet software systems is password protected and limited to authorized and authenticated users.

Governance Committee Responsibilities	Partner Site Requirements
• Recommend PopMedNet for querying MENDS data • Provide feedback from partner sites on PopMedNet • Raise issues related to PopMedNet for discussion • Recommend enhancements to PopMedNet functionality	• Identify one primary individual and one alternate to be trained to fulfill or support PopMedNet functionality • Administer access to PopMedNet web portal for data users and data contributors • Register users for PopMedNet accounts and create credentials • Maintain connectivity among the PopMedNet data mart client, ESP data mart, and the PopMedNet web portal • Report PopMedNet problems to the MENDS IT vendor

MENDS Project Team Responsibilities

- Support partner site installation and maintenance of the data mart client
- Monitor use of PopMedNet software
- Ensure PopMedNet software documentation is available for users
- Address PopMedNet software functionality issues and communicate with MENDS partners about problems and their solutions

3.5 PopMedNet Queries

Objective: Use PopMedNet queries to generate chronic disease surveillance information within MENDS partner sites and across the network

Description of MENDS Functionality

- Nationally, PopMedNet is used to query MENDS partner sites (i.e., network queries). Network queries are created by the MENDS Coordinating Center on behalf of requestors.
- MENDS queries are primarily custom SQL scripts. To build a SQL script, submitters need an understanding of the MENDS ESP system data model.
- Appendix 8: PopMedNet Request and Query Process describes the MENDS PopMedNet request and query process.
- PopMedNet queries are developed ad hoc and used to generate surveillance information.
- Although PopMedNet can be used for record level or patient level query results, only queries returning aggregate count queries are permitted in MENDS. Query results are stratified by demographic group or geographic area. The most granular geographic level permitted in query results is a five-digit ZIP Code or census tract, depending on availability at a data contributor and each data contributor's data privacy standards.
- Partner sites determine local governance and processes for querying their own ESP data mart (i.e., local queries).
- The MENDS Coordinating Center retains MENDS data (query results compiled across partner sites) only for as long as necessary to fulfill the purpose of the request. After 5 years from the date of the receipt of the query results from the partner site, the MENDS Coordinating Center will destroy MENDS data and associated query results within the PopMedNet query portal. However, MENDS data products generated from MENDS network data will not be destroyed.

Governance Committee Responsibilities	Partner Site Governance Requirements
• Provide feedback on the request and query process • Discuss past and future queries	• Identify one primary individual and one alternate to be trained to respond to PopMedNet queries delivered to the partner site through the data mart client • Provide resources to monitor the PopMedNet data mart client for queries • Respond to queries through the PopMedNet data mart client

MENDS Project Team Responsibilities

- Oversee and monitor query functionality and process
- Monitor query participation across data contributors

3.6 RiskScape

Objective: Use RiskScape to visualize surveillance information within MENDS partner sites

Description of MENDS Functionality

- RiskScape is an open-source software developed by the IT vendor.
- RiskScape is a data visualization tool that leverages the ESP data to examine public health surveillance information within each MENDS partner site for MENDS priority indicators, once validation is complete. RiskScape visualizations include graphs, charts, tables, and maps at the five-digit ZIP Code or census tract level.
- RiskScape allows users to create custom reports that include counts (numerator and denominator) and rates such as prevalence estimates.
- A separate MENDS-enabled RiskScape instance is created for each partner site for use by authorized users at the data contributor and data user organization(s).
- RiskScape may be hosted by the MENDS IT vendor or within the data contributor environment.
- Data contributors may administer access to their RiskScape instance or ask the MENDS IT vendor to administer access to their RiskScape instance on their behalf.
- If a partner site collects and aggregates patient data across multiple data owners and de-duplicates patients across the sites, RiskScape cannot stratify indicators by data owners within a partner site.
- RiskScape updates are implemented by the MENDS IT vendor. Every RiskScape update includes release notes to describe changes.

Governance Committee Responsibilities	Partner Site Governance Requirements
• Recommend RiskScape to visualize MENDS data • Provide feedback from partner sites on RiskScape • Review use of RiskScape • Recommend RiskScape enhancements	• Work with MENDS IT vendor to establish a local instance of RiskScape and administer access to users • Appoint a RiskScape site administrator • Report RiskScape problems to the MENDS IT vendor

MENDS Project Team Responsibilities

- Oversee and monitors the RiskScape functionality and process
- Ensure RiskScape software documentation is available for users
- Address RiskScape software functionality issues and communicate with MENDS partners about problems and their solutions

3.7 National Visualization Tool (iVEST)

Objective: Use a software tool to visualize weighted surveillance estimates compiled from all MENDS partner sites

Description of MENDS Functionality

- The MENDS national visualization tool was developed by the MENDS IT vendor and is named the interactive Visualization of EHR-based Surveillance Tool (iVEST).
- iVEST is hosted and maintained by the MENDS IT vendor.
- iVEST is open source. Source code and detailed technical information are available on GitLab.
- iVEST presents graphs and maps of adjusted prevalence for MENDS priority indicators at the national and state levels.
- iVEST visualizes adjusted estimates resulting from the advanced analytic work that are de-identified under HIPAA.
- iVEST access is available to authorized users; access permission is granted by the MENDS IT vendor.

Governance Committee Responsibilities	Partner Site Governance Requirements
• Recommend iVEST to use MENDS data • Provide feedback from partner sites about iVEST functionality • Review use log and discusses use of iVEST	• Share information about iVEST with interested parties • Route iVEST access requests to the MENDS project team • Report iVEST problems to MENDS IT vendor

MENDS Project Team Responsibilities

- Oversee and monitors iVEST functionality and process
- Address iVEST software functionality issues and communicate with MENDS partners about problems and their solutions

3.8 Security

Objective: Ensure the security of data and software at each MENDS partner site and across the network

Description of MENDS Functionality

- The security of patient data is the priority. MENDS's design and implementation are structured to ensure security, a priority for network functionality.
- Data contributors are responsible for maintaining the security of their ESP data mart and any MENDS software applications installed within their environment.
- The MENDS IT vendor is responsible for the security of MENDS software applications (i.e., PopMedNet and RiskScape).
- PopMedNet and RiskScape undergo security reviews at least annually (see Additional Security Information below). Updates are made to each software system based on the results of security reviews. System access is restricted using IP whitelisting or two-factor authentication.
- Access to PopMedNet and RiskScape is limited to password-protected accounts from

<table>
<tr><td colspan="2">authorized and authenticated data users. User accounts, initial passwords, and ongoing account maintenance are provided by the MENDS IT vendor.
• Because the mechanics of software implementation and data contributor environments vary, a data security plan specific to the data contributor is developed in collaboration with the MENDS IT vendor and is required to ensure the software and environment security needs are met.
• When third-party code scans or penetration testing are performed, either contracted by the MENDS IT vendor or by a data contributor, the results are shared with MENDS' IT vendor and reviewed immediately upon delivery. When findings have implications for MENDS, they are shared with the MENDS Coordinating Center and MENDS Governance Committee.
• When a security event or issue emerges, a response plan and/or remediation plan is developed and implemented by the MENDS IT vendor. When implementing a response or remediation plan, the MENDS IT vendor expects urgent and collaborative engagement by data contributors.</td></tr>
<tr><th>Governance Committee Responsibilities</th><th>Partner Site Governance Requirements</th></tr>
<tr><td>• Raise MENDS security concerns
• Discuss MENDS security issues as needed</td><td>• Create a data security plan with the MENDS IT vendor
• Maintain a secure environment for the ESP data mart
• Report security concerns or issues to the MENDS project team
• Share results from code scans or penetration tests with the Coordinating Center
• Provide resources to support responsive remediation to security events or issues</td></tr>
<tr><th colspan="2">MENDS Project Team Responsibilities</th></tr>
<tr><td colspan="2">• Monitor the security of the MENDS software and network
• Report security issues to partner sites within 48 hours of identification, including any remediation plans or network interruptions
• Work with partner sites to resolve any security issue or event</td></tr>
</table>

Additional Security Information

- The IT vendor has established an Information Security Management System (ISMS)[8] to meet SOC 2 compliance standards for its AWS hosting environment. It employs several third-party consulting organizations to perform regular penetration testing and access control testing within the hosting environment and to regularly scan the public code repositories for unauthorized use attempts or insecure code/content.
- The IT vendor has ISMS policies and procedures. Security breach event detection and handling policies and procedures are designed to address sections A.16.1.1., A16.1.5, A16.1.6, A.16.1.6, A.16.1.7. of the ISO/IEC 27001 standard, as well as §164.308(a)(6) and §164.400 to §164.414 of the HITECH Act of 2009.

[8] Definition of Information Security Management System: https://www.techtarget.com/whatis/definition/information-security-management-system-ISMS

3.9 Software Maintenance and Enhancement

Objective: Ensure that MENDS software is proactively maintained to preserve continuity of network functionality

Description of MENDS Functionality

- PopMedNet, ESP, RiskScape, and iVEST are maintained by the MENDS IT vendor, which is responsible for addressing defects, ensuring performance and functionality, performing upgrades, and adding enhancements.
- Data contributors are responsible for maintaining the servers and network upon which ESP and the PopMedNet data mart clients are installed.
- Partner site instances of ESP and RiskScape are maintained by the MENDS IT vendor.
- Maintenance includes emergency or planned maintenance. Planned maintenance is designed to be non-disruptive and scheduled for times when software use is unlikely (overnight or the weekend).
- Maintenance is typically undetectable by the users but may result in periods of downtime for MENDS software. If maintenance will cause downtime or affect performance or front-end functionality, partner sites will be notified.

Governance Committee Responsibilities	Partner Site Governance Requirements
• Discuss software performance • Recommend software enhancements	• Maintain server(s) and network for MENDS software • Report any unexpected results or bugs to the MENDS IT vendor

MENDS Project Team Responsibilities

- Ensure the resources needed to maintain MENDS software
- Compile and evaluate recommended software enhancements
- Manage implementation of approved software enhancements

Appendix 1: Glossary of Terms

The terms used throughout the document are defined below as they relate to the Multi-State EHR-Based Network for Disease Surveillance (MENDS).

- **Algorithm:** A defined set of data elements and logical expressions used to identify one or more chronic disease case definitions (sometimes referred to as e-phenotypes).
- **Aggregate data:** Counts, prevalence estimates (direct, weighted or modeled), or their derivatives (e.g., between group differences or ratios), overall or by specific population subgroups or geographies.
- **Data mart**: A database that typically represents a subset of a larger enterprise data warehouse. For MENDS, the data used to populate Electronic medical record Support for Public health (ESP) are considered a data mart.
- **Data aggregator**: An organization housing clinical data from multiple data owners, such as a health information exchange. Data aggregators can serve in the role of a data contributor in MENDS (see definition below).
- **Data contributor:** An organization constructing and maintaining an ESP data mart that is connected to PopMedNet or RiskScape for MENDS. Data contributors may be an individual health system or a data aggregator such as a health information exchange.
- **Data owner:** A healthcare organization, payer, or other organization that collects and stores clinical data in electronic health record systems and is responsible for the uses and stewardship of that data. Data owners may also be data contributors if they contribute data directly to MENDS. Alternatively, data owners may delegate authority to a data aggregator such as a health information exchange to make their data available to MENDS as a data contributor.
- **Data user:** An individual using MENDS data for data quality, validation, or surveillance. A data user may be a member of the MENDS project team, a partner site data contributor, a state or local health department, or other organization authorized by the data contributor. A partner site may have zero, one, or multiple data users.
- **Electronic medical record Support for Public health** (ESP; www.esphealth.org): An open-source software platform that organizes and maps electronic health record data, analyzes the data for conditions of public health interest, and can transmit either case reports or aggregate summaries to health departments. ESP also serves as the platform for a distributed data network that can be queried by authorized public health officials to assess conditions of interest in aggregate, in a secure and transparent fashion under the oversight and control of the data owner. (See Appendix 7: Technical Requirements for more details.)
- **Governance Committee:** A working and decision-making body responsible for developing and implementing information and software governance in MENDS.
- **Health information exchange (HIE):** An organization that provides services related to the electronic movement of health-related information among organizations according to nationally recognized standards. An HIE is also sometimes referred to as a health information network (HIN). (See term data aggregator above.)
- **Indicator:** An inclusion and exclusion criterion for the numerator and denominator to express a measure.
- **Information Technology (IT) Vendor:** An organization under contract with MENDS to implement MENDS technical infrastructure and support national and site governance and data use. MENDS currently contracts with Commonwealth Informatics.
- **iVEST:** The MENDS Project's interactive, web-based data visualization platform for modeled estimates. iVEST is an open-source product distributed under the BSD 3-Clause license. iVEST provides modeled estimates of prevalence for MENDS priority indicators.

iVEST uses the ESP data mart and partner site's individual-level, de-identified dataset, which is updated monthly to generate graphs based on condition of interest while allowing users to filter or stratify by other factors such as age, sex, race/ethnicity, health indicators, and comorbidities.

- **MENDS Coordinating Center:** Individuals and/or organizations that coordinate queries through the national instance of PopMedNet, access query results and MENDS network data, and generate MENDS data products. For MENDS, this is the National Association of Chronic Disease Directors.
- **MENDS Project Team:** Team that consists of representatives from the Centers for Disease Control and Prevention, the National Association of Chronic Disease Directors, Commonwealth Informatics, Harvard Pilgrim Health Care Institute's Department of Population Medicine, and National Opinion Research Center (NORC).
- **Partner site:** An organization or group of organizations participating in MENDS by contributing data and/or using MENDS data. Partner sites have executed a contract to receive funds for MENDS participation and a business associates' agreement to install MENDS software and share data.
- **PopMedNet:** Open-source software, distributed under the Apache License, Version 2.0. Source code for PopMedNet is posted to GitLab and is maintained by the IT vendor. This open-source repository is a fork of the main PopMedNet code line, maintained by Harvard University Department of Population medicine, and available at GitHub. The instance used by MENDS includes ESP-specific query modules no longer supported by Harvard Department of Population Medicine, which are maintained by Commonwealth Informatics. PopMedNet is a scalable and extensible open-source informatics platform designed to facilitate the implementation and operation of distributed health data networks. PopMedNet is composed of two software components:
- **PopMedNet Web Portal:** A web application for composing and distributing queries to run and return results from a set of connected PopMedNet Data mart clients.
- **PopMedNet Data mart Client:** A required MENDS component application installed on a system within the data provider's network enclave that provides a connection between the PopMedNet web portal and the ESP database. It requires an operator to receive and review queries distributed from the web portal, and if approved, to submit these queries to run against the ESP database. The ESP database returns results to the data mart client, and again an operator is required to review the results and approve them for return to the web portal. It is possible to pre-approve queries for regular distribution and execution.
- **Publicly available data:** MENDS de-identified aggregate crude and/or weighted data products that have been approved by the contributing partner sites and shared with the public (e.g., on a publicly facing website) in accordance with Appendix 6: Data Product Review and Dissemination Guidelines.
- **RiskScape:** The MENDS interactive, web-based data visualization platform. RiskScape is an open-source product distributed under the BSD 3-Clause license. Source code and detailed technical information are available on GitLab. It provides timely, high-level summaries of specific conditions of interest to public health officials. RiskScape uses the ESP data mart and partner site's individual-level, de-identified dataset, which is updated monthly to generate graphs based on condition of interest while allowing users to filter or stratify by other factors such as age, sex, race/ethnicity, health indicators, and comorbidities.
- **Weighting and estimation vendor:** An organization under contract with MENDS to contribute to modeling national and state estimates and to support national and site data use. MENDS currently contracts with NORC.

Appendix 2: Data Use Guidelines

Background

This document defines guidelines for Multi-State EHR-Based Network for Disease Surveillance (MENDS) data use and expectations of users of MENDS data. For data uses not specifically mentioned here, these guidelines can inform decisions and processes.

There are several types of MENDS data users—members of the MENDS project team, and individuals from partner site data contributors or data user organizations. The MENDS project team is charged with building and sustaining the MENDS network and promoting the use of MENDS data for surveillance at the national, state, and local levels. Partner site data users are individuals from a partner site who are authorized to use MENDS data for surveillance, such as an individual at a state or local health department. A partner site might have zero, one, or multiple data users. Partner site data use and access across the data types differ based on the partner site configuration and local governance. Although members of the public cannot have access to raw or crude source data, due to privacy concerns and confidentiality rules, they may be able to view MENDS de-identified aggregate crude and/or weighted data products that have been approved by the contributing partner sites, in adherence with Appendix 6: Data Product Review and Dissemination Guidelines.

Overview of MENDS Data Use

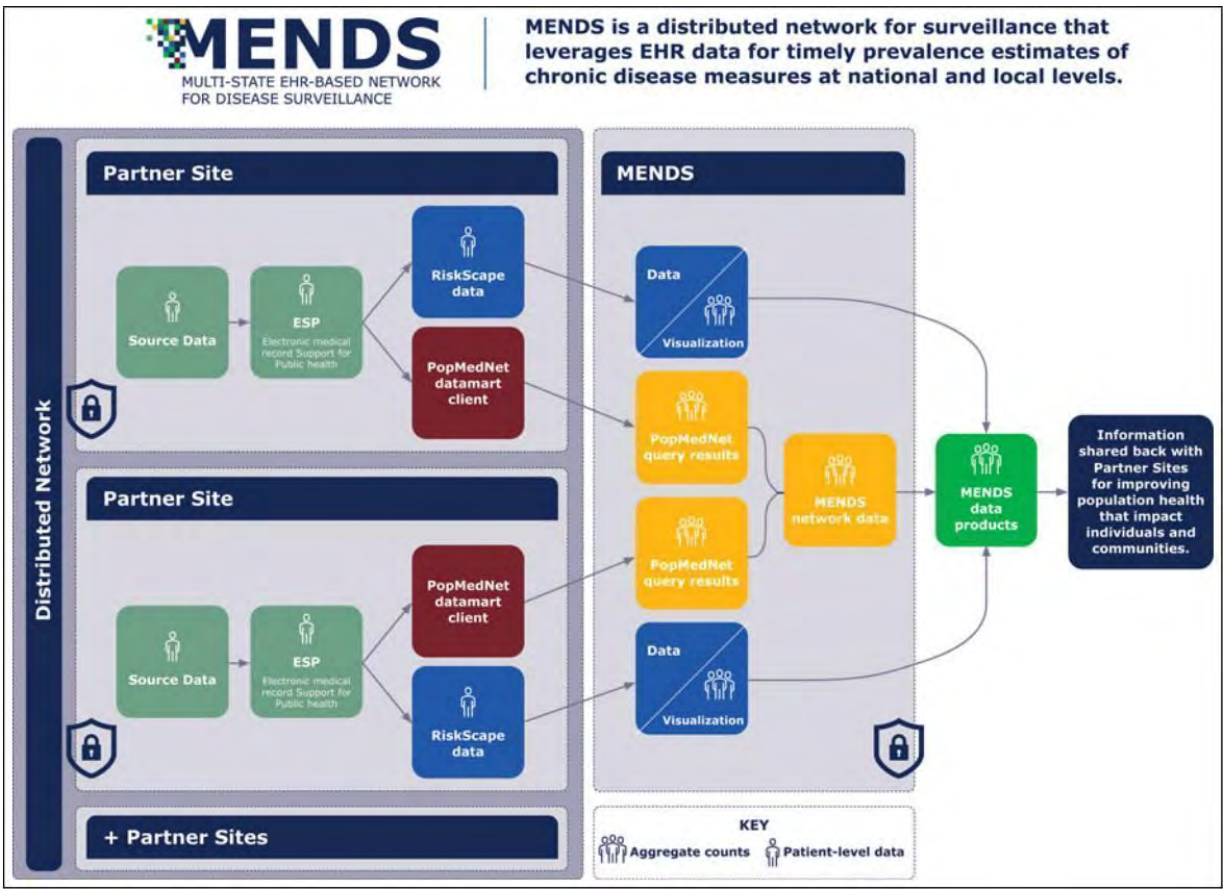
- MENDS is a distributed network built for population health surveillance.
- Partner sites control access and use of their data. Only individuals authorized by a partner site have access to RiskScape, only PopMedNet queries that a partner site permits can run against its data, and only results that the partner site approves can be released.
- MENDS data are only available for use after they have been validated, as appropriate, in accordance with the MENDS validation process. Validation is required at every partner site for each indicator. (See Appendix 5, Data Quality and Validation Process.)
- All MENDS data described in this document are health information and are subject to various privacy and security sections of the Health Insurance Portability and Accountability Act of 1996 (HIPAA). Descriptions of how data protections, including HIPAA, apply to each type of data are included in the sections for each data type below.
- HIPAA Section 45 CFR 164.512(b) permits covered entities to disclose protected health information without authorization for specified public health purposes (i.e., surveillance).
- State laws or organizational policy may apply to MENDS data and are considered local governance and are the responsibility of the partner site.
- Users receiving access to MENDS data as part of validation activities will never retain, share, or disseminate MENDS data before validation has been completed.
- MENDS does not currently support national research-related use of data drawn from multiple partners.
- MENDS data use guidelines may be amended or expanded in the future to include novel use cases based on the approval of the Governance Committee.

Types of MENDS Data

MENDS generates multiple types of data and has two software platforms to facilitate data use: RiskScape and PopMedNet (see Figure A2.1). How data can be used and who can use them differ by data type and platform. Accordingly, MENDS data use is discussed below in the following domains: source data, RiskScape underlying data, RiskScape data visualizations,

partner site PopMedNet query results, MENDS network data (from PopMedNet), and MENDS data products, both crude and weighted data products (includes iVEST as a visualization tool for a type of data product). For each data type, the data are defined as patient-level or aggregate (i.e., counts).

Figure A2.1. MENDS Technical Structure, Software, and Data Flow Diagram



Source Data

- **What are the data?** Source data include structured patient-level electronic health data (e.g., outpatient and inpatient electronic health records, demographic information, insurance enrollment information) and the data stored in the Electronic Medical Record Support for Public Health (ESP) database.
- **What data protections apply?** Source data include personally identifiable information (PII) and protected health information (PHI) stored at the patient level. Because source data reside within a partner site's secure environment, partner sites are responsible for maintaining all necessary HIPAA privacy and security protections.
- **Who can use the data?** Partner sites retain complete possession and control of their source data and use. While members of the MENDS team may have access to source data for ESP implementation, data quality activities, and validation activities, this access is controlled by HIPAA-required business associates agreements (BAAs), and MENDS team members are not permitted to use source data for any other purposes.
- **How can the data be used?** Source data are accessed by PopMedNet and RiskScape for uses described in the subsequent sections. Partner sites may use their own source data for non-MENDS purposes if they comply with applicable state and federal laws and

regulations, including HIPAA[9] and the Common Rule[10].
- Data considerations: None.

RiskScape Underlying Data

- **What are the data?** The RiskScape underlying data are patient-level data (i.e., one row per patient per month).
- **What data protections apply?** RiskScape underlying data are obscured by replacing continuous variables with categorical variables (i.e., BMI >30 rather than BMI = 33) and by substituting variable names with anonymous terms (a, b, c, etc.) but do contain PHI (five-digit ZIP Code, census tract, and the date elements month and year). As such, RiskScape data constitute a limited dataset. Therefore, any direct use of the raw data provisioned to RiskScape requires a data use agreement (DUA) or BAA.
- **Who can use the data?** RiskScape underlying data are used by the MENDS IT vendor to provision RiskScape data visualizations. RiskScape underlying data are used by the weighting and estimation vendor to generate weighted and model-based prevalence estimates. Partner sites may use their own RiskScape underlying data for non-MENDS purposes if those purposes align with their local governance policies.
- **How can the data be used?** See statement above under the heading "**Who can use the data?**"
- **Data considerations:** None.

RiskScape Data Visualizations

- **What are the data?** RiskScape is a visualization platform that provides aggregate counts and prevalence estimates.
- **What data protections apply?** Aggregate data are generally considered de-identified if the data cannot be used, alone or in combination with other reasonably available information, to identify an individual who is the subject of the information. Aggregate data with identifiers (e.g., five-digit ZIP Code, census tract) can be considered a limited dataset, and use can require a BAA or DUA between the data contributor and data user organization.
- **Who can use the data?** RiskScape is accessible only by authorized and authenticated (e.g., password-protected) users. Authenticated users are granted access and managed by the partner site. Partner sites provision their own RiskScape access and access to MENDS project team users if requested. Processes to grant RiskScape access differ based on the partner site configuration and local governance.
- **How can the data be used?** RiskScape is designed for surveillance and provides visual tools for users to map prevalence by geographic area (e.g., state, county, three-digit or five-digit ZIP Code), obtain sociodemographic group-specific estimates, and evaluate temporal trends. Partner sites may use RiskScape for non-MENDS purposes if those purposes align with their local governance policies.
- Data considerations:
    - RiskScape is configured to suppress small cells according to the MENDS small cell suppression guidelines. (See 2.4 Data Suppression and Privacy)

Partner Site PopMedNet Query Results

[9] Health Insurance Portability and Accountability Act of 1996 (HIPAA) for professionals: https://www.hhs.gov/hipaa/for-professionals/index.html

[10] Federal Policy for the Protection of Human Subjects ('Common Rule'): https://www.hhs.gov/ohrp/regulations-and-policy/regulations/common-rule/index.html

- **What are the data?** Partner site PopMedNet query results are aggregate counts of crude data generated from PopMedNet queries. The stratifications of query results are determined by the query creator and may include age group, sex, racial ethnic group, five-digit ZIP Code, or census tract.
- **What data protections apply?** Aggregate data are generally considered de-identified if the data cannot be used, alone or in combination with other reasonably available information, to identify an individual who is the subject of the information. Aggregate data with identifiers (e.g., five-digit ZIP Code, census tract) can be considered a limited dataset and can require a BAA or DUA between the data contributor and data user organization.
- **Who can use the data?** The MENDS Coordinating Center uses partner site PopMedNet query results to generate MENDS network data (see next section of these guidelines). Partner sites can use their own PopMedNet query results for surveillance or other purposes.
- **How can the data be used?** Partner site query results are used to generate MENDS network data.
- **Data considerations:**
    - PopMedNet allows partner sites to review and approve queries before they can be executed and gives partner sites the option to review query results before they are released back to the query requestor.
    - Use and users of query results are specified in the PopMedNet query description that is provided to partner sites prior to query submission and included with the query. Partner sites evaluate these criteria when determining query participation.
    - Partner site PopMedNet query results can be reused for other public health purposes if the secondary application is within the original intent of the query defined in the query description.

MENDS Network Data

- **What are the data?** MENDS network data are query results generated from PopMedNet that have been aggregated across partner sites. Query results include aggregate counts by geographic and demographic group.
- **What data protections apply?** Aggregate data are generally considered de-identified if the data cannot be used, alone or in combination with other reasonably available information, to identify an individual who is the subject of the information. Aggregate data with identifiers (e.g., five-digit ZIP Code, census tract, or ages greater than 90 years) can be considered a limited dataset and can require a BAA or DUA.
- **Who can use the data?** The data can be used by authorized users of the MENDS PopMedNet query creator who have rights to submit PopMedNet queries and use query results to generate data products for themselves or other MENDS partners.
- **How can the data be used?** The MENDS Coordinating Center uses the MENDS network data to derive prevalence estimates and related precision measures and generate data products.
- **Data considerations:**
    - Use of MENDS network data is specified in the PopMedNet query description that is provided to partner sites prior to query submission and included with the query.
    - Periodically, healthcare organizations terminate their participation with data aggregators. When this occurs within a MENDS partner site, MENDS network data that reflect patient data from a data owner that is no longer a MENDS participant can still be used for their intended purpose. Data products generated

from those query results are still available for use.
    - MENDS network data can be reused for other public health surveillance purposes if the secondary application is within the original intent of the query defined in the query description. When reuse for other purposes leads to data products, the products follow the same review process outlined in MENDS Review and Dissemination Guidelines (See Appendix 6: Data Product Review and Dissemination Guidelines).

MENDS Data Products

- **What are data products?** MENDS data products are based on aggregate data and can reflect national or local estimates. National and local products can be generated by the MENDS Coordinating Center, and other local data products can be generated by partner sites. Data products can be crude or weighted. Products include summaries of RiskScape and MENDS network data from PopMedNet (e.g., slides, maps, graphs, reports), analytic results derived from those data (e.g., weighted and model-based prevalence estimates), and weighted estimates presented in iVEST, the national visualization tool.
    - Weighted prevalence estimates are calculated using statistical tools based on both MENDS data and auxiliary population data (e.g., census sociodemographic data) to minimize bias and control for differences in socio-demographic attributes.
- **What data protections apply?** Because MENDS data products are intended for dissemination, they must be de-identified. Weighted prevalence estimates, including those at the five-digit ZIP Code level, are de-identified because they are derived based on information from a large number of persons using statistical procedures and cannot be linked to specific individual persons. Such aggregate data are generally considered de-identified if they cannot be used alone, or in combination with other reasonably available information, to identify an individual who is the subject of the information. To ensure de-identification, MENDS data products undergo a de-identification review process prior to release. De-identification is accomplished using one of the de- identification methods provided by the HIPAA Privacy Rule: Safe Harbor or Expert Determination[11]. The attributes of each specific data product determine which method is used.
    - Data products that do not include the 18 identifiers follow the Safe Harbor method to be designated as de-identified.
    - Data products for which the 18 types of identifiers cannot be removed are reviewed for de-identification by a committee of experts (i.e., Expert Determination method) as recommended by the HIPAA Privacy Rule. (See also Appendix 6: Data Product Review and Dissemination Guidelines.)
- **Who can use data products?** Any MENDS data contributor, data user, data owner, or member of MENDS project team can use the released crude and weighted data products. Publicly released de-identified aggregate and/or weighted data products could be used by the public. (See Appendix 1: Glossary of Terms for details on any of these roles.)
- **How can data products be used?** Data products are subject to review processes and dissemination guidelines.
- **Data considerations:**

[11] Guidance Regarding Methods for De-identification of Protected Health Information in Accordance with the Health Insurance Portability and Accountability Act (HIPAA) Privacy Rule: https://www.hhs.gov/hipaa/for-professionals/special-topics/de-identification/index.html

- MENDS data products reflect contributed data, software and visualization tools, and subsequent analytic activities. Therefore, national data products are jointly owned by the partner sites that contributed data and the MENDS project team. Locally generated data products reflect an effort of partner sites, and, accordingly, this work will be attributed as a product of the MENDS infrastructure.
- Small cells are suppressed according to the MENDS small cell suppression guidelines. (See 2.4 Data Suppression and Privacy.)
- Prevalence estimates with an unsatisfactory precision (i.e., very wide 95% confidence intervals), either weighted or model-based, at various geographic levels, will not be released to the public. The criteria will be determined by the Governance Committee in consideration of industrial standards and public health program needs.
- Partner sites are guaranteed an opportunity to review and comment on data products if partner site data was contributed.
- Once data products have been approved by partner sites, data products can be reused for other public health purposes.

Data User Guidelines

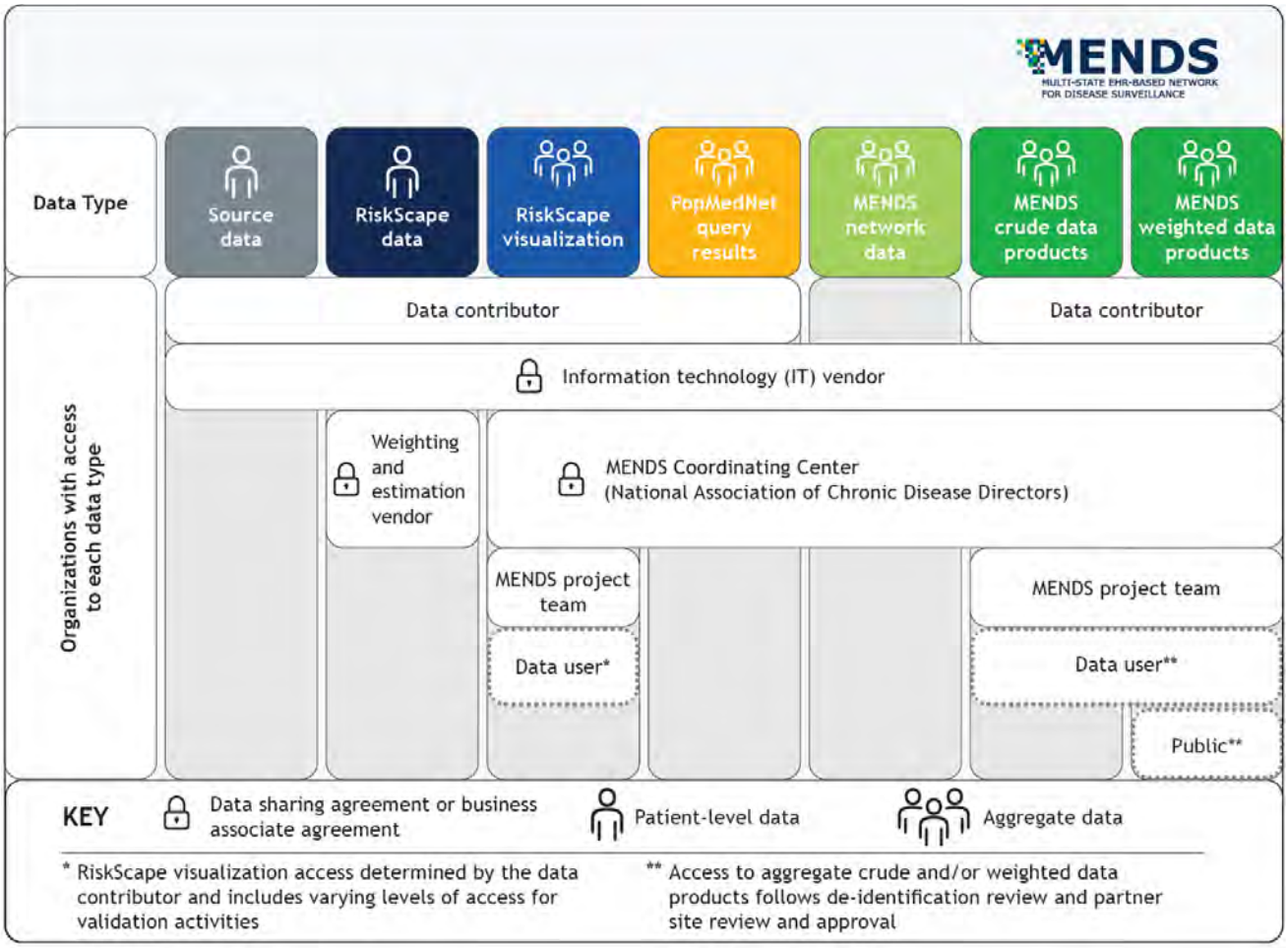
Based on the above descriptions of the different kinds of data, the following guidelines are defined for data users who may include but are not limited to individuals from local, state, or national public health organizations, healthcare systems, payers, and public health institutes. Data users:

- Use MENDS data to enhance public health surveillance, decision-making, and practice
- Use MENDS data only for intended public health purposes agreed upon by partner sites and/or the MENDS Governance Committee
- Adhere to these data use guidelines and MENDS policies and procedures, as defined in the MENDS governance documents
- Are vigilant in protecting patient, provider, health system, and partner site privacy, based on local governance and partner site preferences.

Access to Data Types

Who has access to the different data types depends on the data type in question and the legal agreements that may be in place. Figure A2.2 depicts who has access to what data within MENDS.

Figure A2.2. MENDS Data Access Diagram


MENDS
MULTI-STATE EHR-BASED NETWORK FOR DISEASE SURVEILLANCE
Data Type
Source data
RiskScape data
RiskScape visualization
PopMedNet query results
MENDS network data
MENDS crude data products
MENDS weighted data products
Organizations with access to each data type
Data contributor
Data contributor
Information technology (IT) vendor
Weighting and estimation vendor
MENDS Coordinating Center (National Association of Chronic Disease Directors)
MENDS project team
MENDS project team
Data user*
Data user**
Public**
KEY
Data sharing agreement or business associate agreement
Patient-level data
Aggregate data
* RiskScape visualization access determined by the data contributor and includes varying levels of access for validation activities
** Access to aggregate crude and/or weighted data products follows de-identification review and partner site review and approval

Appendix 3: MENDS Weighting and Modeling Summary

Rationale

Multi-State EHR-Based Network for Disease Surveillance (MENDS) leverages electronic health record (EHR) data from partner sites in various locations in the United States to support timely, detailed, and reliable public health surveillance for priority conditions. The large size, long-term repeated measurements, and rich clinical detail of MENDS data support novel in-depth analyses and timely prevalence information on priority chronic diseases and conditions at the national, state, and sub-state levels. Although the clinical data leveraged by MENDS are robust, the procured source data are not statistically representative of the United States, or even of the regions from which they are drawn, for two reasons. First, the populations covered by partner sites do not reflect the total populations living in their respective areas, and second, the selection of partner sites is non-random. Therefore, the MENDS weighting and estimation vendor uses statistical procedures to generate adjusted prevalence estimates that more accurately reflect the underlying populations and produce estimates for small geographic regions and populations that are underrepresented in MENDS.

Approach

The primary approach is statistical weighting that uses raking (or iterative proportional fitting) to adjust weights to population benchmarks from the American Community Survey (ACS)[12]. This approach reduces potential selection and coverage bias by adjusting for systemic differences in sociodemographic profiles between the MENDS patient population and the underlying population for the specific geographic area. A similar weighting approach is used both at the national and the state-level. For smaller geographic areas (for example, select counties and ZIP codes) a modeling technique known as small area estimation is used. This method integrates auxiliary data from the ACS with MENDS patient population data and is specifically designed to produce reliable estimates in the presence of sparse data.

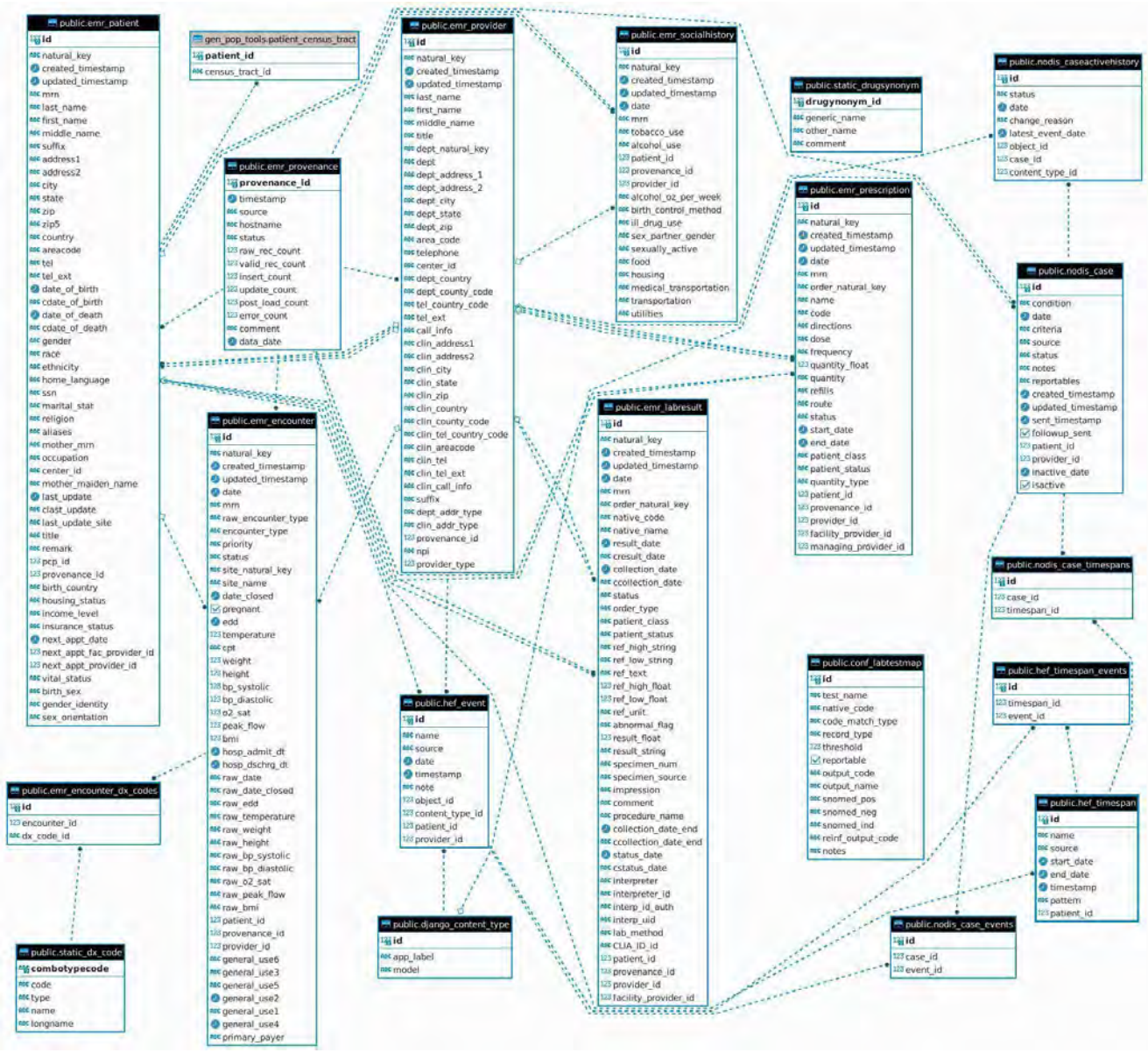
MENDS strives to make weighting and modeling methods uniform across partner sites, indicators, and geographic levels as much as possible. However, when appropriate, methods are customized to meet the unique needs of each partner site's data and each indicator. Initially, the MENDS project focused on data associated with hypertension and hypertension control. More detail about weighting and modeling methods is discussed in the Manual of Modeling Procedures (available on request).

[12] American Community Survey: https://www.census.gov/programs-surveys/acs.html

Appendix 4: MENDS Electronic medical record Support for Public Health (ESP) Data Model

The ESP data model used by the MENDS network is shown below.

Figure A4.1 MENDS ESP Data Model



Appendix 5: Data Quality and Validation Process

Summary

As part of the installation and configuration of MENDS data systems at partner sites, the MENDS project team validates the underlying data and the performance of all active indicators through a five-stage validation process. Stages 0, 1, and 2 validations are performed on the source data to measure overall quality. Stages 3 and 4 validations are performed on each indicator. For each indicator, a number of tests are performed for internal and external validation. The processes described in this appendix are designed to assure a uniformly high level of data quality across MENDS partner sites at the point of onboarding and, as appropriate, throughout the period of participation in the system.

Stage 0—Data confirmation validation

Data quality work begins with Stage 0, which was added to the MENDS validation process in 2022. This stage examines the suitability of the patient data at the data contributor for use within the MENDS network. This effort is completed by the data contributor against the source data before any data are moved to Electronic medical record Support for Public Health (ESP). Data are confirmed in the following domains: patient, visit/encounter, medication, laboratory, and social history. Data contributors receive a document that outlines a set of data investigations to be completed and shared with the MENDS project team. Findings are discussed with the data contributor, and if gaps are identified, options for moving forward are discussed.

Stage 0 Tasks

Testing task	Performed by	Anticipated duration of task	Data will be shared with
Complete data investigations on source patient data	Data contributor	1 month	IT vendor, MENDS Coordinating Center
Review report results and provide feedback as necessary	MENDS Coordinating Center	1–2 weeks	IT vendor, data Contributor

Stage 1—Testing the ESP installation

Testing follows ESP data mart installation and data provisioning. To determine whether the ESP system is storing a valid representation of the source data a series of tests are performed to confirm that all data processing steps are performing as expected. These tests are a type of internal validation. They occur at the time of ESP installation and implementation at each site and are overseen by the IT vendor. Tests include visual inspection of installation logs, review of logs from initial processing of historic data, as well as running SQL to provide basic characterization of the data. To as great an extent as it can, the IT vendor compares the results it obtains against the data source and corrects any processing errors that can be attributed to installation or configuration issues.

Each site receives its own ESP Data Extract, Transform, Load Testing Guide and Report. This report includes information on the historic data that were loaded into ESP and allows sites to check that the data are as expected (e.g., counts of patients, counts of visits, counts of

laboratory tests without a result, counts and basic temporal distribution of various critical variables, etc.). Partner sites and the IT vendor work together to resolve any issues identified during this initial stage.

Stage 1 Tasks

Testing task	Performed by	Anticipated duration of task*	Data will be shared with
Review logs to ensure installation and configuration did not generate errors	IT vendor	1–2 weeks	IT vendor, MENDS Coordinating Center, and partner site
Run SQL scripts to produce basic data characterization; review results and compare against source information, where available, to ensure data provided is loaded correctly	IT vendor	1–4 weeks	IT vendor, MENDS Coordinating Center, and partner site
Generate ESP Data Extract, Transform, Load Testing Guide and Report	IT vendor	<1 week	IT vendor, MENDS Coordinating Center, and partner site
Review report results and provide feedback as necessary	Partner site	1–2 weeks	N/A

*These are estimated timeframes. If significant issues are detected, additional time may be necessary.

Stage 2—Characterizing the data

Once there is a high level of confidence that the ESP system is acquiring and loading data correctly, a more detailed set of SQL queries are run against the data to characterize the distribution of numeric results and the frequency distributions of result categories—namely, data elements that are used in the indicators being implemented (e.g., blood pressure for hypertension-related indicators; See MENDS plugins and indicators). For numeric results, this process includes identification of high and low values; percentile ranges and counts at each percentile; the proportion above or below specified thresholds; as well as the proportion of missing and null results, etc. For categorical results, the set of categories and the frequency of each is produced. To the greatest extent possible, these distributions are assessed at a granular level (such as blood pressure distribution stratified across clinics, geographic regions, or partner sub-sites). If data anomalies are observed, the MENDS project team investigates with the partner site to determine an appropriate resolution.

This work is done to identify outliers and any data problems that may exist in the source patient health records. Some of these tests may uncover data processing errors not identified during Stage 1, but the primary purpose of these tests is to answer the question, "Does the source data provide valid indicators of clinically meaningful population parameters for epidemiology?" For example, if a large proportion of patients had a body-mass index >90, or a significant number of patients had systolic blood pressure values <70 mmHg or >250 or diastolic blood pressure values <50 or >150, and these were values as recorded with the

patient medical record, the IT vendor would need to consider how to qualify or exclude results based on these data elements in the MENDS network data. The tests to characterize the data are well defined and automated—and can be updated so that the IT vendor can monitor the data quality over time and adjust the tests as needed. If anomalies in the data are identified, a determination of the suitability of that data for a specific surveillance purpose must be made on a case-by-case basis. This work informs the implementation of acceptable value sets to be applied to each indicator. If a value falls outside of this value set, it is flagged (not deleted) and this is communicated to the data contributor for awareness and potential correction (i.e., support data quality assurance process). Furthermore, the IT vendor implements automated longitudinal checks of the data to assess for anomalies (e.g., year-over-year change in values above a certain threshold), with the understanding that some anomalies may reflect meaningful public health changes, changes in the underlying population assessed, changes in clinical practice or changes in clinical partners' operations, or data quality issues. The alerts generated during these routine checks of the system need to be assessed by an epidemiologist, or another expert with technical and data analytics expertise, and sometimes brought to the attention of the data contributors to determine the underlying cause.

Stage 2 Tasks

Testing task	**Performed by**	**Anticipated duration of task**	**Data will be shared with**
Write and run SQL scripts to generate data characterization reports	IT vendor	1–2 weeks	IT vendor, MENDS Coordinating Center and partner site
Review and investigate results to confirm that anomalies are not caused by ESP system data processing errors	IT vendor	1 week	IT vendor , MENDS Coordinating Center, and partner site
Manage changes to data characterization scripts to keep current	IT vendor , support by MENDS Coordinating Center	Ongoing	N/A
Generate data visualizations for the data characterization reports	MENDS Coordinating Center	1–2 weeks	IT vendor , MENDS Project Team, and partner site
Review data characterization reports and visualizations, identify problem data, develop plans to account for data problems*	MENDS Coordinating Center Partner site, CDC	2–4 weeks	IT vendor, MENDS Project Team, partner site

The characterization report and accompanying figures are shared with the MENDS project team. Any further sharing of the reports and figures must be with permission/approval of the site leads.

Stage 3—Indicator algorithm internal validation

The defining feature of the ESP system is its ability to process data in patient health records and identify disease conditions of interest. The algorithms for detecting these conditions are called *plugins,* and the condition detection process may generate one or more health *indicators*. Internal algorithm validation is performed to answer the question, "Is the ESP system's algorithm code correctly identifying conditions as specified?" Tests for this validation may be

performed at each MENDS partner site for each indicator, either when the indicator is built or when a partner site joins MENDS and indicators are implemented at the new site.

The tests to validate indicator algorithms rely on review and confirmation of identified conditions. The tests require review of the site's electronic health record (EHR) data by conducting a targeted review of each priority indicator. The following is an example involving hypertension:

- From ESP, select 20 random cases for each specific indicator (e.g., 20 *controlled hypertension*, 20 *uncontrolled hypertension*, and 20 *no hypertension*) from an appropriate population (i.e., if the indicator is for adults, then choose randomly among all adults). The IT vendor writes the code to generate this line list.
- Using the ESP case identification algorithm, someone at the site (or an external, authorized user) reviews each selected patient's electronic clinical history in the *source data* to determine whether the ESP case identification algorithm has been applied correctly. (Source data can be a health information exchange database or EHR data from the clinic; the site determines the source data it has access to for this purpose.) This test ensures that patient clinical data, as represented in ESP, are a correct and complete representation of the data in the source system.
    - Documentation of this work should include a summary of what was done and a summary of what was found, accompanied by a table quantifying and detailing the cases reviewed and the findings.
- If the percentage of cases that are true positives is less than 90%, the test fails. If the test fails, the disqualifying data from the source data system must be examined against the ESP system's data, and any systemic discrepancies resolved. The process is then repeated with sample replacement records for those that failed.
    - The data contributor and the IT vendor are involved with troubleshooting data issues identified.
    - Note that if it is apparent that, after fewer than 20 reviews, the true positive rate will be <90%, then this information can be fed back to the MENDS project team before 20 reviews have been completed to make interim corrections.
    - If the true positive rate is between 90% and 99%, the indicator is validated for Stage 3, but the misidentified cases are reviewed, as necessary.
    - If concordance is 100%, then the indicator is validated for Stage 3.

Stage 3 Tasks

Testing task	**Performed by**	**Anticipated duration of task**	**Data will be shared with**
Specify the condition of interest, the cohort of patients from which to sample, and the period for which the sample is taken; identify the data fields to be provided, along with the patient IDs for the data listing	MENDS Coordinating Center	1 week	N/A; plan is shared with partner site and MENDS project team
Generate the SQL to randomly select the patients matching the criteria; generate the specified listing	IT vendor	1–2 weeks	IT vendor , partner site

Using the algorithm specification, review the patient listing against the source data and confirm or negate correct identification	Partner site-based clinician, analyst, or informatician	2–3 weeks	IT vendor, MENDS Coordinating Center, partner site
Correct for any identification errors found, reiterate listing process	IT vendor	Depends on need for iterations	IT vendor, MENDS Coordinating Center, partner site
Produce summary report of results when stage is complete (no individual level data are included in the report)	MENDS Coordinating Center	1 week	Partner site, MENDS project team

Stage 4—Indicator algorithm validation

This stage answers the question: "Are the algorithms used by the ESP system providing valid estimates of the conditions being studied?" These tests validate the algorithm itself, whereas Stage 3 tests the validity of the implementation of the algorithm. There are two approaches to Stage 4.

Stage 4A: Clinical Validation—*applicable in some circumstances*

For some conditions, it may be possible for sites to generate a list of patients that they have independently identified as having the condition of interest—this list can then be compared with the ESP list of patients with the condition identified by the algorithm.

This approach to validation focuses on understanding causes for discrepant cases that were flagged by one process (e.g., ESP) but not by the other (e.g., site's master list of patients with condition X) and vice versa. This can be done via chart review if the added level of granularity and understanding is desired or can be left as a quantitative summary alone, particularly if discrepant cases are rare relative to concordant cases. Positive predictive value and sensitivity can be calculated—acceptable thresholds for each need to be assessed on a condition-by- condition basis.

This method is relatively resource heavy if the team decides to include chart review. This stage of validation is not included for the five indicators being implemented in MENDS.

The following table summarizes the tasks for this stage (initial steps are identical to Stage 3 testing). Please note that details on timeframe will be added when this stage of validation is used in the future.

Stage 4 Tasks

Testing task	Performed by	Data will be shared with
Specify the condition of interest, the cohort of patients from which to sample, and the period during which the sample will be taken; identify the data fields to be provided, along with the patient IDs for the data listing	MENDS Coordinating Center	N/A; plan is shared with partner site and MENDS project team
Generate the SQL to randomly select the patients matching the criteria; generate the specified listing	IT vendor	IT vendor, MENDS Coordinating Center, partner site
Identify a comparable condition as identified by the clinical site; review the ESP identified patients and determine whether they have the condition identified by the clinical site	Partner site-based clinician, analyst, or informaticist	IT vendor, MENDS Coordinating Center, partner site
Generate statistics of agreement and disagreement as specified; assess threshold of agreement	MENDS Coordinating Center	Partner site, MENDS Project Team
Produce summary report of results when stage is complete (no individual level data are included in the report)	MENDS Coordinating Center	Partner site, MENDS Project Team

Stage 4B—Validation of algorithm by comparison with external data sources of the same indicator/condition

When an external data source(s) is available, a simple comparison of ESP results to other sources should be possible (e.g., Behavioral Risk Factor Surveillance System (BRFSS), National Health and Nutrition Examination Survey (NHANES), other literature, MDPHnet results), considering differences in the data sources, populations, how the conditions were specified, etc. High-level agreement of this nature may be adequate in some circumstances. This work is led by the MENDS Coordinating Center.

The specific tasks for this activity will vary and will be developed and shared with site and CDC leads prior to implementation, along with timeframes and with whom data will be shared. Only aggregate level data can be used for this work.

Effort and funding permitting, the analysis of this sort of data can be extensive. For example, for each partner site, select the states with the highest representation of records for the indicator being evaluated. (i.e., if the indicator is for children, choose the state in the partner sites' data set that has the most children).

A summary of tasks for this stage could include:

- Select a comparison variable within a dataset, (e.g., BRFSS, NHANES) with CDC approval. Note, good comparisons may not exist for all indicators/conditions. Other state-based estimates may be more appropriate than BRFSS, such as those from all-payor claims.
- Using BRFSS (or other source), compute state, county, or regional prevalence estimates for the selected measure.
- Using BRFSS (or other source), compute prevalence estimates by demographic categories (age, gender, race/ethnicity).
- Discard all estimates where BRFSS' (or the other source's) 95% confidence intervals extend beyond the zero point or when the width of the 95% is greater than a predetermined percentage of the estimated prevalence.
- Correlate the comparable geographic and demographic prevalence estimates for MENDS with those from the BRFSS (or other source) for the state(s) in question. Correlation must be significant to consider the indicator validated.
- Note that comparisons of estimates between EHR-based surveillance versus BRFSS specifically need to be interpreted with considerable nuance given that neither BRFSS nor the EHR is a perfect reference standard. Both are subject to error due to the populations included (e.g., people who elect to participate in interviews, care seeking behavior in clinical samples); variability in self-report by patients; variability in testing and diagnosis patterns by providers; and differences in sample size and breadth. Discrepancies are not clear evidence that either source is 'wrong' but need to be interpreted in context.

Documentation—Documentation is applicable to all stages of validation. An example approach and template for Stage 3 algorithm validation is shown below.

A summary of technical validation findings is compiled by the partner site or person authorized to conduct analysis for each MENDS indicator.

A. <u>Create a Patient-Level Line List</u>
    1. Information that may be helpful to include:
        a. Reference IDs
            i. ESP ID
            ii. Patient ID from data contributor system
        b. Date of identification
        c. Reason case was identified in ESP (which criteria did they meet?), if relevant
        d. Notes on what was learned about a case during validation
    2. Example of what this table might look like:

Case Status	ESP ID	Patient ID	Date Identified	ESP Criteria	Notes
Identified by ESP and is a true case	1234	9101	1/1/2017	A	N/A
Identified by ESP, but not a case	1112	1516	1/2/2017	B	Lab was mapped incorrectly in ESP. Mapping was fixed, and case would no longer be identified by ESP
Controlled in ESP, but uncontrolled in EMR	2930	3334	1/15/2017	C	Note in EMR states...
Not treated in ESP, but treated in EMR	3536	3940	1/20/2017	D	Note in EMR states...

B. Create a Validation Summary
   1. Using the patient-level line list, put together a summary of findings from the validation.
   2. The following information should be included in a summary report:
      a. Number of cases identified
      b. Time period in which cases were identified
      c. Number of cases identified by ESP that were true positives
      d. Number of cases identified by ESP that were false positives
         i. Include reasons that cases were identified by ESP
      e. Number of cases where the ESP classification (e.g., acute versus chronic, active versus latent) was incorrect, if applicable
         ii. Include reasons ESP misclassified these cases
      f. Any other findings that should be brought to the MENDS team's attention

Note, if Stage 4a validation is conducted, the number of cases not identified by ESP that should have been (i.e., false negatives) should be reported and described, including reasons these cases were missed by ESP.

Appendix 6: Data Product Review and Dissemination Guidelines

Background

This document provides guidelines for the review and dissemination of Multi-State EHR-based Network for Disease Surveillance (MENDS) data products, including maps, graphs, reports, slide presentations, peer-reviewed manuscripts, conference presentations, and other outputs for the dissemination of project activities and results.

The following topics are addressed below:

- De-identification review
- Partner site review process
- Data product guidelines, inclusive of style recommendations, acknowledgment, and any guidelines specific to a particular data product.

Figure A6.1: MENDS High-Level Data Product Review and Dissemination Process

De-identification Review

All MENDS data products must be de-identified and undergo a review process prior to release. De-identification is executed using one of the de-identification methods provided under the Privacy Rule: Safe Harbor or Expert Determination. The specific data product determines what method is used and who is responsible for de-identification. MENDS national and some local data products are created by the MENDS Coordinating Center, and additional local data products can be created by local partner sites. The National Association of Chronic Disease Directors (NACDD) is responsible for conducting the de-identification review for national data products, while partner sites are responsible for conducting the de-identification review for locally generated data products. Data products must be subjected to the Safe Harbor method to be designated as de-identified if they do not include any of the 18 Health Information Portability and Accountability Act (HIPAA) specified identifiers and when no residual information can identify specific individuals. Data products for which any of the 18 HIPAA-specified identifiers cannot be removed are reviewed by a panel of experts (i.e., Expert Determination method under the HIPAA Privacy Rule) to confirm that the data used in the product has incorporated sufficient obfuscation measures for the dataset to be considered de-identified. The Expert Determination method provides for an individual to be determined as an expert in de-identification through professional experience, academic or other training, and actual experience, using health information de-identification methodologies.[13] For MENDS, the primary application of Expert Determination is for MENDS data products that contain a five-digit ZIP Code or census tract.

This method of de-identification of protected health information requires a HIPAA-covered entity or business associate to obtain an opinion from a qualified statistical expert that the risk of re-identifying an individual from the data set is very small. The principles and rationale used to make that determination and justification of the expert's opinion must be documented and

[13] HIPAA Expert Determination De-Identification Method. The Network for Public Health Law. www.networkforphl.org/wp-content/uploads/2020/01/De_Identification_HIPAAExpertDetermination_CT.pdf

retained by the covered entity or business associate and made available to data contributors.

The expert must be a person with appropriate knowledge and experience in using generally accepted statistical and scientific principles and methods for removing or altering information to ensure that it is no longer individually identifiable. When those methods and principles have been applied, the expert must determine that the risk of re-identification of an individual is very small. HIPAA does not define the level of risk of re-identification other than to say it should be "very small." The expert should define "very small" in relation to the context of the dataset, the specific environment, and the ability of an anticipated recipient to be able to re-identify individuals. Experts may come from different fields and do not require any specific qualifications. What is important is that the experts have experience in de-identifying data. It is that experience that regulators look at in the event of an audit, not specific qualifications or certifications.

The MENDS expert panel is managed by the MENDS Coordinating Center and has at least two MENDS project team members for each review. Upon request from partner sites, one additional individual, external to National Association of Chronic Disease Directors (NACDD), can be added to the expert panel from an organization with an appropriate legal agreement to view a limited dataset in place. The members must have the data protection knowledge and experience to determine whether the risk of re-identification is indeed very small and that the information cannot be used, alone or in combination with other reasonably available information, to identify an individual who is a subject of the information. The expert panel uses the guiding principles of replicability, data source availability, and distinguishability to assess risk in its review. Additional details on these guiding principles can be found in the table below.

Principles used by experts in the determination of the identifiability of health information.[14]

Principle	Description	Examples
Replicability	Prioritize health information features into levels of risk according to the chance it will consistently occur in relation to the individual.	*Low:* Results of a patient's blood glucose level test will vary
		High: Demographics of a patient (e.g., birth date) are relatively stable
Data source Availability	Determine which external data sources contain the patients' identifiers and the replicable features in the health information, as well as who is permitted access to the data source.	*Low:* The results of laboratory reports are not often disclosed with identity beyond healthcare environments.
		High: Patient name and demographics are often in public data sources, such as vital records—birth, death, and marriage registries.
Distinguishability	Determine the extent to which the subject's data can be distinguished in the health information.	*Low:* It has been estimated that the combination of *Year of Birth, Gender,* and *3-Digit ZIP Code* is unique for approximately 0.04% of residents in the United States. This means that very few residents could be identified through this combination of data alone.
		High: It has been estimated that the combination of a patient's *Date of Birth, Gender,* and *5-Digit ZIP Code* is unique for more than 50% of residents in the United States. This means that more than half of U.S.

[14] Guidance Regarding Methods for De-identification of Protected Health Information in Accordance with the Health Insurance Portability and Accountability Act (HIPAA) Privacy Rule https://www.hhs.gov/hipaa/for-professionals/special-topics/de-identification/index.html

		residents could be uniquely described just with these three data elements.
Assess Risk	The greater the replicability, availability, and distinguishability of the health information, the greater the risk for identification.	*Low:* Laboratory values may be very distinguishing, but they are rarely independently replicable and are rarely disclosed in multiple data sources to which many people have access.
		High: Demographics are highly distinguishing, highly replicable, and are available in public data sources.

Every MENDS data product de-identification review process must be documented and retained by the MENDS Coordinating Center. The expert review panel must reach a unanimous decision that the data product is de-identified and may require revisions and multiple rounds of review to reach a decision.

De-identification Review Process

1. The data product and review process are logged and issued a unique identifier by the MENDS Coordinating Center for archiving and referencing.
2. The MENDS Coordinating Center determines whether the Safe Harbor or Expert Determination approach will be used for de-identification. Safe Harbor is used for everything except MENDS data products that contain five-digit ZIP Codes.
3. If the Expert Determination approach is warranted, the MENDS Coordinating Center communicates, in writing, with panel members requesting a de-identification review. Panel members are from entities that have legal agreements (business associate agreement (BAA) or data use agreement (DUA)) with partner sites.
    a. Panel members conduct the review and provide comments or approval.
        - **Decline review:** If a panelist is unavailable for a review, the panelist notifies the MENDS Coordinating Center, and an alternate panelist is identified by the MENDS Coordinating Center.
        - **Active approval:** Panelists determine that the data product is de-identified. Determinations are documented and archived.
    b. The MENDS Coordinating Center incorporates and addresses all comments provided during the review process and distributes an updated version.
    c. **Additional Review:** Because the decision must be unanimous, the MENDS Coordinating Center continues the review process until no modifications are suggested and the panel members have reached consensus that the dataset of interest has been adequately de-identified. If the product cannot be deemed de-identified, the data product will not proceed.
4. The MENDS Coordinating Center formalizes the de-identified determination with addition of the following language on the data product: *"The resulting data product has been de-identified using the [Expert Determination method/Safe Harbor method] required by the HIPAA Privacy Rule."*

Partner Site Review

The review process described below is followed for all MENDS data products once they have been determined to be de-identified. The MENDS Coordinating Center manages MENDS data products through the review process. Review processes are time-bound and structured to efficiently gather appropriate review comments and approval with the intention of ensuring rapid dissemination.

Reviewers

For each partner site, Governance Committee members and alternates (i.e., four individuals—two from the data contributors and two from data users) receive review requests and may include other partner site representatives and subject matter experts as appropriate. All partner sites that contributed data to a product are engaged for review. Partner sites are responsible for determining which individuals and organizations must approve each MENDS data product and for coordinating comments and edits. If Governance Committee members are unavailable, they may delegate review responsibilities to other staff by notifying the MENDS Coordinating Center.

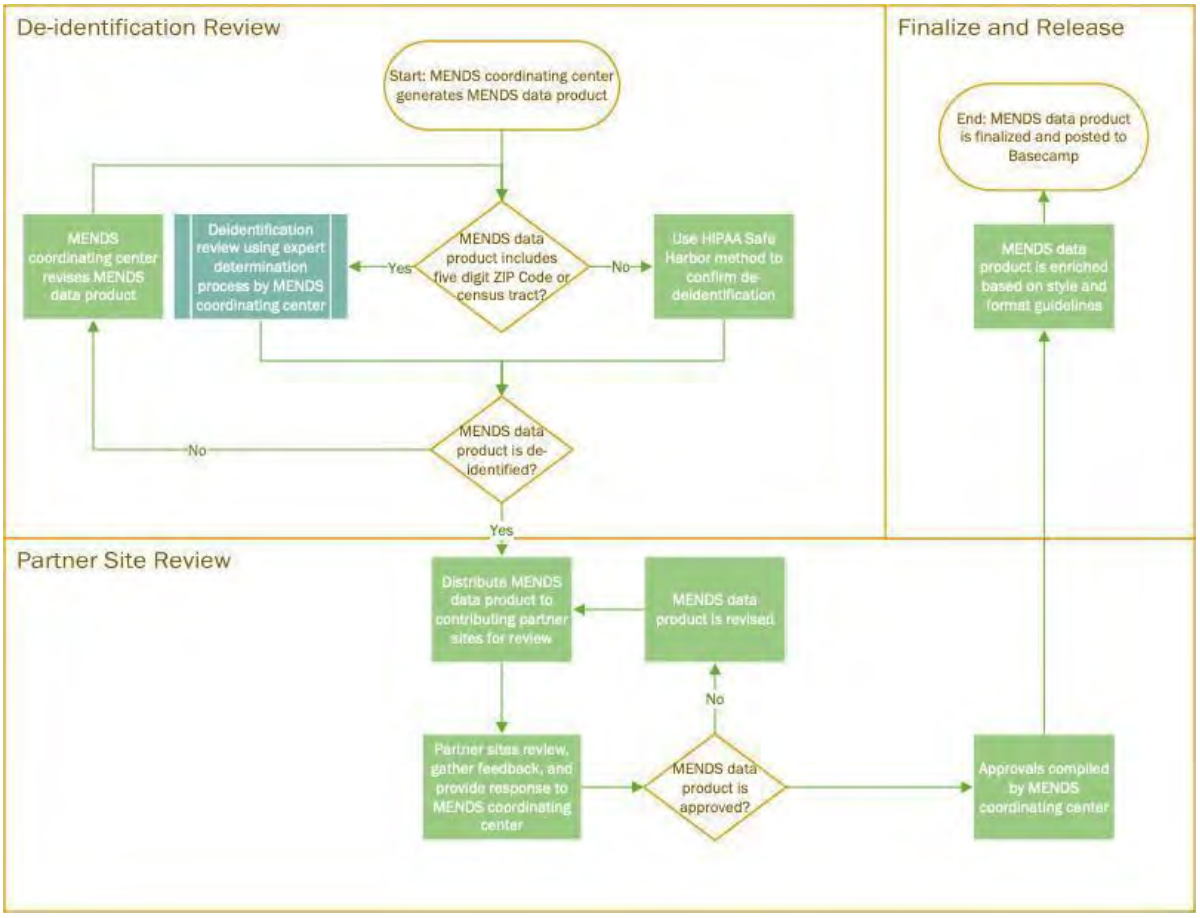
Review Process

1. The MENDS Coordinating Center emails partner site(s) with a review request that includes a description of the data product, the intended use and primary audience(s) (which could include the public), reference information about related data activities (i.e., query information), and the review timeline.
2. Partner sites conduct the review, collect feedback across reviewers, and respond. *One review response per partner site is highly preferred.*
    a. **Extension:** If a partner site cannot complete the review but identifies substantial changes or edits, the partner site contacts the MENDS Coordinating Center to discuss an alternative timeline.
    b. **Additional review:** Partner sites can request an additional round of review based on the significance of the comments provided.
    c. **Reminder:** MENDS Coordinating Center sends a review reminder on the final day of the review period.
    d. **Approval:**
        - *Data product that names one or multiple partner site(s).* For data products that include one or multiple partner sites' data and the data contributor or contributing health organization is specifically named, explicit partner site approval is required from every partner site whose data were used in the creation of the data product.
            a. Partner sites include the following in their response email: "I approve [*DATA PRODUCT NAME*] for dissemination on behalf of [*PARTNER SITE*].
        - *Data product that does not name partner site(s)*
            1. *Data product from one partner site.* For data products that include only one partner site's data, explicit partner site approval is required.
                i. Partner sites include the following in their response email: "I approve [*DATA PRODUCT NAME*] for dissemination on behalf of [*PARTNER SITE*].
            2. *Data product from multiple partner sites.* For data products that include multiple partner sites' data, such as data included in a state report, partner sites have the opportunity to provide comments and ask questions. If no comments are received within the specified review period, approval is implied.
            3. *Network results from all partner sites.* For data products that include all partner sites' data, partner sites are given an opportunity to provide comments and ask questions. If no comments are received within the specified review period, approval is implied.
    e. **Expedited review:** The MENDS Coordinating Center can request an expedited review and solicit comments on an abbreviated timeline.
3. The MENDS Coordinating Center documents, incorporates, and addresses comments provided during the review process, providing follow-up when

warranted.

4. While not always necessary, some data products may need to be presented at a Governance Committee meeting for comment.
5. The MENDS Coordinating Center finalizes the data product and posts a notification to MENDS Basecamp for all MENDS Governance Committee members.

Figure A6.2 MENDS Data Product Review and Dissemination Process



Data Product Style and Format Guidelines

The following guidelines are provided for the creation of data products. The degree to which these guidelines apply depends on the data product.

MENDS Style Recommendations

The following recommendations support the intention to ensure a consistent and recognizable MENDS data product.

- Use the full project title in first use of MENDS acronym: Multi-State EHR-Based Network for Disease Surveillance (MENDS)
- Include MENDS in the data product title, when possible
- Include the MENDS project website link in the document: https://chronicdisease.org/cphl/technical-assistance-hub/data-modernization/mends/
- If the data product was generated through RiskScape, consider the addition of the RiskScape logo or an acknowledgement statement

- If the data product was generated through PopMedNet, consider addition of the PopMedNet logo or an acknowledgement statement
- Follow CDC's Clear Communication Index (https://www.cdc.gov/ccindex/index.html) when possible.

Versioning and Date

Data products must include a production date and version/version history, when applicable.

Acknowledgment

MENDS data products should include an acknowledgment of the project, partner sites, and project funding. Partner sites can decide how they want to be acknowledged on a product-by-product basis within the partner site review process. Unless otherwise indicated, all partner sites must be included in the acknowledgment.

MENDS Acknowledgment (Full) ***includes all data contributors and data users***

We acknowledge the contributions of the MENDS partner sites and project team that participated in the creation of this document: AllianceChicago, Chicago Department of Public Health, Illinois Department of Health, and Cook County Department of Public Health; REACHnet, including the Louisiana Public Health Institute, New Orleans Health Department, Louisiana Office of Public Health, and Texas Department of State Health Services; HCP OneHealthPort, LLC and Washington State Department of Health; Regenstrief Institute and Trustees of Indiana University, Indiana Department of Health, and Marion County Public Health Department; Health Data Compass and the University of Colorado Anschutz Medical Campus; the Centers for Disease Control and Prevention; Commonwealth Informatics; Department of Population Medicine, Harvard Medical School and Harvard Pilgrim Health Care Institute; National Association of Chronic Disease Directors; University of Massachusetts Lowell; National Opinion Research Center; and the Public Health Informatics Institute.

MENDS Acknowledgment (Abbreviated)

We acknowledge the contribution of MENDS partner sites and project team that participated in the creation of this information (https://chronicdisease.org/cphl/technical-assistance-hub/data-modernization/mends/).

Funding Acknowledgment

The "Enhancing the Multi-State Electronic Health Record-Based Network for Disease Surveillance" and "Creating Publicly-Accessible Chronic Disease Prevalence Estimates from the MENDS Network" projects are supported by the Centers for Disease Control and Prevention (CDC) of the U.S. Department of Health and Human Services (HHS) as part of a financial assistance award totaling $1,830,000 with 100 percent funded by CDC/HHS. The contents are those of the authors and do not necessarily represent the official views of, nor an endorsement, by CDC/HHS, or the U.S Government.

Description of Network Infrastructure

When appropriate, the following language should be used to describe the MENDS technical infrastructure:

The MENDS network leverages four software applications: Electronic medical record Support for Public Health (ESP), PopMedNet, RiskScape, and the interactive Visualization of EHR-Based Surveillance Tool (iVEST). ESP (esphealth.org) is an open-source software application that extracts electronic health data, organizes the data into a standard format stored across multiple data tables, and applies algorithms to identify conditions of public health interest. PopMedNet (popmednet.org) is a software application that allows querying of the ESP data

tables. RiskScape (esphealth.org/riskscape) is a software application that provides summaries and visualizations of the ESP data. iVEST is an open-source product that provides modeled estimates of prevalence for MENDS priority indicators. MENDS is the national implementation of these four applications.

For a current list of published journal articles related to MENDS, visit: https://chronicdisease.org/cphl/technical-assistance-hub/data-modernization/mends/presentations-and-publications/.

Authorship

Authorship assigns responsibility and provides appropriate credit for the development of intellectual work. Assigning authorship should reflect the honest contributions made to both the development and finalization of the data product as outlined by the International Committee of Medical Journal Editors (https://www.icmje.org/recommendations/browse/roles-and-responsibilities/defining-the-role-of-authors-and-contributors.html).

The use of "the Multi-State EHR-Based Network for Disease Surveillance Team" may be considered on the authorship line (e.g., when limited authors can be listed but additional authors should be recognized.) Individuals' names will be listed according to journal guidelines, and these people can include the publication in their CVs.

CDC-Specific Considerations

When a CDC employee or contractor is an author or co-author of an abstract, manuscript, or other publication, the publication must adhere to CDC authorship and clearance policies.

Appendix 7: Technical Requirements

Electronic medical record Support for Public Health (ESP) overview

ESP is an open-source, free-license software originally developed under a Centers for Disease Control and Prevention (CDC) Centers of Excellence grant to the Harvard Medical School's Department of Population Medicine. Information is available at esphealth.org. Software source code is available for download at gitlab.com/ESP-Project. ESP performs notifiable and chronic disease case detection against a data mart of patient clinical data. ESP works only with clinical data, including patient demographics, patient visit and visit diagnoses, medications, laboratory test results, immunizations, and social behaviors (smoking and alcohol use). The ESP data mart maintains patient clinical data, including protected health information (PHI), and therefore is typically installed and maintained in the healthcare data partner's data center. The standard model of operation is to extract and load patient medical history from the source data system going back at least to January 1, 2017. Subsequently, all new or updated patient clinical data are loaded on a regular basis—the Multi-State EHR-Based Network for Disease Surveillance (MENDS) target being monthly updates. ESP then uses a set of plug-in disease detection modules to identify disease cases and collect all relevant available condition medical history. ESP data are used to generate aggregate data for provisioning the RiskScape visualization system and for querying via PopMedNet. Once disease detection plugins are configured and validated, ESP requires very little maintenance beyond mapping new laboratory test types when these are added to the electronic health record (EHR) system that feeds ESP.

Sizing Hardware or Virtual System

The following overview is provided to give a sense of scope for determining ESP server requirements. All MENDS sites work with Commonwealth Informatics to fully investigate the site-specific ESP server requirements as part of initial site setup and ESP installation.

Storage

At its core, ESP is a data mart of patient EHR data. The standard system configuration has the ESP Python application installed on the same Linux server as a dedicated PostgreSQL database. Other configurations are supported. Storage requirements can be significant. Storage requirements primarily depend on the number of patients to be included in the ESP system, and the number of years of data the ESP system needs to support (going back historically and going forward for the planned life of the server). Based on Commonwealth Informatics experience working with several large, primarily outpatient healthcare organizations, a good sizing rule is:

> 50 GB base + (1 GB X Number of active patients in your system/5000) X Number of years of patient data kept in the ESP system
>
> For example: An organization with approximately 200,000 active patients intends to maintain up to 12 years of patient data in the ESP data mart. This data mart would require: 50 GB + (1 GB X 200,000/5000) X 12 = 530 GB of storage.

A third factor to consider when sizing storage is the density of patient data, meaning the number of distinct observations, orders, and results collected per patient. If the organization deals with long-term care of very sick patients, patient data density is much higher and requires more storage capacity. If the organization holds a large volume of inpatient data (from hospitalizations), the patient data are highly dense and require additional storage. Alternatively, a healthcare organization that primarily deals with transient short-duration patients with limited care provision have much lower data density and require less storage. Commonwealth Informatics works with organizations to develop more precise estimates in

these cases.

As storage size increases, I/O bandwidth must increase as well. I/O requirements are particularly high during monthly or quarterly data update cycles, so virtual systems will benefit from an elastic I/O provisioning model.

Memory

A recommendation for adequate performance is 4 gigabytes (GB) of memory for each 100 GB of used storage. A system with 1 TB of used storage, for example, would require 40 GB of memory. Used storage refers to the current storage size of the organization's ESP database. This should grow over the lifetime of the ESP installation, and the organization should plan to either initially over-provision memory or plan to add memory to the server as the database grows in size. If there are large disparities in the volume of data collected by a MENDS data contributor across different clinical domains, more memory may be required to provide adequate performance across all clinical data areas.

CPU

ESP can run disease detection algorithms in parallel threads, so more CPUs increases performance. Two CPU cores are a minimal requirement, and two CPUs are needed for each 250 GB of used storage. Optimally, data contributors with 1 TB of storage should provision eight CPUs.

OS and Software Stack

ESP runs on Linux OS. It is developed on Ubuntu Server LTS systems and has run on many other Linux distributions, including RedHat, SUSE, and CentOS.
A basic Linux server would need the following additional software:

- Administrative logins are permitted from anywhere, but the system console requires Open-SSH service.
- Iptables, a Linux computer firewall, should be installed and configured to manage and restrict system access according to policy.
- Git is used for ESP distribution.
- The ESP data mart requires PostgreSQL as the relational database management system (RDBMS), although it can be configured for use with MS SQL Server.
- ESP software is developed using Python 3 and the Django Object Relational Mapper (ORM). An ESP installation uses the Python virtual environment infrastructure.
- An administrative web interface uses Apache web server.
- The system can be configured to run under SELinux if required.

Network Requirements and Data Security

ESP stores and provides access to Health Insurance Portability and Accountability Act (HIPAA) regulated PHI data, so network design for data security issues is of critical importance. The ESP server must exist within a firewalled and access-controlled network environment. Ensuring controls in this environment is the responsibility of the data contributor.

ESP software and updates are distributed via gitlab.com. A firewall rule enabling outgoing requests to gitlab.com/ESP-Project is necessary to obtain ESP software and to obtain updates for maintenance.

Data provisioning for ESP occurs via a set of delimited text files, Fast Healthcare Interoperability

Resources (FHIR®) bulk data client data, or HL7 Consolidated Clinical Document Architecture (CCDA) documents, generated at least quarterly, so there must be a network connection between the ESP server and the machine that provides the extract files. Dataflows to the ESP system should use an encrypted protocol.

Data provisioning from ESP to RiskScape is aggregated at the patient-month level and includes five-digit ZIP Code, census tract, and year/month at the patient level. This is a limited dataset in accordance with HIPAA Safe Harbor rules. Secure Sockets Layer (SSL) and connection authentication is required for transfers. This requires an outgoing rule in the data center firewall to allow traffic from ESP to the RiskScape server.

Similarly, PopMedNet data queries are distributed to a data mart client tool installed alongside the ESP server. This data mart client requires a firewall rule enabling communication between the PopMedNet server and the data mart client.

Appendix 8: PopMedNet Request and Query Process

A request is defined as a solicitation of data from the Multi-State EHR-Based Network for Disease Surveillance (MENDS) network. A network request is a request for data from multiple MENDS partner sites that is implemented through the national PopMedNet server. Network requests are coordinated by the MENDS Coordinating Center. Requestors of MENDS network data include public health entities. The process below is used to request data from the MENDS network through the PopMedNet platform. This process applies network queries (i.e., requests of all partner sites) using either approach to PopMedNet query creation: the query composer and SQL distributions. For information on how to leverage PopMedNet to answer chronic disease surveillance questions using MENDS data, see the PopMedNet for Multi-State EHR-Based Network for Disease Surveillance (MENDS) Data Contributors and Users—Answering Novel Chronic Disease Surveillance Questions tip sheet.

Frequent, thorough, and timely communication about requests is at the forefront of the request process. In keeping with the MENDS guiding principle of transparency, request communication includes partner site Governance Committee representatives, partner site PopMedNet data mart administrators, and the MENDS project team.

Network Request Process

1. The requestor populates and submits the MENDS network request form.
2. The MENDS Coordinating Center receives the request and reviews the request information for completeness.
    a. The MENDS Coordinating Center may contact the requestor for additional relevant request information.
3. The MENDS Coordinating Center drafts and distributes the request receipt notification to partner sites via email.
    a. The notification includes the recommended query response period.
    b. Partner sites can submit feedback or questions related to the request.
    c. Partner sites identify a subject matter expert for the request topic as appropriate.
4. The MENDS Coordinating Center works with the MENDS information technology (IT) vendor to develop and test the PopMedNet query based on parameters provided in the request.
5. The MENDS Coordinating Center distributes the PopMedNet query to data contributors and sends the query distribution notification to partner sites via email.
    a. The query goes to the PopMedNet data mart client at each partner site.
    b. Partner sites should have at least two PopMedNet data mart client administrators (i.e., primary and backup).
6. PopMedNet data mart administrators are notified of query arrival via automated email from the PopMedNet data mart client.
7. PopMedNet data mart administrators review the query and approve the query to run.
    a. Most queries are completed in less than 2 hours,
    b. If query-related errors occur, partner sites contact the MENDS Coordinating Center or share error information with the MENDS Coordinating Center through PopMedNet.
8. Upon query completion, the data contributors review the query results.

9. Partner sites contact the MENDS Coordinating Center with questions. Partner sites submit query results to the MENDS Coordinating Center through PopMedNet.
10. The MENDS Coordinating Center compiles query results across partner sites into MENDS network data.
    a. The MENDS Coordinating Center conducts data quality checks and contacts partner sites regarding irregularities.
    b. The MENDS Coordinating Center strips partner site identifiers.
11. The MENDS Coordinating Center distributes the request completion notification to the partner sites via email confirming request completion.

MENDS
MULTI-STATE EHR-BASED NETWORK FOR DISEASE SURVEILLANCE

MENDS PopMedNet Network Request Information

Q1 Please provide your contact details below:

- O Name: _____
- O Organization: _____
- O Email: _____

Q2 What are your surveillance questions?

Q3 Which risk factor(s) is this request related to?

- ☐ Hypertension
- ☐ Cholesterol
- ☐ Smoking
- ☐ Diabetes
- ☐ Obesity
- ☐ Asthma
- ☐ Other, please specify_____

Q4 How will you use the resulting information?_____

Q5 Please provide specifics for this query to support SQL development. This can be done by adding text below, providing public links to details, or uploading files.

- O Who is the target population? _____
- O Observation period: _____
- O Exclusions: _____
- O Other related information: _____

Q6 Supporting files for query requests can be uploaded below. *[Attachment function]*

Q7 What type of MENDS data product is needed? (e.g., summary table, slides, etc.)

Q8 Is there a deadline associated with this request?

Local Request Process and Information

The MENDS Coordinating Center provides limited query support to generate chronic disease surveillance information within MENDS partner sites, using a local request form. The MENDS-local-queries GitHub site is a repository of validated queries and information for data users.