Vermont Health Care Uniform Reporting and Evaluation System (VHCURES)

Request for Information

Prepared for:

The Green Mountain Care Board

Submitted by:

The Lewin Group, Inc., an Optum Company

Contact: Sue Bembers
703.269.5684

3130 Fairview Park Drive
Suite 500
Falls Church, VA 22042
Fax: 703.269.5501

sue.bembers@lewin.com

September 27, 2017, 2:00 pm

LEWINGROUP®

September 27, 2017

Erin Collier, Purchasing Agent
State of Vermont, Green Mountain Care Board
89 Main Street
Montpelier, VT 05620

Reference: Response to State of Vermont, Green Mountain Care Board VHCURES 3.0 RFI

Dear Ms. Collier:

The Lewin Group, Inc. (Lewin) is pleased to provide our response to the State of Vermont, Green Mountain Care Board (State) for the Vermont Health Care Uniform Reporting and Evaluation System (VHCURES) 3.0 request for information (RFI) issued September 6, 2017.

Please direct all future communications regarding this opportunity to:

The Lewin Group, Inc.
Attn: Sue Bembers, Director of Contracts
3130 Fairview Park Drive, Suite 500
Falls Church, VA 22042
Phone: 703-269-5684
Fax: 703-269-5501
Email: sue.bembers@lewin.com

We look forward to working with you on this important project.

Sincerely,

Lisa Chimento
Chief Executive Officer

Table of Contents

Budget Estimate

The annual cost for operating an All Payer Claims Database (APCD) like VHCURES will be determined by several choices that we anticipate the Green Mountain Care Board (GMCB) will specify in the forthcoming RFP. Lewin estimates that the annual cost for the first implementation year could have a minimum cost of $500,000 to over $2 million. Annual operating costs could be lower in following years depending upon the amount of new development that occurs. Some of the factors that will have an impact on these estimates that GMCB needs to consider in developing the RFP are listed below.

- **Frequency of Data Submissions –** Each data submission will require vendor resources to intake the file and validate the file. Monthly submissions will require up to three times as many resources than quarterly submissions. Requesting quarterly versus monthly submissions will reduce the annual operating costs for VHCURES 3.0.
- **Hosting Requirements –** Hardware costs to host VHCURES will be dependent upon the number of environments, amount of data that will be hosted and server capacity specified by GMCB. Each additional environment, uptime requirements and disaster recovery requirements will impact hardware costs.
- **Data Warehouse Features –** In addition to housing the data, GMCB can require additional functionality which will impact the data warehousing costs. Additional features such as metadata management, work flow, business rules and additional features will impact the information available to GMCB, but also increase the cost of building the data warehouse.
- **Data Analytics Reports –** The number and type of data analytic reports required for VHCURES 3.0 will also impact vendor costs during the first few years of implementation. The tradeoff is the marketability of VHCURES 3.0 and reductions in the need to hire vendors to build data analytics on other GMCB projects.
- **Authorized Users –** The number of users authorized to query the data warehouse and access the data analytic reporting system will impact hardware costs and potentially license costs. Another factor that will influence cost is the permissions that each user has to access the data. Power users that can conduct queries will consume more resources than users who are allowed to view published reports. Vendors will need to size their hardware solution accordingly based upon the number and types of users requested.

Based upon the choices that the GMCB makes for each of the aforementioned factors there is potentially a wide range in estimated costs. Based upon our prior experience the table below provides scenarios for potential low, medium and high cost solutions.

VHCURES 3.0 Budget Estimates

Budget Scenario	Data Submission Frequency	Hosting Requirements	Data Warehouse Features	Data Analytic Reports	Authorized Users
Low Cost $[redacted]	Quarterly	Production Environment	SQL Data Warehouse with no Ad-Hoc Access	20 Data Analytic Reports, 1 format for data extracts	Tableau Reports Delivered by Tableau Reader
Medium Cost $[redacted]	Monthly	Production and Fail-Over Environment	SQL Data Warehouse with Ad-Hoc Access	20-50 Analytic Reports, 1 format for data extracts	Tableau Reports Delivered by Tableau servers with up to 20 users with report view access
Higher Cost $[redacted]	Monthly	Production and Fail-Over Environment	SQL Data Warehouse with Ad-Hoc Access, Meta Data Manager, Work Flow Tracking, Business Rules Editing	20-50 Analytic Reports, custom data extracts	Tableau Reports Delivered by Tableau servers with up to 50 users with report view access and up to 10 super users

5 Anticipated Requirements

5.1 Data Collection, Cleansing, Consolidation, and Distribution

Data intake, validation, data submission management, data preparation, consolidation, distribution, and documentation are critical activities that form the foundation for any data warehouse solution.

As a part of the data collection process, we recommend that the data submissions should be thoroughly validated using control totals, formatting checks, data volume over time, key field validity, missing values, distinct values and data integrity linkage checks. These validation processes ensure integrity and stakeholder confidence in the final analytics that are accessible to end users. The cleansed data should then be consolidated into a standard format for enrichment purposes. We also recommend that the results of these validations be distributed in the form of interactive reports to the data submitter and be reviewed and approved by both the contractor and GMCB to ensure transparency.

Lewin has performed data collection and validation, data transformation and enhancement with groupers, and quality control analysis for both state and federal clients and has a strong portfolio of expertise in claims analysis, aggregating and enhancing datasets, and secure storage and

management of data. Lewin understands that VHCURES 3.0 will process claims similarly to the current 1.0 implementation. Our solution for data collection will mesh seamlessly with the current vendor's Submission Guide, using the same fields from current and new data contributors including mapping on Medicare data. Lewin will work with new and existing VHCURES data contributors to identify any new data fields that may be necessary to include Medicare data or merge data sets appropriately.

Lewin has extensive experience working with All Payer Claims Database (APCD) data collection and validation with Medicare data through our work with the Wisconsin Health Information Organization (WHIO) and state Medicaid data through the CMS Payment Error Rate Measurement (PERM). Lewin is able to accommodate claims data submission as often as monthly and perform detailed data validation for all data contributors. Lewin's work on PERM includes collecting Medicaid claims, eligibility, and provider data from multiple states and from multiple sources within states. Lewin works with each data contributor to understand the nuances of the data. In doing so, we ensure that claims data is correctly and accurately recorded, securely transferred, stored, and combined into useable datasets. In order to maintain a level of standardization across the data sources, we also apply a standard stratification technique.

Lewin has collected detailed data from multiple payers across Medicaid, Medicare, and commercial lines of business. Data is collected via secure means and stored on encrypted servers. The data is then run through a standard validation and review process prior to use. Lewin produces validation reports which are reviewed by both Lewin and the APCD client to determine any issues with the data. Should issues be found with the data, we communicate with the data contributor and our APCD client to identify and resolve the issue, including collecting new data, if necessary. Occasionally, a data contributor may not be able to submit a timely and satisfactory data set. When this occurs, it is important to have a policy regarding the handling of late and/or unsatisfactory data. The policy should be clear whether the APCD will postpone the data load due to a data contributor submission issue or if the data processing proceeds without the data contributor's data set. In the latter case, the data is processed in the next data collection cycle. This process should be reviewed and revised regularly to incorporate any lessons learned and process improvements that will improve the efficiency of the collection and validation process.

Data Integration and Enrichment

The cleansed data can be enriched through episode groupers and risk assessment software and transformed into a format that lends itself for analytics. We recommend that the data should be transformed into a series of dimensions and measures, known as a dimensional model that is easily understood by end users and technical staff. Measures are numeric representations of a set of occurred events (e.g., total dollars paid, number of claims, member months, etc.). Dimensions provide structured labeling information to these measures (e.g. Category of Service, Claim Type, Location, Episode Treatment Group, Condition, etc.). The cleansed and enriched data should then be consolidated into a final named version that can be loaded into the data warehouse in compliance with the principles of Master Data Management (MDM). This ensures that stakeholders can track any data elements from source to destination and keep track of data lineage. The data in the data warehouse should act as a source of truth for any analytics and reporting needs.

Lewin's data integration routines create value added analytical variables and run data through the data enrichment programs such as episode groupers, risk groupers, and quality measurement software. The MS-DRG grouper should be used to assign Diagnosis Related Group's DRG's to inpatient claims. In addition, an APCD may want to utilize software to normalize service pricing, employ separate, configurable modules to create common identifiers for patients and providers, and enforce an agreed-upon process for data masking to produce Protected Health Information/Personally Identified Information (PHI/PII) free reports. These de-identified and masked reports can be available to the public.

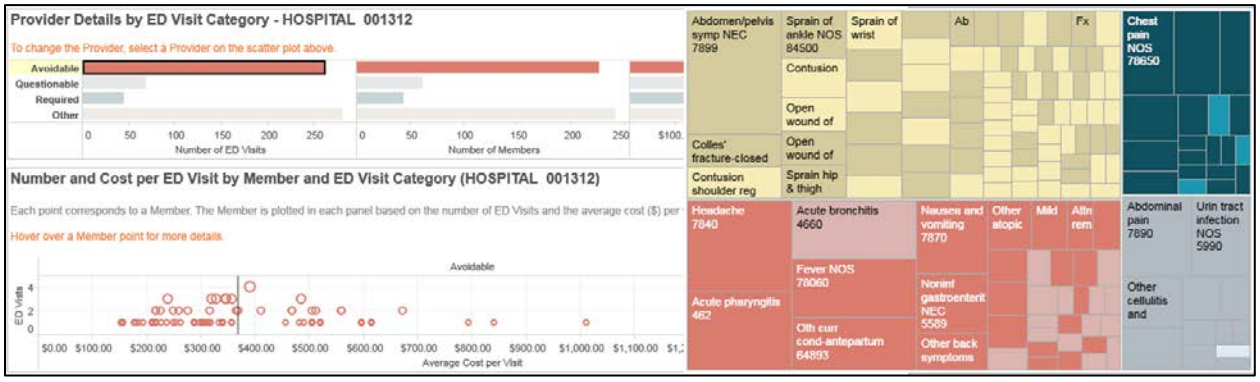
In order to deliver a data management and analytics solution, our skilled team of health care data analysts has developed HealthView with a goal of eliminating fragmented and inconsistent information. HealthView emphasizes the data integration, flexibility, and required customization to arrive at timely actionable intelligence.

HealthView's data warehouse provides a comprehensive shared metadata environment for consistency in data definitions. HealthView's analytics system pre-determines dimensions and measures for ease of use, with the capability of adding even more. The flexibility in our data model combined with the powerful metadata modeling and mapping capabilities within our Business Intelligence (BI) platform enable us to deliver data-dense and complex analyses in the form of visually appealing and discerning visualizations. For example, **Exhibit 1** shows a multifaceted visualization that classifies emergency department visits, displays associated members and their diagnoses, and their costs in a clear and simple view.

HealthView's features include:

- Integrated claims with non-claims data
- Comprehensive, flexible and intuitive visualizations
- Role-based security (Table, column, row, summary and detailed level of access)
- Standard and custom reporting packages
- Ad hoc visualization development
- Custom data sets

Exhibit 1: Visualizing Emergency Department Visits in HealthView



5.2 Master Person Index

Lewin has successfully created a master person index (MPI) for other APCD clients. When creating a master person index, a Universal Identifier (UID) should be assigned to each member. A complete member lookup dataset should be maintained with UID, corresponding members and contributors. After data cleansing, members are joined using the UID, then by their identifiable

information. We standardize and clean multiple member identification fields. Invalid social security numbers (SSNs) should be removed, street addresses standardized for suffixes and prefixes, and first names are formalized. Many different combinations of name, address, zip code, sex, date of birth, phone number, and Soundex names should be used to optimize member matching. Different keys can be generated based on whether the match is an intra-plan match and on the availability of a SSN. New members should be assigned a UID while reoccurring members are joined with their old records. Multiple field combinations compensate for missing or one-off values while reducing false positives. Using this approach, an APCD can successfully match members across many different data submitters to create an MPI.

5.3 Master Provider Directory

In building a Master Provider Directory, an APCD should require data submitters to provide the National Provider Identifier (NPI) information with each claim. NPIs can be validated using the NPI Check Digit and the Luhn formula. The Luhn formula and check digit ensures that the NPI meets the specifications detailed in the Final Rule for the Standard Unique Health Identifier for Health Care Providers (69 FR 3434). Unique providers are identified from the claims and matched with NPIs from the National Plan & Provider Enumeration System (NPPES) to create a Master Provider Directory. This approach is straightforward and ensures accuracy in reporting.

5.4 Data Warehouse/Analytic Enclave

5.4.1 Secure Container

A data warehouse cannot be more secure than the physical platform on which it runs. Therefore, it is critical that GMCB plans and evaluates the security of the data warehouse and the overall solution.

The facility hosting the data warehouse should be designed to run 24 x 7 and employ various measures to help protect operations from power failure, physical intrusion, and network outages as described below.

- Facilities Access: The facility should utilize industry standard access mechanisms to protect physical infrastructure and datacenter facilities. Access should be limited to a very small number of operations personnel. Physical access, and the authority to approve facility access, should be controlled by authorized operations personnel.
- Power Redundancy and Failover: The hosting facility should have the ability for off-grid operations in order to plan for a power outage.
- Media Disposal: Upon systems end-of-life, rigorous data handling procedures in compliance with Health Information Portability and Accountability Act (HIPAA) should be utilized to dispose data and hardware.

In addition to these physical security controls, we also recommend additional security devices as described below.

- Network access controls for administrator access: Role-Based Access Control (RBAC) should be employed to enable fine-grained access for administrators. Only designated staff with network access credentials should be allowed to administer the network.

- Security devices used to protect the infrastructure: Security controls such as service endpoints to ensure that the firewall allows traffic only from designated sources can be employed to secure access to sensitive information.
- Change control processes for all systems: GMCB should work with the data warehouse contractor to ensure that change control requests are performed only after approval is obtained from designated staff such as the project director and project manager.
- Host-based firewall and anti-malware protection: Antimalware that helps identify and remove viruses, spyware, and other malicious software should be employed to ensure that the data should not be corrupted due to a malware attack.

In order to provide a secure container for our solution, we also recommend that an APCD solution should be deployed in a FedRAMP and NIST compliant data center.

GMCB can also consider a cloud-based solution. A cloud based-solution is ideal when there is a need to scale up or scale down the hosting infrastructure on a very frequent basis. However GMCB should consider the relatively high costs of implementing security and compliance controls within a cloud-based solution.

5.4.2 Controlled Access

In order to control access to various section of the data warehouse, the solution must be capable of assigning varying levels of permissions to its diverse data customers. Customer access must be restricted to approved data sets and uses, at the organizational and individual user levels.

To achieve such secure access to data and retain flexibility in data operations, our solution, HealthView, employs robust role-based security as described below.

- Users should be assigned to roles, based on approved access levels. For example, a submitter access is limited to their organization's data quality reports only.
- The APCD should be flexible enough to obfuscate/mask sensitive data fields to certain users at runtime based on their role. For example, in the same visualization, a user with elevated privileges can view the patient's address but another user with lower privileges cannot.
- The APCD's analytic system can also be configured to restrict access to individual data elements at the column or row level within a visualization or report. For example, a user can be configured to access non-PII fields of just pharmacy claims, within the entire claims dataset.
- Additionally, the APCD should be able restrict or enable a user's access to data at a summary level or a detailed record/row level.

Such data and user access should also be auditable to answer common questions as they pertain to the data warehouse.

- Who did "what" in the data warehouse?
- Who changed the records?

- Who deleted the records?
- When was that data changed?
- What are the data changes on a specific table in a specific time period?
- What rows have changed for a user table?
- Has a row changed?

The APCD should be fully capable of auditing and logging events at three levels of detail:

1. The database: data backups, add/delete/edit users, permission changes at the database level etc.
2. Individual data objects: Select/Insert/Execute/Delete queries, permissions on the objects (tables, functions etc.)
3. Audit objects: Changes to the audit policy specification and logs

Such data warehouse audit logs and analytics audit logs should be made available to your staff according to your data governance policies.

5.4.3 Improved Processing Capabilities

One of the primary objectives of building a data warehouse is to propagate the data in the warehouse to the data analysts and business intelligence visualizations in order to deliver actionable intelligence to the right person at the right time with minimal data preparation lag. In order to minimize the time span between data collection and analytics, we recommend the following guidelines.

- **Collaborative Development Process**: An iterative and collaborative methodology ensures that data is processed as efficiently and accurately as possible. We recommend that analytic solutions must be designed with continual involvement from end users and project leadership from stakeholders.
- **Automated Reusable Extracting, Transforming and Loading ETL Processes:** We recommend a development approach is focused on building ETL components as reusable libraries. This ensures flexibility and reusability and allows the incorporation of new data elements as necessary in a cost effective manner.
- **Scalable Infrastructure:** The solution must scale-up and/or scale-out as necessary to adjust to user and data volume growth, while reducing data processing times.
- **Data Optimization for Analytics**: A data warehouse solution should minimize latency when executing queries on large tables. We recommend using data compression techniques and in-database analytics to improve the performance of analytics queries.
- **Targeted Data Marts:** We recommend building targeted data marts that stores data from the last two years. In our experience, this data will be accessed very frequently. Hence, queries and visualizations that run against this data mart will be executed and results will be rendered quickly for the end user.

HealthView's data warehouse platform supports the features mentioned above. We also realize that our clients would benefit from advanced analytics and data science through machine learning. Hence we have incorporated support for such analytics directly into our platform. We also recommend that the visualizations built on top of the APCD should support intuitive drag and drop features and should be accessible through any industry standard browser and on mobile devices.

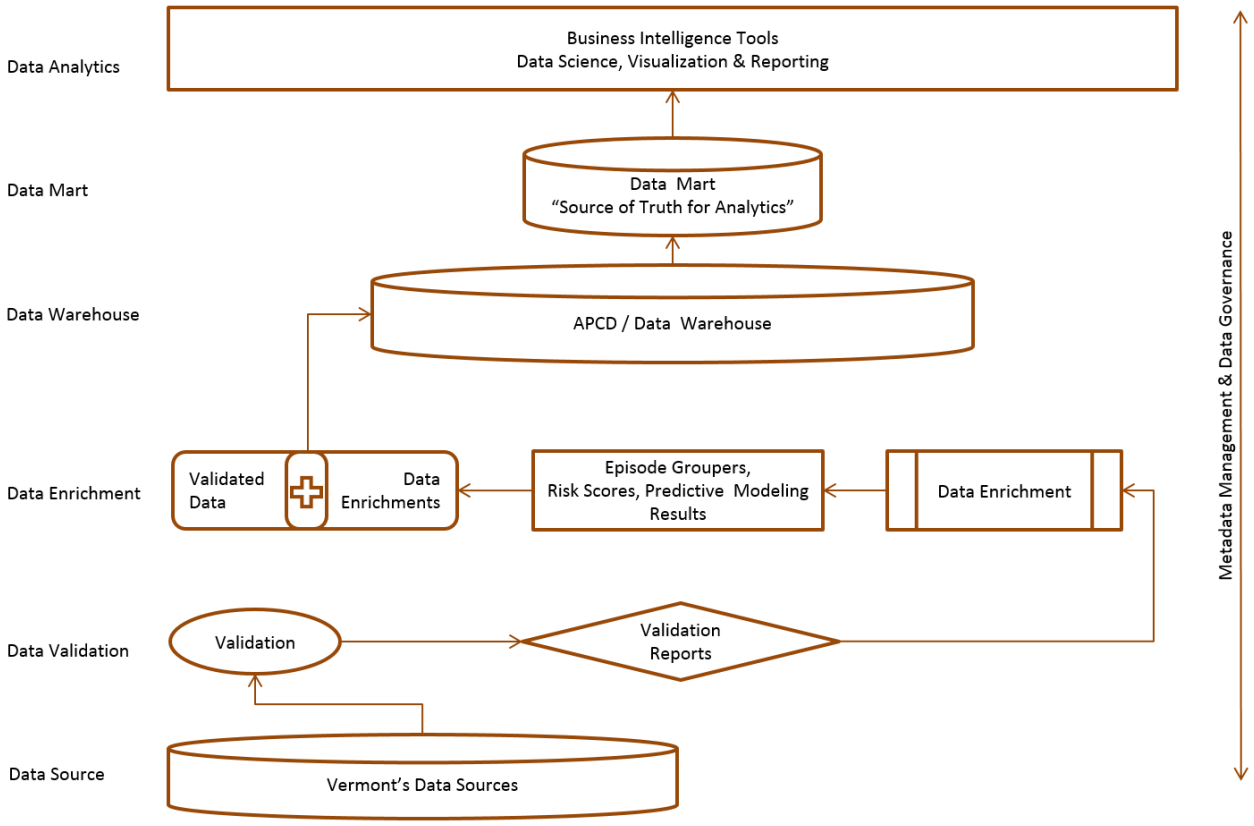
5.4.4 Data Management

The objective of Data Management is to ensure the highest level of data quality in a data model that lends itself to productive analytics leading to timely and accurate decisions. Such a data warehouse should enable end users to extract datasets and standard reports on demand. It should also have the capability to allow the end users to create custom datasets and ad-hoc reports when necessary. We described the functionality associated with such features in Section 5.4.3-Improved Processing Capabilities.

In order to handle requests for data, ad-hoc reports, information queries and their associated approval processes (data use agreements, store supplemental material, user affidavits, periodic status checks on agreements etc.), we recommend the use of standard web based software such as SharePoint. SharePoint is available in most organizations and has the ability to setup and customize workflow processes.

We believe that an analytics solution should be designed to enable users to see the cause, effect and meaning in the data, leading to actionable intelligence. Therefore, our approach will emphasize clarity through data visualization and encourage users to dig deeper into analysis of the data. HealthView's overall solution architecture, as applicable to VHCURES, is illustrated in **Exhibit 2**.

Exhibit 2: HealthView's Architecture and Data Flow



In order to deliver such analytics to GMCB, we envision the process described below.

Extract, Transform and Load (ETL): Extract data from the data sources and perform standardized validation checks in preparation for enrichment and downstream analytics. The data is then de-identified to handle sensitive PHI/PII data. Data lineage should be tracked to re-identify the data if necessary. For any reason, if GMCB prefers not to de-identify the data, then features such as data masking could be utilized to handle sensitive data while ensuring the appropriate level of privacy.

Enrichment: Enrich the data using industry standard groupers to aggregate claims into episodes of care and provide risk scores by member. In addition to these features, we recommend that GMCB utilize groupers that calculate quality measures specified by CMS and other data stewards. The enrichment process should also add analytical flags for key measures such as readmissions and emergency room visits.

Data Mart: Transform the enriched data into a flexible data model and make it readily available to analysts and decision makers. The objectives of the data mart are complete visibility, interactivity and ease of use. In order to achieve these objectives, we recommend that the data be stored in the data mart in a flexible data model that allows easy analysis across all dimensions and measures. The analytical system should also analyze GMCB's data usage pattern to

proactively suggest, refine and tune the schema for the best performance to satisfy your analytical requirements.

5.5 Public Use Data and/or Analytic Files

In order to support the enhanced data analytics reports in VHCURES 3.0, GMCB should consider requiring that enriched Analytics Data Marts (ADM) be used for public use data or analytic files provided to VHCURES customers. The ADM can be used to support the analytic reports and enable power users to conduct ad hoc analyses to answer more complex data analytic questions.

The ADM should incorporate data enrichments that GMCB can use to evaluate the health status of members, the quality of care they receive, and the cost of the episodes of care provided to members. Including these types of enrichments will support a wide range of data analytic reports and ad hoc analytics. These enrichments will also provide GMCB with the basic data analytic building blocks that are needed to build the data analytic reports required to monitor the Vermont All-Payer Accountable Care Organization Model (APM) waiver.

VHCURES should consider a combination of commercial off-the-shelf (COTS) software and open source software for data enrichment including: risk adjustment, quality measurement, episode of cares, disease condition grouping, and normalizing service pricing. Open source software, such as Clinical Classification Software (CCS), developed by AHRQ, is useful in categorizing individuals with specific disease conditions and calculating costs of individuals with those conditions. Open source software, such as CCS, allows researchers to compare their own findings with other research. In addition, synthetic databases can be produced in the ADM that allow a researcher to test their own analysis using the open source code.

COTS software, on the other hand, can provide value-added analysis of an APCD data set. Commercially available groupers provide the advantage of frequent updates to keep pace with evolving health care data analysis requirements. The specifications for quality measures, for example, change frequently and commercially available quality measurement tools are more capable of keep up with the changing technical specifications.

GMCB should also consider requiring that VHCURES 3.0 incorporates a reporting tool that provides them with rapid development capabilities and flexibility. A hard coded reporting environment will increase development efforts and potentially lock GMCB into an inflexible release schedule where new reports are published on an annual and semi-annual basis. In our experience we have found that data visualization and BI software provides the rapid development and flexibility to meet the data analytic reporting requirements of the ever changing health care system.

The approach described above would also enable GMCB to collaborate with other APCDs to incorporate similar data analytics. This is the approach that Optum and Lewin are currently following in building the APCD for the State of New York. Claims, eligibility and provider network data submitted by commercial payers, Medicaid and Medicare will be validated and stored in a data warehouse. This data will then be enriched by claims groupers and other data analytic modules and stored in the ADM. The data visualization and BI reports will then utilize the ADM to provide data analytic reports to users.

Business Intelligence and Ad Hoc Reporting

An APCD should deliver a solution designed to provide a flexible BI solution to meet the diverse analytic needs of your clients. Our experience providing data analysis and actuarial services for the GMCB State Innovation Model (SIM) Model Testing Grant and with other states APCDs puts us in a unique position to successfully integrate complex and large volumes of data from many and varying data submitters and deliver actionable intelligence. HealthView delivers a comprehensive, flexible, and innovative data warehouse, business intelligence, and data visualization platform.

By combining the strengths of Commercial Off-The-Shelf (COTS) and open source groupers, HealthView will provide VHCURES' internal and external users secure role-based access to the underlying data and interactive visualizations. HealthView integrates health care claims with clinical, socio-economic, and patient reported outcomes to expand your analytic potential. Authorized users can access a standard set of reports, custom reports, or a combination thereof as directed by the VHCURES. Roles can be set to expire after a specified time period, enabling the delivery of pre-built, standard offerings and custom reporting packages as subscriptions to your customers. HealthView also delivers data in the form of datasets, text files, database export files, and data extracts to facilitate visual analysis. VHCURES' designated users will be able to develop and save custom visualizations through an intuitive drag and drop user interface for ad hoc analysis.

Through these powerful capabilities, HealthView provides an empirical framework that systematically delivers on-demand health care analytics to enable VHCURES to optimize performance of the changing Vermont health care system. HealthView's functionality and release schedules are flexible to meet your customizations and change orders. We will also add new visualizations to our analytics suite.

HealthView is designed to bring data to life. Using an innovative and creative approach to visualizing health care data, HealthView enables researchers, analysts, and decision makers to identify and understand cost drivers, quality performance, trends, and many other aspects of the health care system. Our visualization team will collaborate with VHCURES to determine the best design for data visualizations. We frequently use stratification to drill down into cost, quality, and utilization metrics to identify specific people, provider groups, and geographies where variation occurs. HealthView's mapping capability allows the user to quickly identify hotspots for your important health care metrics. Our data visualizations are designed to allow the user to drill down to the patient/provider level. For example, a quality measure such as HbA1c testing in diabetics can be visualized on a map by county. Further drilling, can take the map to the zip code or census tract level. Doing this, shows the user locations performing well or poorly. In addition, using patient/provider attribution, HealthView will list providers' performance with quality measures, such as HbA1c test and allow the user to list diabetic patients who have not had an HbA1c test.

HealthView delivers actionable information where it is needed most. HealthView's purpose is the fulfillment of the Triple Aim: improving the patient experience, improving population health, and lowering the cost of care. HealthView's five areas of analytic services (Population, Provider, Financial, Ecosystem and Engagement) provide an empirical framework to systematically feed

operational processes with the right information at the right time to drive continuous improvement.

Data Extracts

Making Vermont health care data available to researchers, state agencies, health plans, and health care providers is the core purpose of an APCD. Health care data, including medical claims, pharmacy claims, and clinical data, can be used to better understand the dynamics of a state's health care system. Since the value of an APCD is derived from independent analysis from APCD customers analyzing the data, The Lewin Group provides a number of options and formats to provide data extracts.