INTERIM REPORT SUMMARIZING DATA
AGGREGATION METHODS IN USE TO DATE AND THEIR
STRENGTHS AND WEAKNESSES

Les Morgan / October 20, 2017

Center for Elder Care and Advanced Illness
Table of Contents

Introduction .................................................................................................................................. 3
Classification Variables ............................................................................................................... 3
Summary of Architectural Models for Aggregation of Care Plan Data ........................................ 4
Architectural Model Charts .......................................................................................................... 7
  Single Data Source versus Multiple Data Sources ................................................................. 7
Data Structure and Storage Methodology .................................................................................... 8
Implementation Strategy ............................................................................................................. 9
Workflow Integration ................................................................................................................... 10
Patient and Caregiver Integration ............................................................................................... 11
Introduction

This is a high-level Information Technology (IT) architectural review discussing major variations in system design approaches for care plan data aggregation, based on a content analysis of interviews with vendors and some key users. A series of architectural diagrams is included to classify approaches. Our interviews with technology vendors are ongoing, and this preliminary summary of methods will evolve as we see more examples of working systems.

Classification Variables

In approaching a systems analysis of IT solutions for aggregation of care plan data, many classification variables affect the overall architecture of a system. In IT vendor analysis, these are issues that could be specified in detail in a Request for Proposals (RFP) or Request for Information (RFI). Some of the most important variables that have emerged during our interviews are as follows:

- Where is the data located? Is there just one data source, or multiple sources? Are all the data sources covered entities, or are there some non-covered entities (such as community organizations)?
- Does the data reside in an EHR (Electronic Health Record) system, or not?
- Is the data stored in a relational database, or not? Data not in a relational database includes Excel files, older systems with record-format data stored in non-relational (flat) files, paper records, etc.
- For electronic data, is the data structured or unstructured? Unstructured data includes free text, PDF files, “blobs” of digital data, images, etc.
- Is the implementation strategy Premise-based or Cloud-based (Hosted)? If Cloud-based, does the Vendor rely on a utility cloud service such as Azure or AWS?
- What analytic tools and strategies does the system support? Systems may offer reports native to product, have a traditional report generation writing feature, or may focus on export of data to an external analytic system (such as Tableau).
- Does the system support use of advanced analytic methods such as Machine Learning and classification by Artificial Intelligence?
- Does the system integrate with workflow at the point of service delivery, or not? If there are workflow features, are there real-time alerts or decision tree features that are integrated with service delivery contact points?
- Does the system incorporate patient and family input directly? Is there a patient portal?
- Can the system accept input data from connected monitoring devices worn by or used by the patient (that is, does it support an Internet of Things (IoT) distributed data input model?
- Does the system have robust tools for handling privacy, security and data governance issues?
Summary of Architectural Models for Aggregation of Care Plan Data

Specific data systems may mix and match the classification variables described above, resulting in a very large number of possible scenarios. An analyst can use the model classification charts at the end of this report to determine where a specific data system fits on several key issues.

Each architecture and its implementation will have different considerations regarding privacy, security, and data governance. Evaluation of those matters is a complex issue that will be addressed by a separate report later in this project.

Some of the major variations in data aggregation models in use are as follows:

1. **Create a centralized data warehouse by extracting data from multiple feeder systems and analyze the data there using traditional report generation methods. Example: San Francisco Department of Public Health.**

   *Strength:* Once the data is in a single location it is easy to process.

   *Weakness:* Gathering the data in a single location requires extensive export activity from feeder systems. Semantic interoperability of the data may be poor unless all the feeder systems use common standards for vocabulary and content. Data currency may be poor if the central repository is out of date. Updating from the central repository to the feeder systems may not be supported. If developed by the provider locally, may be expensive to maintain and enhance over time.

2. **Use a centralized data warehouse provided by an HIE. This is similar to the first model, but uses an HIE as the outsourcing provider for the warehouse. Example: “CRISPing a patient”, a term used to refer to viewing all the data available on a patient from CRISP, a major HIE.**

   *Strength:* Once the data is in the HIE (a single outsourced location) it is easy to process.

   *Weakness:* Gathering the data in a single location requires extensive export activity from feeder systems. Semantic interoperability of the data may be poor unless all the feeder systems use common standards for vocabulary and content. Data currency may be poor if the central repository is out of date. Updating from the central repository to the feeder systems may not be supported. Data that is available will probably be limited to a restricted set of standard record types supported by the HIE.

3. **Use an EHR as the primary data repository, analyze the data using traditional report generation methods native to the EHR. Example: This will be the most common case for a hospital.**

   *Strength:* Once the data is in the EHR (a single location) it is easy to process. Data quality for clinical medical issues should be excellent, updated in real time.
Weakness: Unless the EHR is only handling data from a single source, gathering the data in a single location requires extensive export activity from feeder systems. Semantic interoperability of the data may be poor unless all the feeder systems use common standards for vocabulary and content. Data currency may be poor if the central repository is out of date. Updating from the central repository to the feeder systems may not be supported. Data that is available will probably be limited to clinical medical data and a restricted set of standard record types supported by a connected HIE. Report generation tools may be limited or expensive to modify unless the system support good native reports by default.

4. **Use an EHR as the primary data source, analyze the data by exporting it to an external data analytic system (such as Tableau).** *Example: This will be a common case for a hospital that has an advanced data analytic group, using an external service in addition to whatever native reporting the EHR supports.*

   **Strength:** Once the data is in the EHR (a single location) it is easy to process. Data quality for clinical medical issues should be excellent, updated in real time. Export of data to an external data analytic system should give additional analytic and display capability.

   **Weakness:** Unless the EHR is only handling data from a single source, gathering the data in a single location requires extensive export activity from feeder systems. Semantic interoperability of the data may be poor unless all the feeder systems use common standards for vocabulary and content. Data currency may be poor if the central repository is out of date. Updating from the central repository to the feeder systems may not be supported. Data that is available will probably be limited to clinical medical data and a restricted set of standard record types supported by a connected HIE. Export to the external system may result in a lag in reporting currency unless the bridge is near real time. Costs will be higher due to the need for an additional service contract.

5. **Use a cloud-based care management system that is optimized for care coordination.** General reports either using native reporting tools or by exporting to an external data analytic system (such as Tableau). *Examples: Eccovia Solutions, Penelope.*

   **Strength:** Once the data is in the cloud (a single virtual location) it is easy to process. Data quality should be excellent, updated in real time. Export of data to an external data analytic system can give additional analytic and display capability. Multiple types of providers can access the system. Systems that have been designed specifically for care coordination typically have strong care plan features. Costs may be lower than investment in traditional EHR systems, and will not require capital investment, typically using fee-based models such as cost per seat or managed population size pricing levels. Potentially faster ramp-up implementation schedules.

   **Weakness:** May have issues with connecting to EHR systems and HIEs.
6. **Use a cloud-based care management system that is based on a Customer Relationship Management (CRM) system. Generate reports using an external data analytic system.**

*Examples: Tribridge, Salesforce Health.*

*Strength:* Once the data is in the cloud (a single virtual location) it is easy to process. Data quality should be excellent, updated in real time. Export of data to an external data analytic system can give additional analytic and display capability. Multiple types of providers can access the system. Systems that have been designed specifically for care coordination typically have strong care plan features. Costs may be lower than investment in traditional EHR systems, and will not require capital investment, typically using fee-based models such as cost per seat or managed population size pricing levels. Potentially faster ramp-up implementation schedules. CRM systems typically have strong workflow management features.

*Weakness:* May have issues with connecting to EHR systems and HIEs. CRM systems are typically designed for sales management applications, not healthcare, and may lack sophistication in some aspects of healthcare management.
Architectural Model Charts

The following architectural model charts summarize key classification variables for care plan data aggregation systems.

SINGLE DATA SOURCE VERSUS MULTIPLE DATA SOURCES

Data Source

Single Source

- EHR (Data will be strong for medical content, probably weak for other domains)
- non-EHR system (Data may cover multiple domains, but may lack medical detail)

Multiple Sources

For unified view, create a centralized repository

- Use proprietary system development methods (e.g. SFDPH)
- Use a care coordination platform with APIs

For unified view, use HIE as repository (e.g. "CRISPing a patient")
DATA STRUCTURE AND STORAGE METHODOLOGY

Data Structure and Storage

- Standards-compliant
  - Relational Database
    - Structured data
    - The ideal case for data analytics
  - Non-Relational storage method
    - Flat files with structured record formats
    - Excel files and other exchange formats (e.g., CSV)
    - Unstructured data (e.g., free text, PDF files, images, "blobs")

- Not Standards-compliant
  - Relational Database
    - Structured data
  - Non-Relational storage method
    - Flat files with structured record formats
    - Excel files and other exchange formats (e.g., CSV)
    - Unstructured data (e.g., free text, PDF files, images, "blobs")

Relational Database

Structured data

Unstructured data (e.g., free text, PDF files, images, "blobs")

Flat files with structured record formats

Excel files and other exchange formats (e.g., CSV)
IMPLEMENTATION STRATEGY

Implementation Strategy

Cloud-Based (Hosted)
- Vendor-Supplied EHR
  - Native data analytics
  - Export for data analytics
- Care Coordination Suite
  - Native data analytics
  - Export for data analytics

Premise-Based
- Vendor-Supplied EHR
  - Native data analytics
  - Export for data analytics
- Care Coordination Suite
  - Native data analytics
  - Export for data analytics
- Home-grown custom solution
  - Native data analytics
  - Export for data analytics
Workflow Integration

Single Provider

- System has intelligent workflow features to affect point of service (now standard with large EHR and care coordination platforms)
- Data can be collected at point of service but does not influence service

Multiple Providers

- System can provide a consolidated view of the client and has intelligent workflow features to affect point of service (Requires a Shared Platform)
- System can provide summary view of other provider relationships but not directly affect them (e.g., SFDPH)
- System supports messaging protocols between providers but does not provide a summary shared view of the client
**Internet of Things (IoT) Example:** A 69-year old man with glaucoma is placed on a trial of treatment of an ophthalmic eyedrop in the beta-blocker class to reduce intra-ocular pressure. Since ophthalmic beta-blockers can be absorbed systemically, a Fitbit wrist tracker continuously monitors the patient’s heart rate during a two-week trial of treatment. Data show a measurable reduction in heart rate during the trial period, which correlates with the patient’s subjective reports of discomfort during exertion. Patient is switched to a different class of glaucoma drug as a result.