HMORN CERT DEcIDE Distributed Query Tool: System Description and Technical Documentation

December 2010

Based on release 2.1.11

CONTACTS:
Jeffrey Brown, PhD
Harvard Pilgrim Health Care Institute
Jeff_brown@hphc.org

Table of Contents

1. Overview and Background

The Distributed Query Tool is a software application developed as part of several contracts awarded by the Agency for Healthcare Research and Quality (AHRQ) to the HMO Research Network (HMORN) Center for Education and Research on Therapeutics (CERT) DEcIDE Center housed in the Department of Population Medicine (DPM) at the Harvard Pilgrim Health Care Institute (HPHCI). The software application has been enhanced using additional funding via the FDA Mini-Sentinel contract with HPHCI as the prime contractor. The system was developed by Lincoln Peak Partners (LPP) under the direction of HPHCI. LPP hosts and maintains the system.

The system design is based on findings from previous studies, [1-7] coupled with our experience in operating a distributed network, the HMORN CERT, and participating in other networks.[8-12] Our prior work assessed the needs of data partners (e.g., health plans) and potential network users (e.g., data partners, federal agencies) with respect to making data available for comparative effectiveness and other secondary uses. [3, 4, 6] Data partners identified several requirements for voluntary participation in a distributed network. These included: 1) complete physical and operational control of their data, 2) strong security and privacy features 3) limited impact on internal systems, 4) minimal data transfer, 5) auditable processes, 6) standardization of administrative and regulatory agreements, 7) transparent governance, and 8) ease of participation and use.

The system allows users to create and securely distribute “queries” to network data partners and to have data partners review, execute, and securely return the results of those queries to the requestor via a secure web-based Portal. Data partners maintain control of their data, and data partners have the ability to review all queries before they are executed locally, and to review all query results before the results are transferred securely back to the Portal. The system design allows data partners to automate any portion of the query process.

To date, two distinct networks have been established using the software application - the HMORN CERT DEcIDE Network and the FDA Mini-Sentinel Network. This document describes the overall system architecture, and details the technical and security approaches implemented by both networks within the system. Although the two networks share the same underlying architecture, the networks are hosted separately and operate under different governance structures. This document includes the governance rules for the **HMO Research Network CERT DEcIDE Network (section 4);** documentation of the governance of other networks is maintained separately.

This documentation is based on Version 2.1.11 of the Distributed Query Tool software.

2. System Overview

The Distributed Query Tool system is comprised of two separate applications, the Portal and the DataMart Client. The **Portal** (there is one Portal per Network) is the starting point for all information requests and controls all system communications, security, and governance policies. Data partners receive queries, process them, and securely return them to the Portal via their local **DataMart*** **Client**. There is exactly one Portal in the network and many DataMarts. All query requests and communications within the network are securely routed from the Portal to the DataMarts and then back to the Portal. The reference material provides additional details on the querying process.

To participate in a network, data partners must:

1. Install and configure the desktop application (i.e., the DataMart Client) on one or more local computers
2. Assign one or more staff members to act as the DataMart Administrator responsible for interacting with the system on behalf of the data partner
3. Set DataMart preferences on the Portal to establish settings, such as what data can be queried and who can submit queries to the DataMart
4. Create data in the required format and make it available for querying.

To date, menu-driven querying requires partners to create summary tables (described below) and make them available via the network. The DataMart Administrator or other staff members do not need any special information technology or computer expertise to install the software, manage the DataMart, or respond to distributed queries.

DataMart Client

The DataMart Client application allows the DataMart Administrator to view queries distributed to the DataMart, execute queries locally, review the results, and upload the results to the portal. The DataMart Client is a .NET/C# Windows desktop application developed by LPP that is installed locally on an Administrator's desktop. All communications between the DataMart Client application and the Portal use HTTP/SSL connections to securely transfer queries and results between the application and the Portal. The application uses ODBC connections to the local DataMart databases used to process queries and generate results.

Query Types

The system currently supports two types of queries: 1) menu-driven queries that execute against summary tables, and 2) file distribution queries.

Menu-driven queries are created by users on the portal and distributed to data partners. These queries must be run against pre-populated summary tables that are created and maintained by the data partners. The summary tables are described in detail in Appendix A. Briefly, these tables provide summary counts of individuals by period, age group, and sex. The summary counts include information on medication use (e.g., number of dispensings, users, and days

* The term "DataMart" is used in an information technology context referring to the place where the data are held for querying. Use of this term does not imply that data partner information is being sold or being made broadly available; data partners maintain control of all their data and all uses.

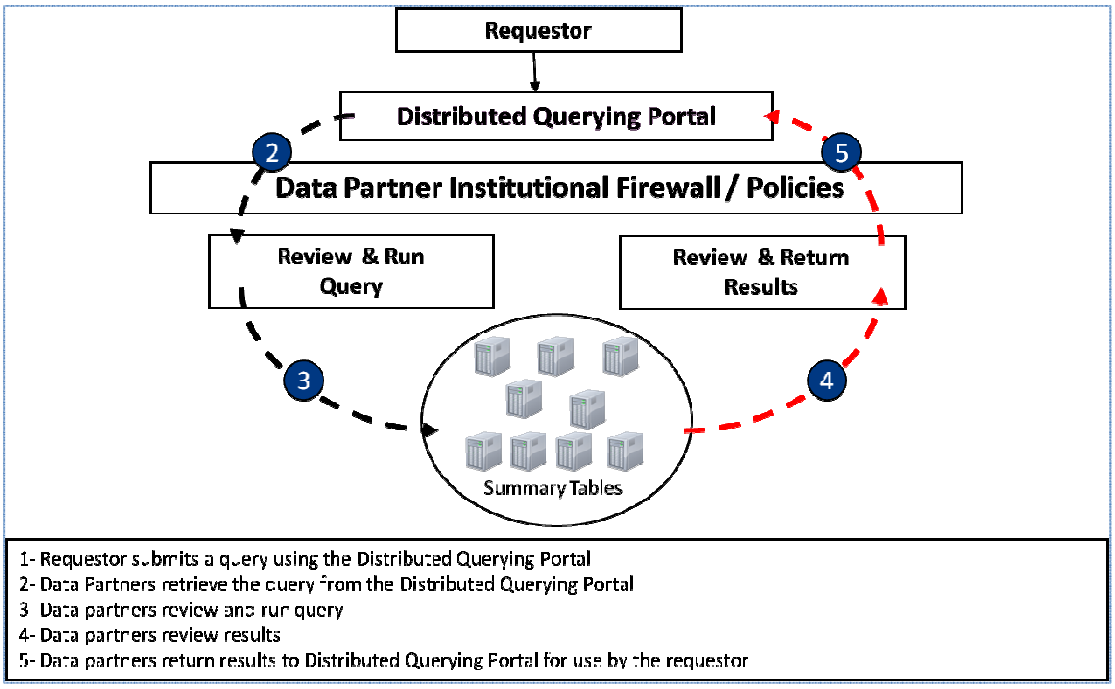
supplied), diagnoses (e.g., number of individuals with the diagnosis), procedures, and the overall data partner population.

A **File Distribution Query** allows users to securely distribute electronic files to data partners. Although any type of file can be distributed, a common use is expected to be the distribution of SAS programs and work plans to data partners who will download and execute the programs and then securely upload results upon institutional approval to do so.

Network Workflow

Figure 1 illustrates the flow of requests and information within the network. The workflow can be divided into activities undertaken by the requestor and those that are the responsibility of the data partner. Each is described below.

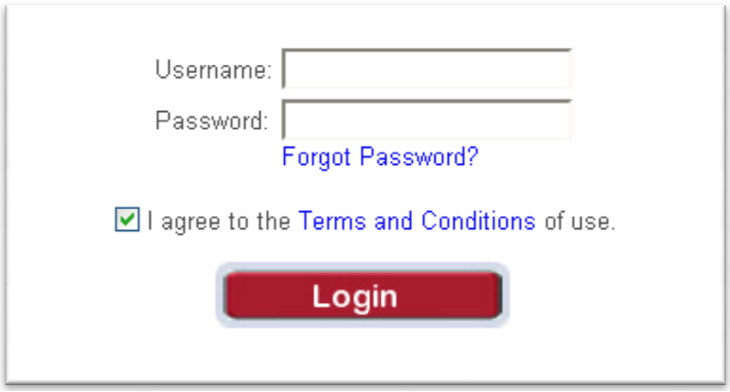
Figure 1. Network Workflow



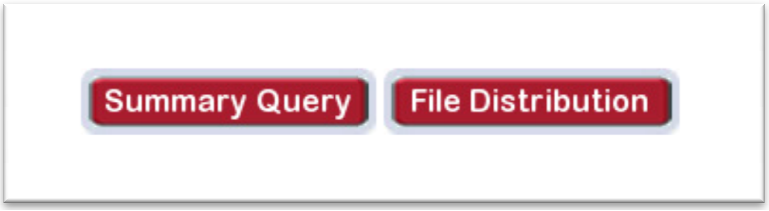
2.1. Submitting a Query (requestor actions)

Submitting a query through the network requires several steps. The steps, along with a screenshot from the software for illustration are listed below.

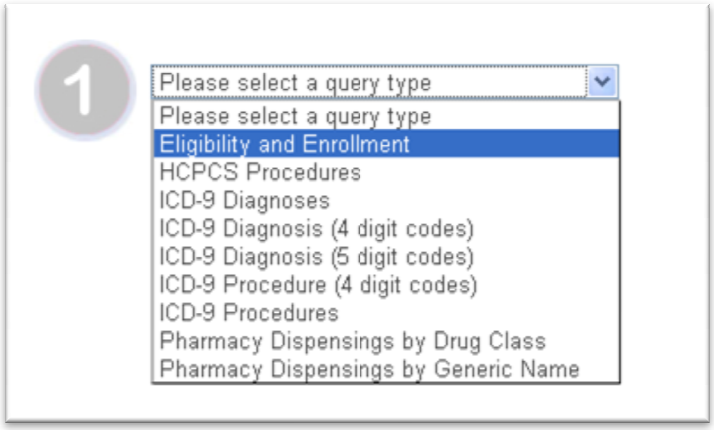
1. User login



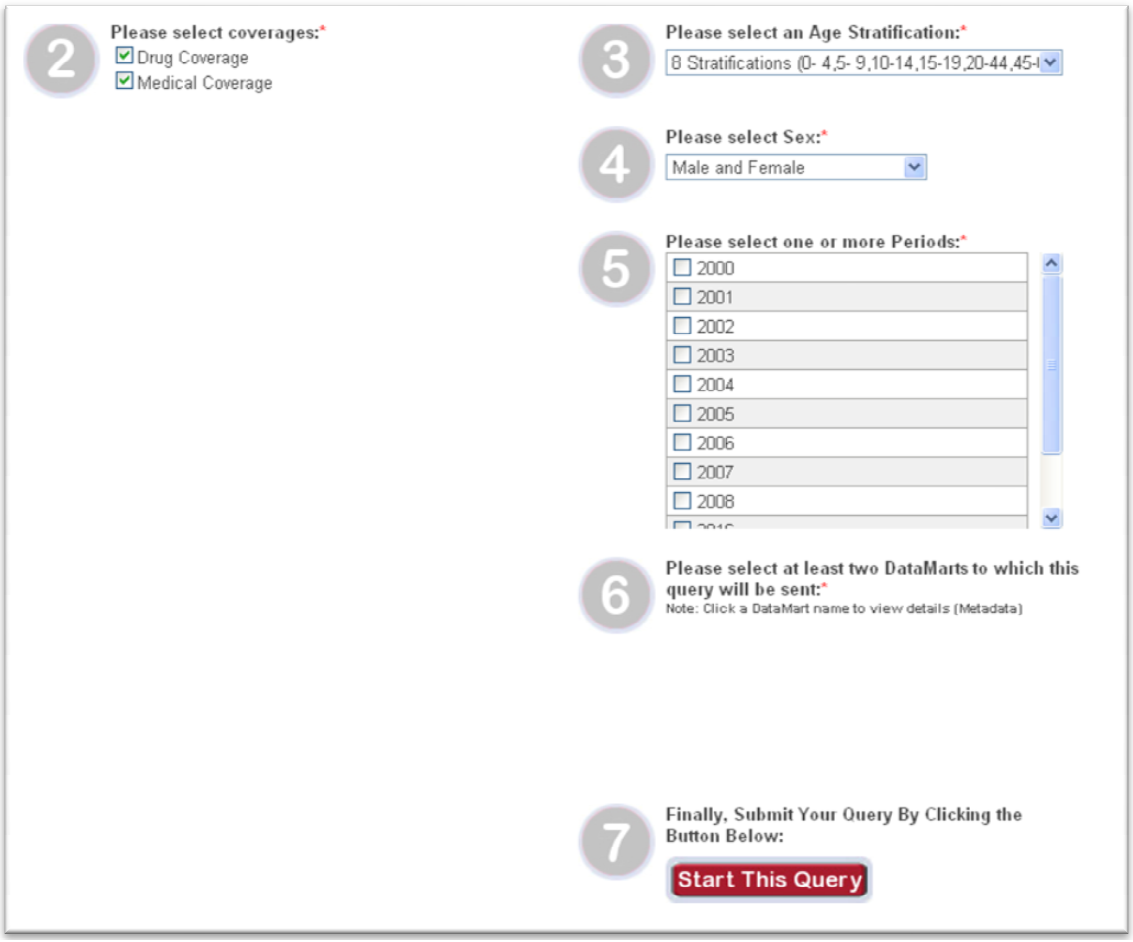
2. Select the query type



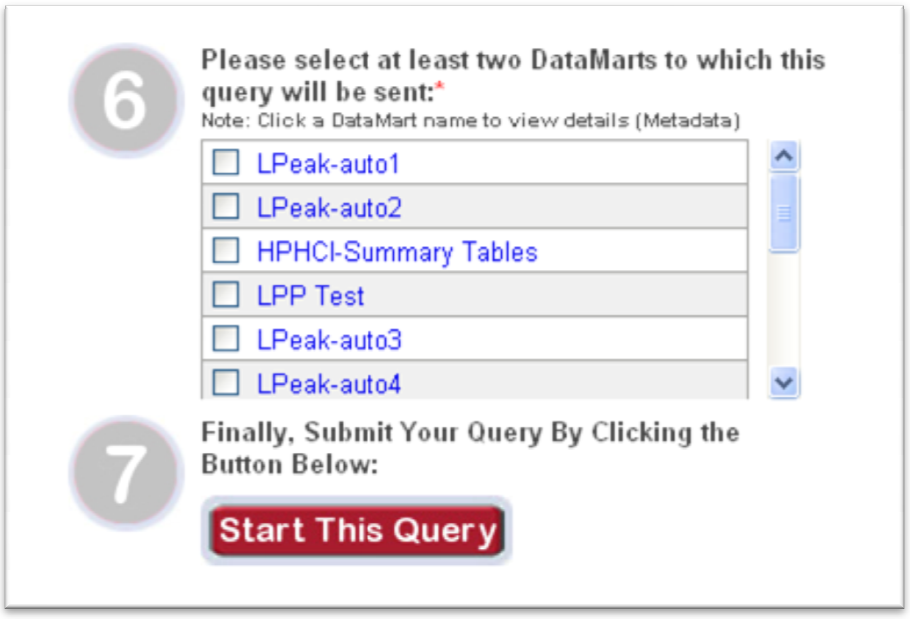
3. For summary queries, select specific summary table



4. Build the query (example of Eligibility and Enrollment query)

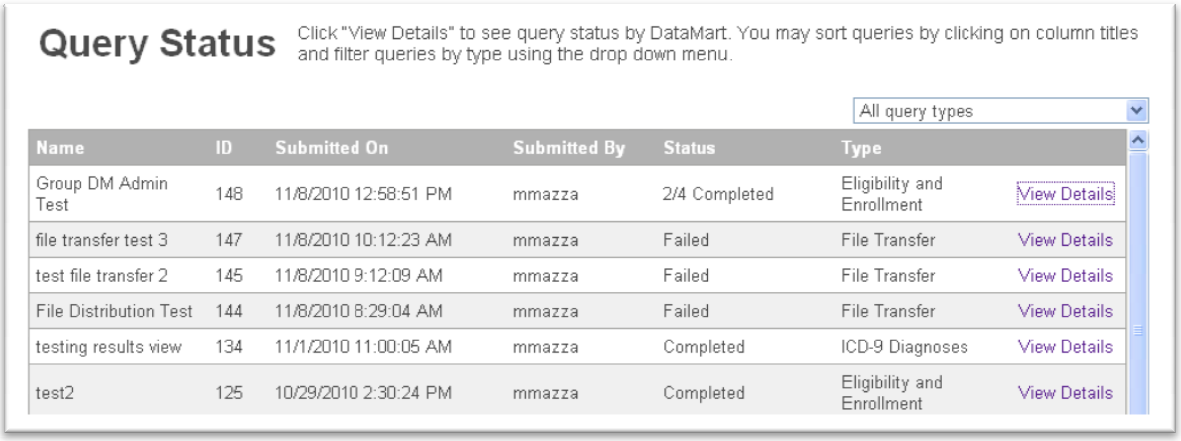


5. Complete your query by selecting the DataMarts you are submitting your query to and then select Start This Query.



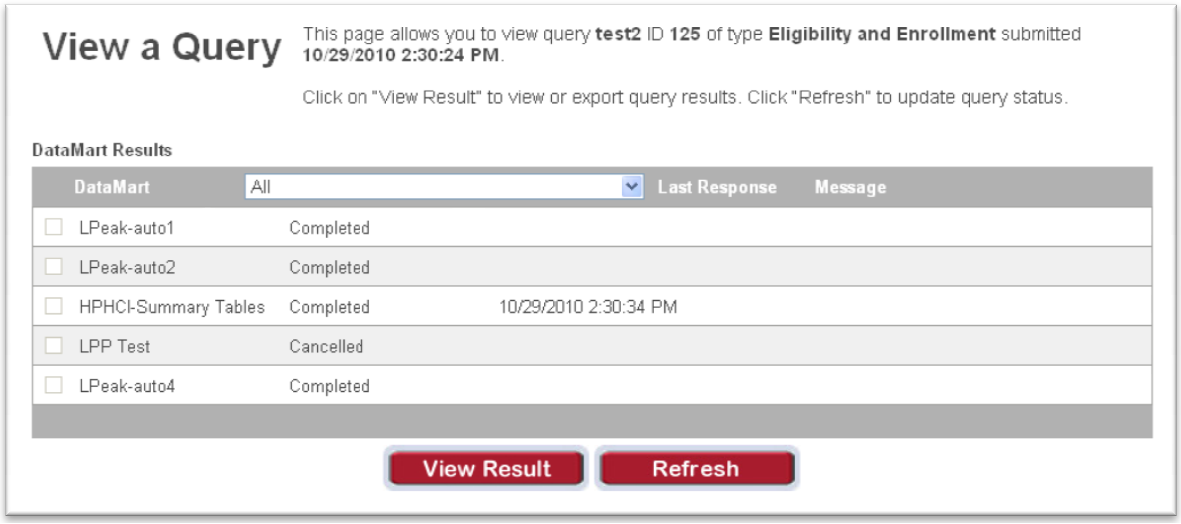
At this point the query is distributed to the selected DataMarts (data partners) for local execution and upload of results. An email notification system notifies the requestor when data partners upload results.

6. View Query Status

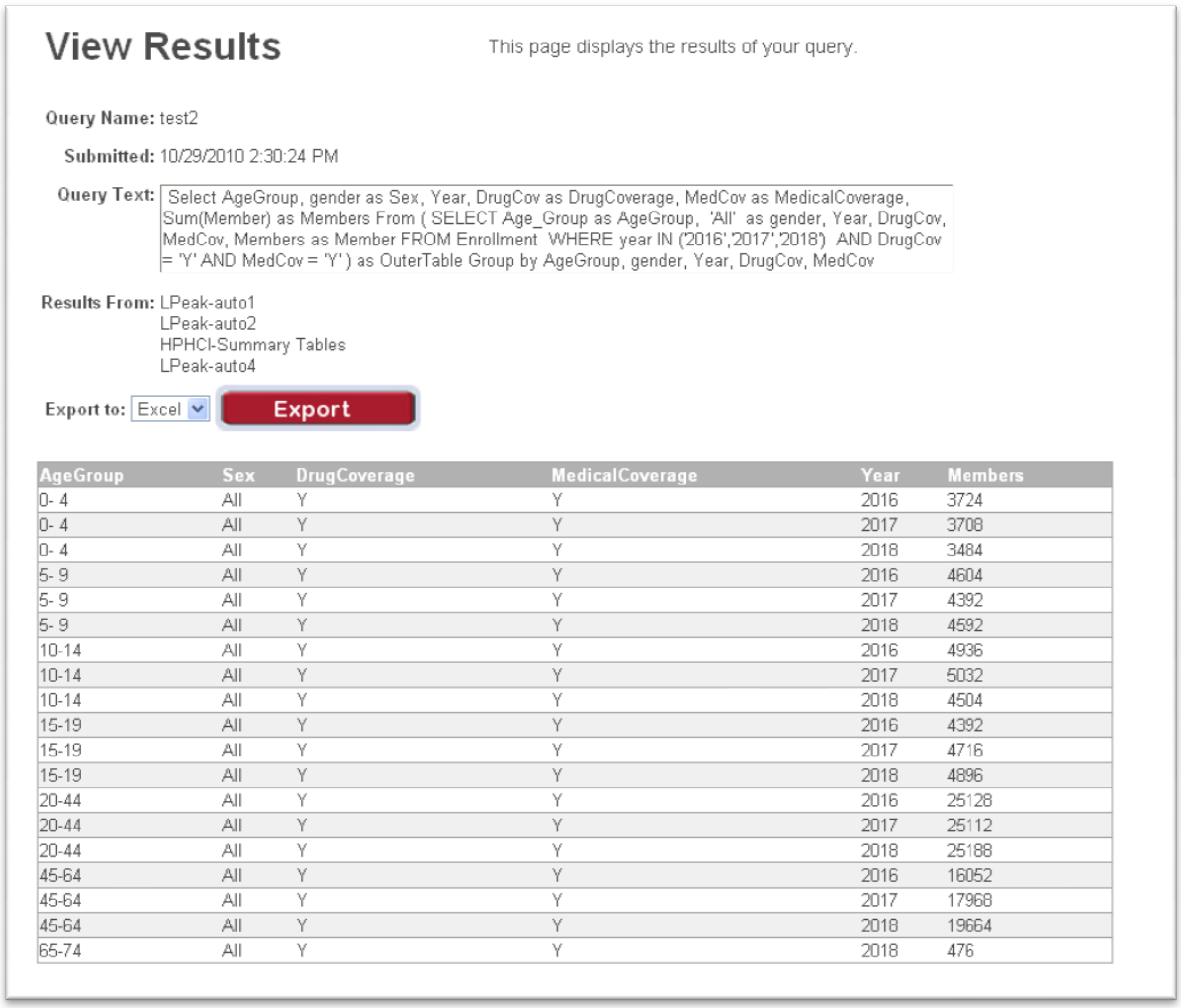


Name	ID	Submitted On	Submitted By	Status	Type	
Group DM Admin Test	148	11/8/2010 12:58:51 PM	mmazza	2/4 Completed	Eligibility and Enrollment	View Details
file transfer test 3	147	11/8/2010 10:12:23 AM	mmazza	Failed	File Transfer	View Details
test file transfer 2	145	11/8/2010 9:12:09 AM	mmazza	Failed	File Transfer	View Details
File Distribution Test	144	11/8/2010 8:29:04 AM	mmazza	Failed	File Transfer	View Details
testing results view	134	11/1/2010 11:00:05 AM	mmazza	Completed	ICD-9 Diagnoses	View Details
test2	125	10/29/2010 2:30:24 PM	mmazza	Completed	Eligibility and Enrollment	View Details

After viewing the Query Status page, you can select View Details for each query you submitted.

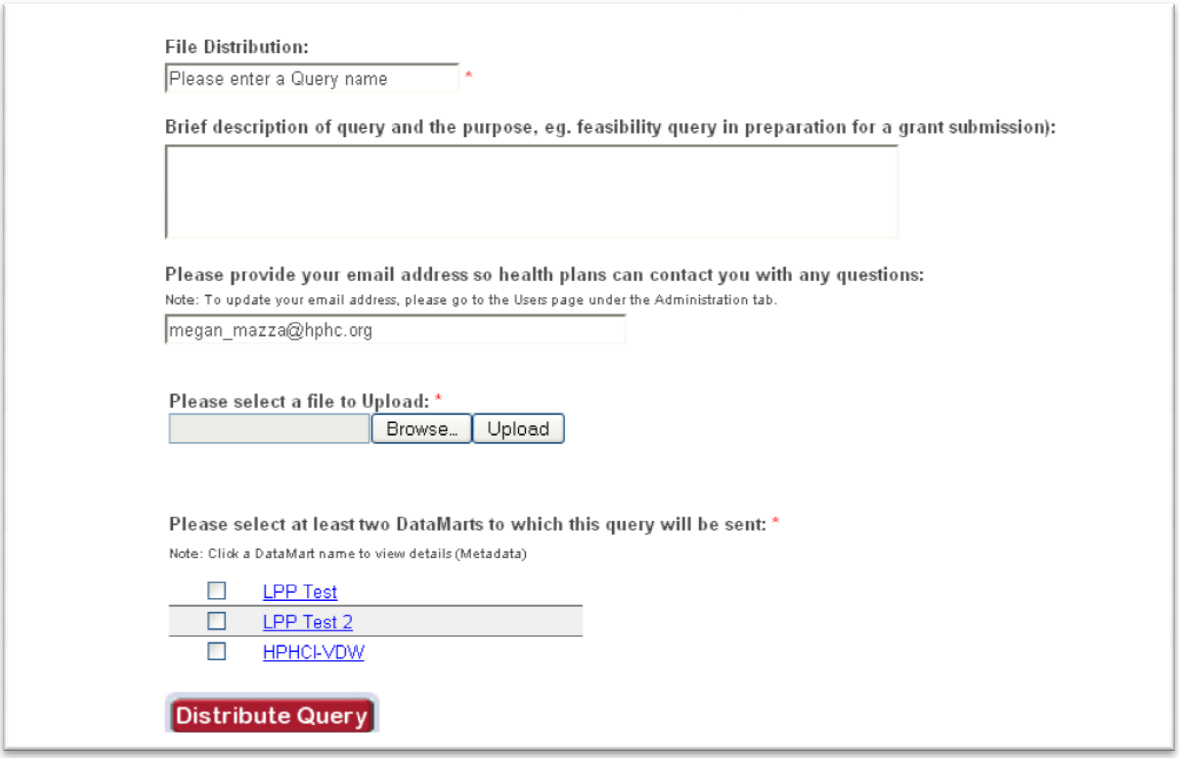


7. View or Export Query Results



AgeGroup	Sex	DrugCoverage	MedicalCoverage	Year	Members
0- 4	All	Y	Y	2016	3724
0- 4	All	Y	Y	2017	3708
0- 4	All	Y	Y	2018	3484
5- 9	All	Y	Y	2016	4604
5- 9	All	Y	Y	2017	4392
5- 9	All	Y	Y	2018	4592
10-14	All	Y	Y	2016	4936
10-14	All	Y	Y	2017	5032
10-14	All	Y	Y	2018	4504
15-19	All	Y	Y	2016	4392
15-19	All	Y	Y	2017	4716
15-19	All	Y	Y	2018	4896
20-44	All	Y	Y	2016	25128
20-44	All	Y	Y	2017	25112
20-44	All	Y	Y	2018	25188
45-64	All	Y	Y	2016	16052
45-64	All	Y	Y	2017	17968
45-64	All	Y	Y	2018	19664
65-74	All	Y	Y	2018	476

The process described above is the same process used for a **file distribution query**, with the only difference being that instead of creating a query (step 4) the user selects files to distribute to data partners. Screenshots of the file distribution query interface are presented below.

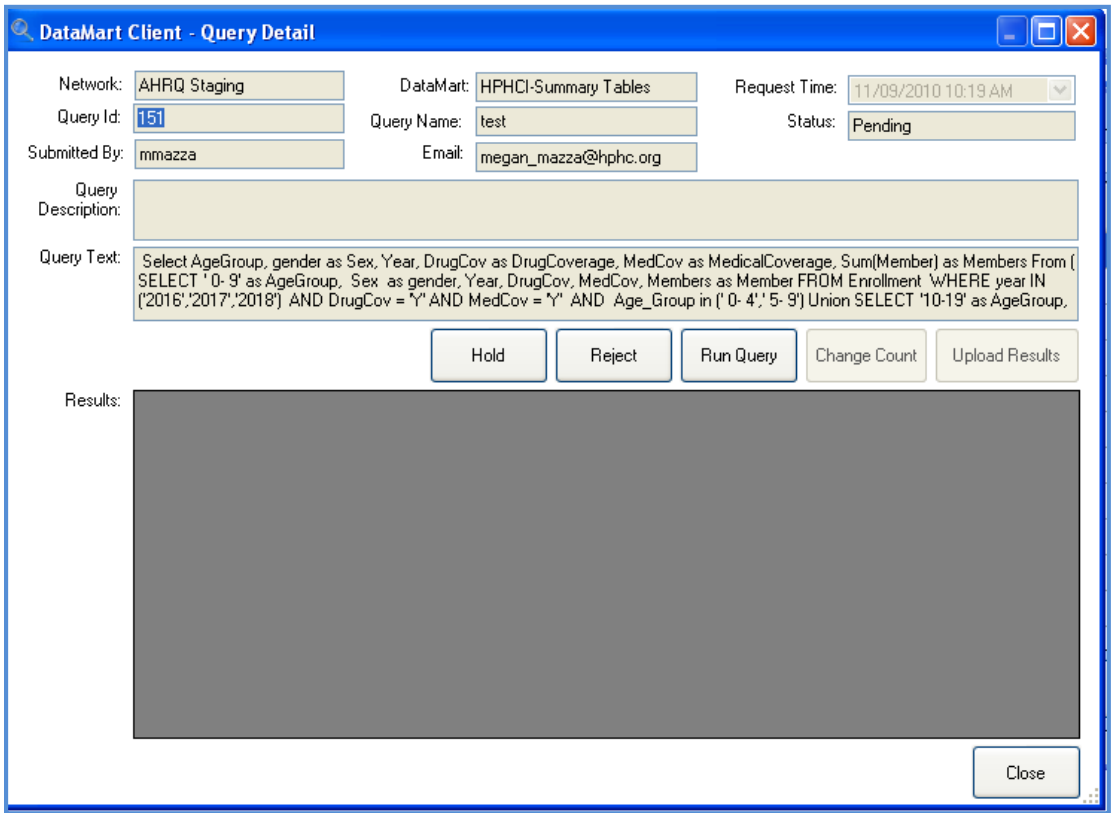


2.2. Responding to a Query (data partner actions)

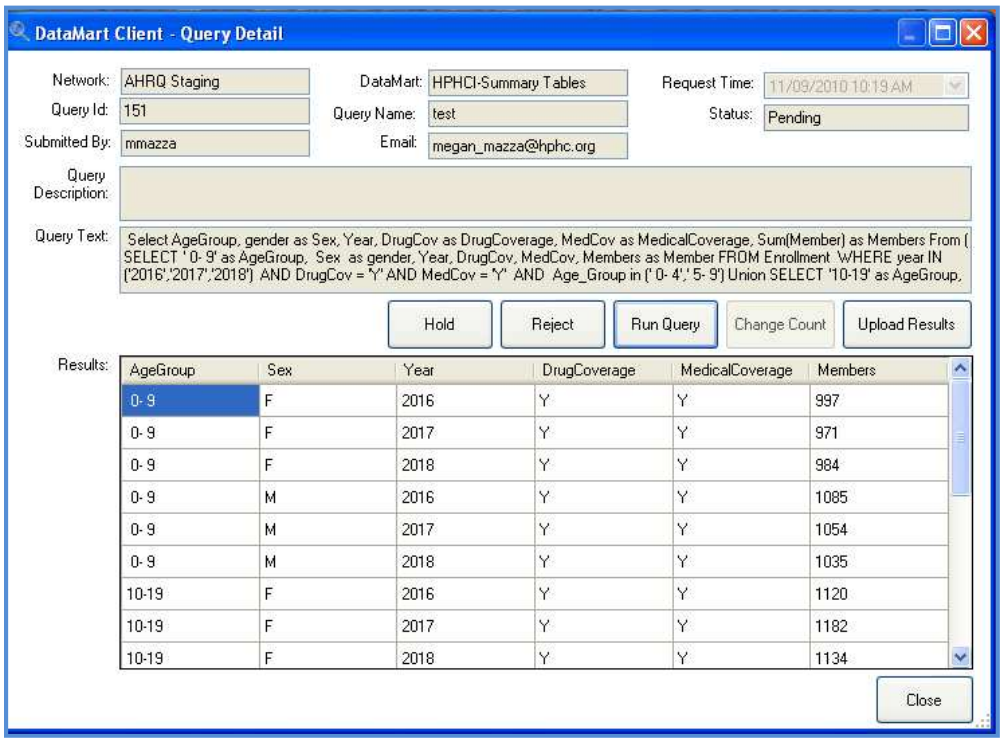
Responding to a query through the network requires several steps by the DataMart Administrator using the locally installed DataMart Client.

The steps necessary for responding to a query are described below. Data partners have the ability to set a notification for small cell counts (a parameter setting), and to re-set those counts to "0" before uploading to the portal. The status of a query will be updated in the Portal according to the actions of the DataMart Administrator.

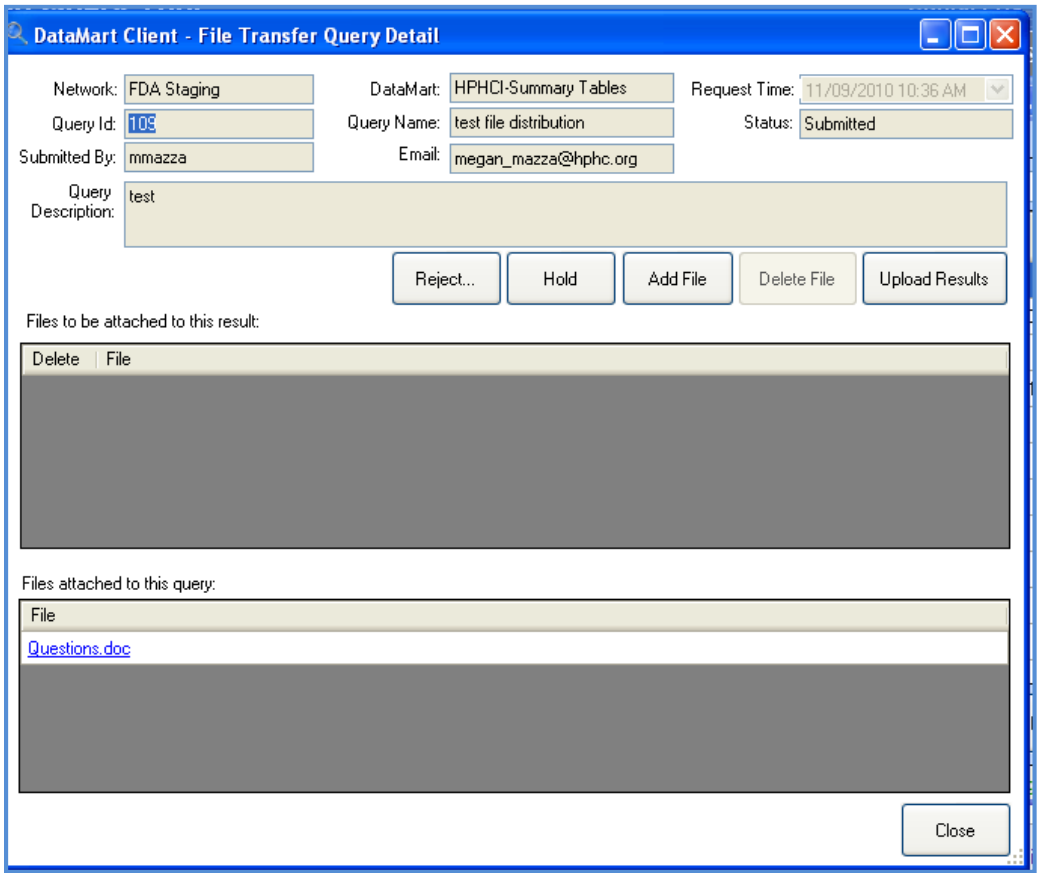
1. Review the query details; then Run Query (accept), Hold or Reject the query.



2. Review the query results; obfuscate low cell counts if necessary. After you review the query results you may then select Upload Results which will send your results to the Portal.



Responding to a **file distribution query** follows the same query review process. Instead of downloading a query for execution, the DataMart downloads the file(s) and can later upload files in response to the request. Screenshot for responding to file distribution queries are below.



3. System Security Policies and Features

3.1. System Users and Roles

The system currently has five main roles. Additional roles may be defined and developed at the discretion of a System Administrator and in accordance with the governance of the system. The user roles are: System Administrator, Group Administrator, DataMart Administrator, Investigator and Enhanced Investigator.

1. System Administrator: can add new data partners; create groups, organizations, and roles; add/delete users; re-set passwords; and view all queries submitted. The System Administrator has complete access to the network and all its functions. This role is limited to staff at LPP and HPHCI.

2. Group Administrator: able to review, aggregate, and release results for the group. A group of data partners can designate a person as the group administrator, and select

rules that require the group administrator to review group results before the results are released to the requestor. Results can be released individually or as an aggregate.

3. DataMart Administrator: manages the local DataMart(s) for each data partner. This role can set DataMart preferences on the Portal and DataMart Client (e.g., what data can be queried and by whom). There can be one or more DataMart Administrators per data partner. DataMart Administrators cannot send queries to other DataMarts.

4. Investigator: can submit queries to DataMarts that have given them or their organization permission to submit queries and view only aggregated query results.

5. Enhanced Investigator: can submit queries to DataMarts that have given them or their organization permission to submit queries and review their query results. This role has the additional right to view site results individually across the organizations within the query.

New data partners and Investigators can only be added to the network by the System Administrator and in accordance with system governance. The system uses role-based access control to give network users permission to perform certain functions (also referred to as giving a user a "right" as described below). Network users who have two roles, for example Investigator and DataMart Administrator, must log-in the system using the proper role; their rights are not combined into a new role.

3.2. Role-Based Access Control

Role-based access control is a common approach for managing a complex system with multiple users, each of whom many have different needs for access and control within the system. For example, an electronic filing system may give all users the ability to store, download, and save files in system folders, but only selected users with the proper permissions can add or delete folders.

Within the distributed network, role-based access control is used to assign users the ability to perform functions in the network. Technically, role-based access control within the network determines if and how an **Entity** (e.g., an Investigator) acquires **Permission** to a **Target** object or capability (e.g., permission to submit queries). Entities are recipients of a grant of access, and there are three entity types organized into a hierarchy. **Users**, which are individuals, are at the bottom of the hierarchy. Users belong to **Organizations**; one User belongs to one Organization, and one Organization encompasses many users. An Organization is defined as a company or institution (e.g., Kaiser Permanente Colorado, FDA, AHRQ). Organizations may belong to Groups. One Organization may belong to many (or no) Groups, and one Group can encompass many users. A Group is defined simply as a set of related organizations (e.g., Kaiser Permanente, HMO Research Network).

Permissions are an allowance granted or restricted by access control. Permissions are associated with a target and granted to an entity. Permissions are inherited by entities through parent / child relationships. An Organization inherits all permissions granted to the Groups to which it

belongs, and a User inherits all permissions granted to its parent Organization. Thus, a User has the union of all permissions granted to the User, the parent Organization, and all parent Groups.

A Target is either a data object or a capability. There are three types of Targets:

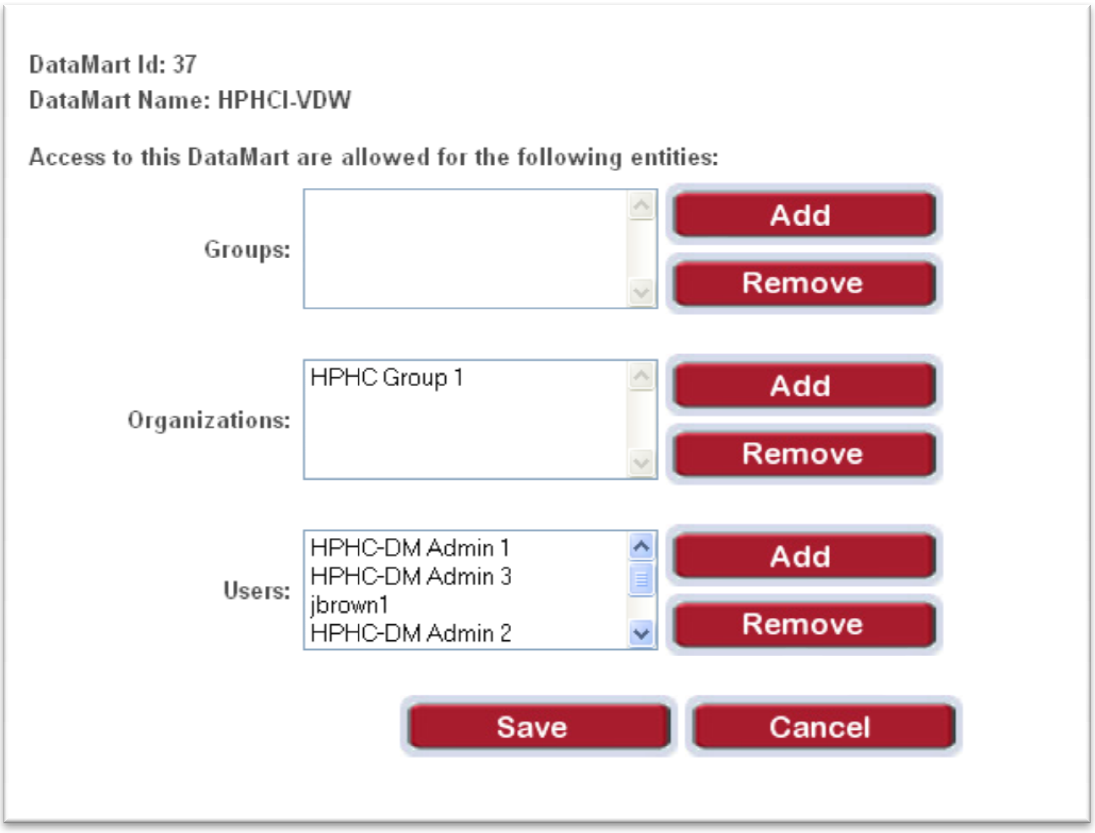
1. DataMart: Instance of a remote database, typically hosted by a data partner. Granting permission to a DataMart allows the target entity (user, organization, or group) to submit queries to that DataMart. Permissions to administer a DataMart and respond to queries submitted to a DataMart are covered under Rights.

2. Right: Capability or function available to users of the system such as "log in" or "submit query". Granting a right allows the target entity (user, organization or group) to utilize a particular function (e.g., "administer user profile") or utilize a particular function in a particular way (e.g., "administer user profile for another user in the same organization").

3. Query Type: Granting permission to a Query Type allows the target entity (user, organization or group) to submit queries of that type.

The following screen shot shows how a DataMart administrator can limit DataMart access to particular entities (i.e., prevent others from being able to submit a query to that DataMart):

DataMart Administrator Page



All communications between the DataMart clients and the Portal occur via a secure Web Service. Transactions are secure and one way; the DataMart client always calls the Portal Web Service. There is no way "rogue" Portal requests can be sent to DataMarts, as there is no mechanism to do so. DataMart Client settings (e.g., requiring manual review of all queries before execution) are stored in a settings file at the DataMart Client and not in the Portal database, and there is no mechanism for the Portal to query or set this DataMart information. Therefore, only the DataMart Administrator can set or change DataMart Client settings. DataMart administrators can choose to receive email notifications whenever a DataMart setting is changed.

Table 1 and Table 2 below specify the **rights** granted within the system for each role. These rights can be assigned to system users; a pre-defined set of rights is referred to as a role. For example, the role of Investigator has a standard set of rights that permit the Investigator to submit queries and review results of those queries. Only the System Administrator can assign additional rights to a system user. Many of the system rights are designed to facilitate administration of the system and are reserved for the System Administrator.

Portal and DataMart Client Rights

Table 1. Distributed Query Tool Portal Rights

	Investigator	Enhanced Investigator	DataMart Administrator	Group DataMart Administrator	System Administrator
Log into Portal	Yes	Yes	Yes	Yes	Yes
QUERIES					
Summary Query	Yes	Yes	Yes - Limited	No	View Only
File Distribution	Yes	Yes	No	No	View Only
Query Own DataMart	Yes	Yes	Yes	No	No
RESULTS					
View own query status	Yes	Yes	Yes	N/A	N/A
View organization's query status	No	No	Yes - Limited	No	Yes
View group query status	No	No	No	Yes	Yes
View any query status	No	No	No	No	Yes
Add/Remove DMs to own query	Yes	Yes	No	No	No
View own result summary	Yes	Yes	Yes	N/A	N/A
View own result detail	Yes	Yes	Yes	N/A	N/A
View organization's result summary	No	No	Yes	Yes	Yes
View organization's result detail	No	No	No	No	No
View individual site results for any organization that your query has been submitted to	No	Yes	No	Yes	No
Aggregate individual site results for any organization that your query has been submitted to	No	No	No	Yes	No

Approve individual site results for any organization that your query has been submitted to	No	No	No	Yes	No
ADMINISTRATION					
Add/Delete User	No	No	No	No	Yes
Add/Delete Organization	No	No	No	No	Yes
Add/Delete Group	No	No	No	No	Yes
Administer User Profile	Yes	Yes	Yes	Yes	Yes
Administer Organization Profile	No	No	Yes	No	Yes
Administer Group Profile	No	No	No	No	Yes
Administer Query Permissions	No	No	No	No	Yes
Administer Rights	No	No	No	No	Yes
Administer Roles	No	No	No	No	Yes

Table 2. Distributed Query Tool DataMart Client Rights

	Investigator	**Enhanced Investigator**	**DataMart Administrator**	**Group DataMart Administrator**	**System Administrator**
Log in	N/A	N/A	Yes	N/A	Yes
View list of queries submitted to DataMart	N/A	N/A	Yes	N/A	Yes
View query detail	N/A	N/A	Yes	N/A	Yes
Hold, Run, Reject or Cancel queries	N/A	N/A	Yes	N/A	N/A
Mark low cell counts in query results	N/A	N/A	Yes	Yes	Yes
Upload Results	N/A	N/A	Yes	Yes	Yes

4. Governance Policies for the Distributed Query Tool

The HMORN CERT DEcIDE Network has a Governance Panel that evaluates and develops their own policies for use of the Distributed Query Tool. The FDA Mini-Sentinel Network also has a governance panel which develops their own policies that are described in separate documentation. A listing of governance policies for the HMORN CERT DEcIDE Network is below. Detailed governance policies for the HMORN CERT DEcIDE Network are available separately.

- Representatives from HPHCI and LPP will serve as system administrators.
- New data partners and network users can only be added to the network by the System Administrator and in accordance with network governance policies.
- Role-based access control gives network users permission to perform certain function; network users who have two roles (e.g., Investigator and DataMart Administrator) must login the system using the proper role; their rights are not combined into a new role.

- Approved partners may view site-specific results, all other will only be able to view aggregated results; network rules will ensure results cannot be disaggregated.
- Data partners will appoint one or more individuals to serve as DataMart administrators for their sites. DataMart administrators will be responsible for responding to queries distributed to their DataMart through the network.
- DataMart administrators will retain full control over access to their data and of the transmission of query results. They will have the ability to accept or reject each query on a case-by-case basis.
- Data partners may use the network to query their own data.
- DataMart administrators can at any time create audit reports of activity related to their DataMart.
- DataMart administrators will determine their DataMart access settings on the Portal, including contact information, the tables available for querying and the users/ organizations/groups able to send queries. These settings can be changed at any time.
- The system administrator will not alter any DataMart settings without prior approval of the DataMart administrator; DataMart administrations can opt to be alerted via email when any DataMarts settings change.
- Users are restricted to a maximum of 10 items in a single query (e.g., users can select up to 10 drugs).
- Query results **may not** be used in a proposal or in any report without the consent of the Network member organization where the data originated.
- No publication or external report other than use in research proposals is permitted.

As new policies are developed by the Governance Panel, the Coordinating Center at the Harvard Pilgrim Health Care Institute will notify users and implement changes.

5. Technical and Security Overview

5.1. Hosting, Security and Support Requirements

This section provides a detailed description of the hosting, security, and support features of the Distributed Query Tool and Portal software system that is currently supporting the HMORN CERT DEcIDE and FDA Mini-Sentinel networks. Each network is hosted separately in the same secure environment; there are two separate portals and two separate implementations of the system. The next two sections describe the system hosting infrastructure and security controls.

Hosting, Security and Support for Distributed Query Tool is provided by LPP and consists of:

- ✓ Hosting that is compliant with Federal Information Security Management Act (FISMA) requirements.
- ✓ Hosting through the full software development lifecycle (including design, implementation, unit testing, user acceptance testing and preparation for production).
- ✓ Deploying the system into production environment.

- ✓ Supporting all production versions of the applications.
  - This involves monitoring and maintaining the application and its operating environment as well as effectively responding to technical questions and issues encountered by the users.

The general requirements and detailed requirements are in Table 3 and 4:

Table 3. Hosting, Security & Support: General Requirements

Requirement	Description
General Requirements	
Multiple Hosting Environments	Separate Development / QA / UAT (User Acceptance Testing) and Production hosting environments are required to isolate active data partners from implementation and testing work being performed for the Distributed Query Tool and Portal 2.0 or any other related activity.
System Software	Development and Production hosting environment each require Windows Server, IIS, .NET and SQL Server as the operating environment.
Production System Monitoring	Internal monitoring for hardware, system software, or application software failures and remediation.
Ticketing System	System for logging, tracking, and auditing resolution of all incidents detected via monitoring or due to support calls.
Technical Support	**Technical / Customer Service and Support Hotline / Process Overview** Anyone experiencing technical issues involving use of the systems may call the hotline for support. The specific process works as follows: 1. **Call the Support Hotline:** (866) 624-2030 (Within the U.S.A) **/** (513) 768-3747 (International) **NOTE: ALL ISSUES THAT NEED IMMEDIATE ATTENTION MUST BE SUBMITTED VIA TELEPHONE.** The call center staff will enter a ticket and contact an "on-call" engineer. The on-call engineer will respond within 15 minutes. **Email Option for Non-Critical Support Needs** Non-critical issues can be submitted via email to: managedservices@lincolnpeak.com. A ticket will be entered into the tracking system. However, the call center will not notify the on-call engineer as these issues are not expected to be critical. On-call engineer will lead the technical support delivery team, keeping the Client Partner and Technical Lead aware of all issues. **2. For each support request, users will:** 1. Tell the call center customer representative which network (e.g., AHRQ, FDA Mini-Sentinel) they are calling about 2. Provide company name, your name, phone number and

	email address 3. Describe the issue
Software Patches	Application of software patches for the operating environment (Windows Server, IIS, .NET and SQL Server) and the DRN Portal application will be applied on a regular basis during regularly scheduled maintenance windows. Publishing of updates to the DataMart will occur on a regular basis.

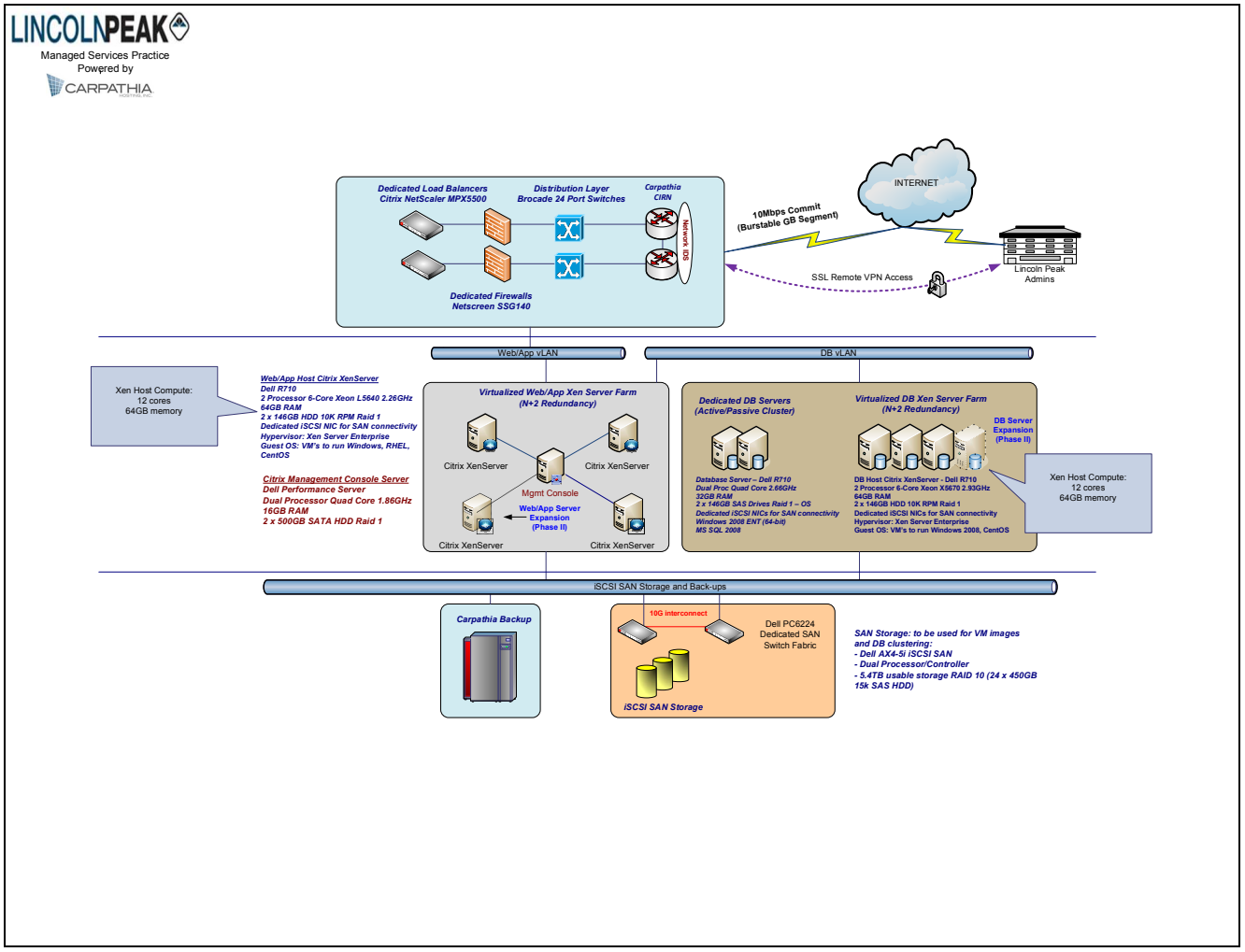
Table 4. Hosting, Security & Support: Detailed Requirements

Requirement	Description
Detailed Requirements	
Ping, pipe, power, connectivity, fire suppression, security.	Redundant TIER IV level network connectivity at LAN and WAN, HVAC, fire suppression, and power along with physical and video security monitoring.
Servers, Virtual Machines	Web servers are hosted in private cloud based on Citrix XenServer with redundant physical servers supporting automated failover and load balancing. Database servers are clustered physical servers. All servers or VMs are connected to RAID 10 iSCSI SAN for storage and SAN based backup.
System software	Windows 2008 Server, IIS, and SQL Server 2008.
Server maintenance	Regular maintenance windows to install system software and application software installation of patches and upgrades as well as server performance analysis.
Solution environment backup	Daily scheduled backup of the solution source and web server runtime environment.
Database backup	Full backup daily and incremental every 15 minutes. Stored onsite.
System event and SNMP trapping and notification	Trapping, alerting and responding to hardware, system software (operating system, database) and application software errors and notifications.

5.2. Hosting Design Overview

The hosting environment is operated at a data center provided by Carpathia Hosting, Inc. in Dulles, Virginia. Carpathia is provider of FISMA/ SAS-70 private cloud services and operates TIER III datacenters (TIER III covers full system redundancy and redundant commercial connections to major backbones). Specifically, Tier III is comprised of multiple active power and cooling distribution paths, has redundant components, and is fault tolerant, providing 99.995% availability. Carpathia has facilities in many major US cities and around the world and provides: redundant HVAC, redundant fire suppression, redundant power with UPS and generator backup. The facility is secured with man-trap entrances, photo identification validation, manned armed security tours, and video surveillance 24 hours per day, 7 days per week. Figure 2 illustrates the system infrastructure.

Figure 2. System Infrastructure



LPP's systems connect to the internet via a dual Juniper Router / Firewall / VPN concentrators that provide redundant connections to the internet with automatic failover. Each device has redundant power supplies connected to separate power circuits in the Tier III data center. The devices provide routing functions from the VLANs implemented on the redundant switches to the Internet. In addition to routing the systems provide firewall and VPN functionality. Firewalls are configured to restrict inbound traffic to only HTTP (port 80) and/or port HTTPS (443) to the web servers. All clients are assigned dedicated web servers on virtual machines. No direct inbound web access is allowed to the database servers. All database traffic is routed through the firewalls and limited to the appropriate web server. VPN is dual authentication requiring the use of an RSA token in addition to username/ password. The VLANs span the dual Ethernet switches

and dual physical NICs are teamed on the servers for production data providing 2GB bandwidth and redundancy in the event of NIC or switch failure.

The Application Portal is hosted in a two server configuration, one server (Portal web Server) to run the application and to service all applications requests that come in via the Web. This server runs the Portal application under IIS and ASP .NET. The second server (Portal Database server) houses the Portal Database in a MS SQL Server 2008 instance. Note that there will be no connection from the Portal Database server to the web. All requests will be made via the Portal Web server. Web servers are on virtual machines with support for load balanced web farms as utilization increases and database servers are physically clustered servers for FISMA compliance. Database server is replicated via log shipping to Carpathia Phoenix data center which is also FISMA compliant. Each server is hardened and performance tuned according to Microsoft best practice documentation. A third Management Server (not open to the Web and only available via Virtual Private Network) will be used by Operations Administrators to monitor the health and tune the Portal Web Server and the Portal Database Server.

5.3. FISMA Controls per NIST SP 800-53 Security Controls

LPP has contracted Caturano & Company (http://www.caturano.com/) to review all Lincoln Peak's Standard Operating Procedures (SOP) pertaining to Managed Services to determine required enhancements for FISMA compliance. Specifically, the system is designed to meet FISMA Moderate Risk security controls as specified in the National Institute of Standards and Technology (NIST) Special Publication 800-53 (http://csrc.nist.gov/publications/nistpubs/800-53-Rev2/sp800-53-rev2-final.pdf). The following is a list of applicable NIST SP 800-53 controls and a summary of Lincoln Peak's policies and procedures for each. These descriptions relate to internal LPP SOPs and policies, not those of the querying system.

Lincoln Peak Standard Operating Procedures per NIST SP 800-53 Security Controls

I. Lincoln Peak User Access Policy
    1. Provides policy to control who is allowed to access systems and how that access is managed.
    2. Logical Access
        i. New Hires/Terminated Users/Modifications/Contractors
            a) Documentation and verification of all account requests
        ii. User Access Review
            a) Periodic review of accounts to eliminate unnecessary accounts
        iii. Segregation of Duties
            a) Limiting functional access by role to ensure only properly trained, authorized MSP personnel have access to production equipment.
        iv. VPN Access
            a) Policy for issuing and managed dual token SSL based VPN for accessing all systems
        v. Domain Policies
            a) Active Directory and LDAP policies to control system access
            b) Passwords - 7 character minimum, strong password, quarterly change

c) Lockouts – 5 failed attempts results in locked account requiring administrator intervention

3. System Security
    i. Server/Network Configuration – security policies
        a) DMZ
            a. Web server and database server firewall configuration to prohibit external access to database servers and limit web server protocols/ ports
    ii. Secure Data Transfer
        a) FTP
            a. Limited to behind firewall for authenticated VPN users only
        b) Encryption
            a. All traffic behind traversing firewall is encrypted other than HTTP access to front end web servers by external users
    iii. Assessments and Certifications
        a) Penetration Testing
            a. Periodic testing of security
        b) Vulnerability Scanning
            a. Periodic scanning of ports and systems
    iv. Authorized Traffic
        a) Firewalls
            a. Firewall rules are created on a server by server basis to restrict inbound traffic to HTTP (port 80) and/or HTTPS (port 443) to web servers. Port 25 is available on request for SMTP. Additional ports are available if required and are documented through Change Management Process. Database servers have no direct inbound web traffic and are not NAT'd. DMZ firewalls limit access to each database server to the associated web server(s).
        b) Anti-Virus
            a. All servers must run NOD32 anti-virus
    v. Physical Access
        a) Third Party SAS70 Review
            a. Type II SAS-70 audit to be performed in Q4 2010.
4. Written Information Security Policy/Risk Policy – provides policy on high level controls for access and security monitoring as well as response in the event of an incident
    i. Protecting Data
        a) Both Physical and Electronic data are covered in this SOP.
    ii. Security Awareness Training
    iii. Incident Response
5. Business Continuity, Disaster Recovery Plan
    i. Policy and Plans for recovery of services in the event of data corruption/loss, component failure, system failure, site failure, and geographic failure (i.e., Natural disaster).

a) Data corruption/loss is addressed via backup/recovery policy
b) Component failure and system failure are addressed by in-device redundancy and overall redundant architecture of infrastructure providing near zero downtime for these conditions
c) Site failure is addressed via cold site in Phoenix AZ that is FISMA compliant with log ship database replication and webserver daily backup and copy to remote SAN allowing 72 hour configuration and recovery RTO and 15 minute RPO.

6. Change Management Policy
    i. Policy and procedure for reviewing and approving all change to production environment to ensure no unexpected results
    ii. Security Impact Analysis
    iii. Change requests
    iv. QA testing/end user testing
    v. System Backup
    vi. Change Approval prior to Implementation
7. Software Development Life Cycle
8. Maintenance Policy
    i. Policy for the control of system maintenance such as OS and application patches
    ii. Establishes maintenance schedule
    iii. Establishes resource and financial budgeting
9. Vendor Management Policy
    i. Policy for the review, approval, and control of vendors as they pertain to managed services
10. Human Resources Policy – Policy and procedure for review and approval of employee and contractor candidates
    i. Candidate screening including background and reference checks.
    ii. System security awareness policy/training

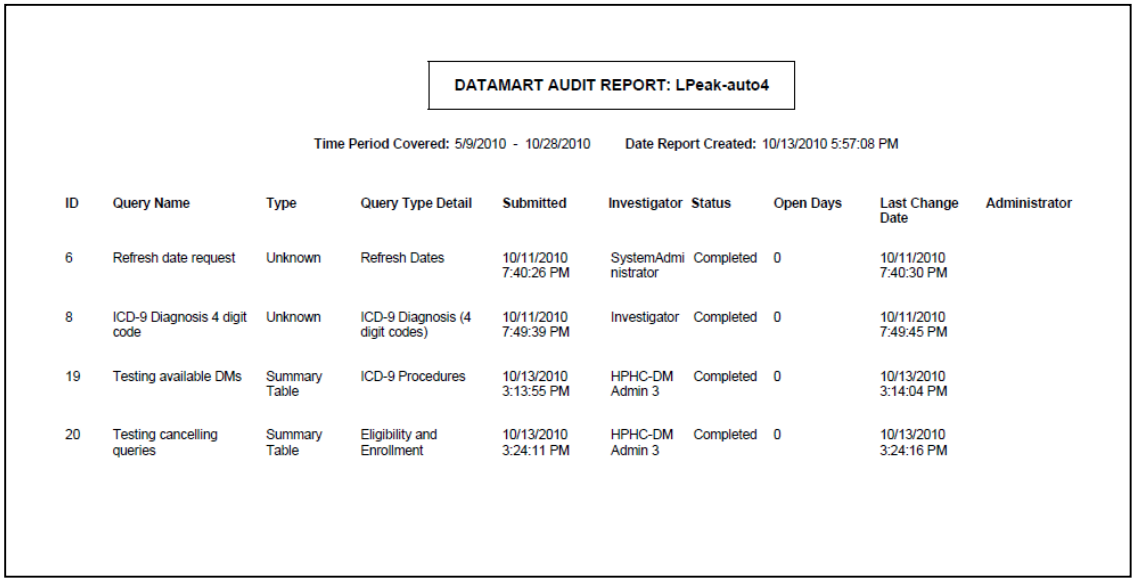
5.4. Security Specifications

The following list contains major security specifications of the system.

- Users are required to select strong passwords with the following rules: at least 7 characters, at least 1 number, at least one nonnumeric character, at least one capital letter, at least one lower case letter. Passwords cannot contain the user name or any part of the user's full name.
- The system will force users to change their passwords every six months.
- Passwords cannot be re-used.
- The system will automatically log users off after thirty minutes of inactivity.
- The system will automatically delete all query results after one year.
- The system will automatically delete file transfers after 21 days.
- System Administrators will verify user identities and email addresses before creating new user accounts.
- Users must use corporate email addressed for network communication.

- The system will audit all system activity (access, user ID changes, query initiation, results upload, etc.) and will regularly review audit logs to look for inappropriate system use.
- Antivirus software will run regularly on all system servers.
- DataMart Administrators will be notified of relevant changes within the system such as the addition of a new user or DataMart.
- DataMart Administrators will be able to create audit logs of all activity related to their DataMart; see screenshot below for an example audit report.

Figure 3. Sample DataMart Audit Report



DATAMART AUDIT REPORT: LPeak-auto4

Time Period Covered: 5/9/2010 - 10/28/2010 **Date Report Created:** 10/13/2010 5:57:08 PM

ID	Query Name	Type	Query Type Detail	Submitted	Investigator	Status	Open Days	Last Change Date	Administrator
6	Refresh date request	Unknown	Refresh Dates	10/11/2010 7:40:26 PM	SystemAdministrator	Completed	0	10/11/2010 7:40:30 PM	
8	ICD-9 Diagnosis 4 digit code	Unknown	ICD-9 Diagnosis (4 digit codes)	10/11/2010 7:49:39 PM	Investigator	Completed	0	10/11/2010 7:49:45 PM	
19	Testing available DMs	Summary Table	ICD-9 Procedures	10/13/2010 3:13:55 PM	HPHC-DM Admin 3	Completed	0	10/13/2010 3:14:04 PM	
20	Testing cancelling queries	Summary Table	Eligibility and Enrollment	10/13/2010 3:24:11 PM	HPHC-DM Admin 3	Completed	0	10/13/2010 3:24:16 PM	

6. List of Additional Material

1. Manuscript 1: Maro JC; Platt R; Holmes JH; Strom BL; Hennessy S; Lazarus R; Brown JS. Design of a national distributed health data network. Ann Intern Med. 2009 Sep 1;5:341-344. (*Annals_Maro_2009.pdf*)

2. Manuscript 2: Brown JS, Holmes JH, Shah K, Hall K, Lazarus R, Platt R. Distributed health data networks: a practical and preferred approach to multi-institutional evaluations of comparative effectiveness, safety, and quality of care. Medical Care 2010; 48:S45-51. (*Brown Distributed Research Network Medical Care 2010.pdf*)

3. PowerPoint presentation 1: Overview presentation of the query tool and Portal presented at the 2009 HMO Research Network conference. (*DRN_HMORN2010_presentation 2010.07.20.pdf*)

4. PowerPoint presentation 2: Overview presentation of the query tool and Portal with additional details presented as part of the FDA Mini-Sentinel project. (*FDA_MS_QueryTool_Introduction06282010.pdf*)

7. References

1. Brown J, et al., *Proof-of-principle evaluation of a distributed research network. Effective Health Care Research Report No. 26. (Prepared by the DEcIDE Centers at the HMO Research Network and the University of Pennsylvania Under Contract No. HHSA29020050033I T05.) Rockville, MD: Agency for Healthcare Research and Quality. June 2010. Available at: http://effectivehealthcare.ahrq.gov/reports/final.cfm.* . 2009, AHRQ.
2. Brown J, et al., *Blueprint for a distributed research network to conduct population studies and safety surveillance. Effective Health Care Research Report No. 27. (Prepared by the DEcIDE Centers at the HMO Research Network and the University of Pennsylvania Under Contract No. HHSA29020050033I T05.) Rockville, MD: Agency for Healthcare Research and Quality. June 2010. Available at: http://effectivehealthcare.ahrq.gov/reports/final.cfm.* 2009, AHRQ.
3. Brown, J.S., et al., *Distributed health data networks: a practical and preferred approach to multi-institutional evaluations of comparative effectiveness, safety, and quality of care.* Med Care. **48**(6 Suppl): p. S45-51.
4. Brown JS, et al., *Design specifications for network prototype and cooperative to conduct population-based studies and safety surveillance. Effective Health Care Research Report No. 13. (Prepared by the DEcIDE Centers at the HMO Research Network Center for Education and Research on Therapeutics and the University of Pennsylvania Under Contract No. HHSA29020050033I T05.)* 2009, Agency for Healthcare Research and Quality: Rockville, MD.
5. Brown, J.S., et al., *Active influenza vaccine safety surveillance: potential within a healthcare claims environment.* Med Care, 2009. **47**(12): p. 1251-7.
6. Maro, J.C., et al., *Design of a national distributed health data network.* Ann Intern Med, 2009. **151**(5): p. 341-4.
7. Moore, K.M., et al., *Potential population-based electronic data sources for rapid pandemic influenza vaccine adverse event detection: a survey of health plans.* Pharmacoepidemiol Drug Saf, 2008. **17**(12): p. 1137-41.
8. Go, A.S., et al., *The Cardiovascular Research Network: a new paradigm for cardiovascular quality and outcomes research.* Circ Cardiovasc Qual Outcomes, 2008. **1**(2): p. 138-47.
9. Magid, D.J., et al., *Creating a research data network for cardiovascular disease: the CVRN.* Expert Rev Cardiovasc Ther, 2008. **6**(8): p. 1043-5.
10. Velentgas, P., et al., *A distributed research network model for post-marketing safety studies: the Meningococcal Vaccine Study.* Pharmacoepidemiology and drug safety, 2008. **17**(12): p. 1226-34.
11. Wagner, E.H., et al., *Building a research consortium of large health systems: the Cancer Research Network.* J Natl Cancer Inst Monogr, 2005(35): p. 3-11.
12. Federal Immunization Safety Task Force. *Federal Plans to Monitor Immunization Safety for the Pandemic 2009 H1N1 Influenza Vaccination Program*. 2009 [cited 2010 January 31]; Available from: http://www.flu.gov/professional/federal/fed-plan-to-mon-h1n1-imm-safety.pdf.