Why Population Health Management Strategies Need Both Clinical and Claims Data

Both Clinical and Claims Data Are Required for Meaningful Population Health Management

Health systems will never achieve population health management if they continue to use only claims data or clinical data to analyze patient populations. Neither type of data provides the necessary depth of information when used alone. The solution is to combine claims and clinical data. While this process won’t be easy, it’s necessary if providers and payers want to achieve the three goals of population health management: improved outcomes, increased patient safety, and decreased costs.

Why the Healthcare Industry Is Suddenly Interested in Population Health Management

The concept of population health management has become intriguing for health systems and payers for two critical reasons. The first reason is to find a way to manage the escalating costs of treating chronic diseases in the United States. According to the Centers for Disease Control and Prevention (CDC),...
chronic diseases are the leading cause of death and disability in the U.S.¹

Consider these facts:

- Seven out of 10 deaths among Americans each year are from chronic diseases.²
- Heart disease, cancer, and stroke account for more than 50 percent of all deaths each year.³
- Arthritis is the most common cause of disability, with nearly 19 million Americans reporting activity limitations.⁴
- Diabetes continues to be the leading cause of kidney failure, non-traumatic lower-extremity amputations, and blindness among adults aged 20-74.⁵

The second reason health systems and payers are interested in population health management is because of the shift in government payment models from fee-for-service to pay-for-value. Led by the Centers for Medicare and Medicaid Services (CMS), the pay-for-value model offers financial incentives for disease prevention and management but penalizes poor outcomes. As a result, health systems are moving patients away from high-cost, acute settings to the most appropriate, lowest-cost settings — and understandably so.

CMS’s mandatory shift in reimbursement models for health systems is necessary because the fee-for-service model is unsustainable. In contrast, the pay-for-value model will significantly reduce the government’s spending on healthcare and curtail the growing budget deficits.

According to a report on the Clinically Appropriate & Cost-Effective Placement Project (CACEP) from the Alliance for Home Health Care Quality and Innovation, Medicare could reduce its spending over a 10-year period by $34.7 billion if patient care settings were shifted from facility-based care to home and community-based care. The report estimates an additional $100 billion could be saved if care delivery were restructured to be less wasteful and more effective. These additional reductions could be accomplished by using bundled payments and by adding a policy to reduce post-acute care spending (excluding the index hospitalization) by 7.5 percent.⁶

Providers and payers have powerful incentives to change with the new payment model. But to survive, they will need to access and analyze vast amounts of population data from many different outpatient settings. In particular, the following groups and programs need accurate data about the outcomes for various patient populations: Patient-Centered Medical Home
(PCMH), Accountable Care Organizations (ACOs), and CMS’s Bundled Payments for Care Improvement initiative (BPCI).

**STRENGTHS — AND LIMITATIONS — OF CLAIMS DATA**

Claims data is administrative data and includes information about patient demographics, billable charges, dates of service, diagnosis codes, procedure codes, insurance, and providers. This type of data is created after every patient encounter with a provider and is subsequently submitted to a payer (the government or an insurance company) for reimbursement.

Historically, payers and health systems have exclusively relied on claims data for analysis associated with population health management because of the following attributes:

- **Claims data is readily available.** The requirements for payment to the provider ensure the data is entered into the claims system on a complete and timely basis.

- **Claims data spans a patient’s full continuum of care.** The payer has a record of every encounter and every prescription filled unless a patient pays for services out of his or her own pocket.

![Figure 1 – CMS 1450 requires claims data be entered into highly structured fields, making it easy to be used by other systems.](image-url)
Claims data is highly structured. Almost all the data must be captured in specific fields on standardized forms in order for claims to be approved and payment to be issued. Because of this consistent format, it’s relatively easy for other systems to consume.

By examining claims data, analysts gain information about the cost and utilization of a patient population across multiple care settings as well as information about the types of diagnoses and procedures performed. In specific, claims data provides insight into the following performance measures:

- Mortality rates
- Complications
- Access to appropriate health services
- Length of stay (duration of treatment) for appropriate settings
- Charges for care provided

However, despite having some value, claims data has limited use for quality and cost improvements for the reasons indicated below.

Claims data lacks important clinical detail. Claims data is designed to capture only as much detail as is required to determine payment. This type of data does not capture all of the clinical details of the patient care process. Although a collection of claims can include information about chronic conditions or history, a single claim does not unless it is relevant to the specific treatment or procedure. As a result, data critical to an accurate analysis may be missing from a particular claim. Likewise, a claim may not include all diagnoses and typically lacks critical data, such as lab results and medications.

Claims data is highly retrospective. There is a significant lag between the date of care or service and the date the claims data is available for analysis which is most often measured in months. Part of this delay is due to the need to verify the accuracy of the coding that is the basis for establishing payment. With the speed at which new discoveries in population health management (and healthcare in general) are made, this type of lag time makes it difficult to deliver accurate, timely analysis.

Claims data does not provide insight into the actual process of care. The key to improving the quality of care is to understand the dynamics of the processes that are involved in the actual delivery of care. Since claims data is a static summary of the diagnoses,
procedures, and costs that are the result of these care processes, efforts to identify the root cause of any variations in care between patients is impossible.

STRENGTHS — AND LIMITATIONS — OF CLINICAL DATA

In addition to claims data, health systems have access to another source of information about their patient populations: clinical data.

Unlike claims data, clinical data provides critical detail and insights into the processes used to deliver patient care. This type of data answers the questions: what is effective, why is it effective, and how can providers leverage the lessons learned to improve care for other patients. Some typical categories of clinical data include patient demographics (age, sex), medical history (vital signs), allergies, immunizations, problem list, diagnoses, procedures, lab tests, medications, and providers (with specialty).

When the care delivery team combines clinical data with patient satisfaction data, they gain additional knowledge about how patients perceive the care processes, which is critical to providing a quality experience. In addition, clinical data shows how any variations in these processes will impact the cost, quality, or outcome of the patient’s encounter. Clinical data is also timely, since it’s collected in real time at every patient visit and allows modifications in the care delivery processes as they occur.

However, despite the strengths of clinical data, its use for population health management is limited because of the reasons below:

- **Clinical data doesn’t span the continuum of care.** Clinical data provides great detail about a patient’s medical profile and treatment patterns during a specific encounter. However, the details are generally limited to the acute care, ambulatory setting or physician’s office and are difficult to aggregate into a cohesive view.

- **Clinical data is siloed.** The data in one facility is generally siloed from the others, another major drawback of clinical data. If a patient goes to several places for care, such as a specialist’s office, a pharmacy, or a rehab center, the disparate systems usually don’t share information across the care continuum.

- **Clinical data is unstructured.** Rather than being input into pre-defined fields in a standard form the way a claims system is, 80 percent of clinical data is entered as unstructured text notes, making extraction and normalization of the data difficult.
CHALLENGES THAT MUST BE OVERCOME BEFORE ACHIEVING THE
BEST SOLUTION: THE COMBINATION OF CLINICAL AND CLAIMS DATA

While neither clinical data nor claims data provide enough information on their
own for population health management, when combined, the two types of data
provide the best of both worlds: up-to-date clinical detail along with a complete
view of a patient’s care throughout their entire patient care experience. With this comprehensive picture, healthcare
providers and payers can manage their populations more
effectively and realize the Triple Aim of higher quality, lower
costs, and a better patient experience. Yet, to achieve this
ideal solution, there are many challenges providers and
payers must first overcome.

Cultural Challenges

The greatest challenge to data sharing has primarily been
cultural because data has been used as a pawn in contract
negotiations between payers and providers. And while
they may talk about sharing data, there’s still significant resistance to actually
sharing the data. For information sharing to work, payers and providers will
have to learn to trust each other, and that’s a huge hurdle to overcome.

Patient Cohort Definition Challenges

Most organizations use groupings of administrative (claims) codes, such as
CPT, ICD-9, and ICD-10 to define cohorts. This approach is a valid starting
point, but such macro level groupings often miss patients who should have
been included. Adding clinical data, such as lab results and imaging studies, is
necessary to refine the cohort and ensure the full eligible population is included.

Even though it will be difficult to group specific patient
populations into disease cohorts, it’s a foundational step
and will give providers and payers the ability to answer
questions such as the following:

- What is this population’s utilization of healthcare
  services across the continuum?
- How much do the population and individual members
cost the organization?
- What is the risk associated with this population and
  specific members?
- What are the rules for accurately attributing
  members to physicians?
Registry Challenges for Patients with Multiple Conditions

Grouping patients is complicated when they suffer from multiple chronic diseases, such as diabetes, heart failure, or asthma — all major focuses of population health management. One reason for this difficulty is the complexity of identifying patients who may not present with a primary complaint related to these chronic conditions.

Figure 2 — Data from payer systems (claims) and provider systems (EMRs, etc.) combine to form Source Marts in a data warehouse, creating the technical platform necessary for collaboration between payers and providers.
For example, suppose the primary reason a patient is admitted to the hospital is for a total knee replacement procedure. Yet the patient also happens to suffer from heart failure and diabetes. The patient is assigned a primary diagnosis related to the knee surgery for billing purposes, but the physicians and coding staff may omit other diagnoses that aren’t relevant from a billing perspective.

Now, based on this patient’s primary diagnosis code, he appears in hospital data as an orthopedic patient. This presents a problem when the organization is defining its heart failure or diabetes population because the patient may be excluded from the other cohorts. Only clinical data offers the insight necessary to identify this patient as one who belongs in the heart failure or diabetes registry. This task is best accomplished through the establishment of an enterprise data warehouse (EDW) and the use of advanced applications for analyzing data.

Identifying patients who belong in multiple registries becomes even more difficult when attempting to create those disease registries across the full continuum of care. Yet, despite the many challenges, finding a solution is necessary in order to manage population health given the tectonic shift occurring in today’s healthcare industry and payment models.

**Billing Code Variation Challenges**

Currently, care settings have many variations with their billing codes and billing methodologies. These variations and the lack of clinical specificity attached to the codes makes it difficult to define specific disease management cohorts.

To illustrate the issues, Medicare has different requirements for how skilled nursing facilities (SNFs), home health agencies, dialysis centers, and other ambulatory care venues can bill them. For example, a skilled nursing facility may bundle all of their services into a single bill or bill for certain services separately depending upon the patient’s Medicare status, but a dialysis center is limited to submitting a single charge per visit. To add further variation, hospital outpatient services bill according to ambulatory payment classifications (APCs). Professional claims must also be included in any comprehensive analysis of a population of patients across the continuum of care.

As claims data from across the outpatient continuum builds up — all with differing requirements for how Medicare will reimburse and many lacking clinical detail — using this data to define and then manage a disease cohort becomes a very complex prospect.
Additional Challenges

In addition to the previously mentioned difficulties that must be overcome, the following challenges also limit the sharing and merging of claims and clinical data between payers and providers:

- Both types of data — but particularly claims data — tend to be “dirty,” i.e., inaccurate, incomplete, or erroneous, requiring the data be scrubbed before using them for analytics.

- Matching patient identifiers from each record to create a single, accurate, comprehensive patient record is a complex and difficult undertaking. The more diverse the sources of data, the more challenging it becomes to bring them under a single patient identifier.

- Payers have to choose which claims data is the most accurate and which version of any particular claim to use. Each claim has multiple iterations, including (at the very least) the original claim the provider submits and the edited, adjudicated claim. While it may be easier for a provider to access its own claims data, there is a risk it may not reflect the final form the data takes.

![Figure 3 — Health Catalyst’s Source Mart Designer analyzes the structure of data sources and recommends pragmatic data types and column names with well-structured, consistent data fields.](image)

<table>
<thead>
<tr>
<th>Source Schema</th>
<th>Source Table</th>
<th>Source Type</th>
<th>PK</th>
<th>Null</th>
<th>Source Column</th>
<th>Column Alias</th>
</tr>
</thead>
<tbody>
<tr>
<td>CatalystAdmin</td>
<td>TableBASE</td>
<td>integer</td>
<td></td>
<td></td>
<td>TableID</td>
<td>TableID</td>
</tr>
<tr>
<td>CatalystAdmin</td>
<td>TableBASE</td>
<td>varchar(255)</td>
<td></td>
<td></td>
<td>DatabaseNM</td>
<td>DatabaseNM</td>
</tr>
<tr>
<td>CatalystAdmin</td>
<td>TableBASE</td>
<td>varchar(255)</td>
<td></td>
<td></td>
<td>SchemaNM</td>
<td>SchemaNM</td>
</tr>
<tr>
<td>CatalystAdmin</td>
<td>TableBASE</td>
<td>varchar(255)</td>
<td></td>
<td></td>
<td>TableNM</td>
<td>TableNM</td>
</tr>
<tr>
<td>CatalystAdmin</td>
<td>TableBASE</td>
<td>varchar(255)</td>
<td></td>
<td></td>
<td>ViewNM</td>
<td>ViewNM</td>
</tr>
<tr>
<td>CatalystAdmin</td>
<td>TableBASE</td>
<td>varchar(255)</td>
<td></td>
<td></td>
<td>DataNumFieldNM</td>
<td>DataNumFieldNM</td>
</tr>
<tr>
<td>CatalystAdmin</td>
<td>TableBASE</td>
<td>dbtimestamp</td>
<td></td>
<td></td>
<td>DataNumStartDT</td>
<td>DataNumStartDT</td>
</tr>
<tr>
<td>CatalystAdmin</td>
<td>TableBASE</td>
<td>dbtimestamp</td>
<td></td>
<td></td>
<td>DataNumEndDT</td>
<td>DataNumEndDT</td>
</tr>
<tr>
<td>CatalystAdmin</td>
<td>TableBASE</td>
<td>numeric(38, 0)</td>
<td></td>
<td></td>
<td>RowCNT</td>
<td>RowCNT</td>
</tr>
<tr>
<td>CatalystAdmin</td>
<td>TableBASE</td>
<td>dbtimestamp</td>
<td></td>
<td></td>
<td>LastSuccessfullLoadDT</td>
<td>LastSuccessfullLoadDT</td>
</tr>
<tr>
<td>CatalystAdmin</td>
<td>TableBASE</td>
<td>dbtimestamp</td>
<td></td>
<td></td>
<td>MaxIncrementalDTS</td>
<td>MaxIncrementalDTS</td>
</tr>
<tr>
<td>CatalystAdmin</td>
<td>TableBASE</td>
<td>integer</td>
<td></td>
<td></td>
<td>StartIncrementalOverTypeMinutesNBR</td>
<td>StartIncrementalOverTypeMinutesNBR</td>
</tr>
<tr>
<td>CatalystAdmin</td>
<td>TableBASE</td>
<td>integer</td>
<td></td>
<td></td>
<td>SourceConnectionID</td>
<td>SourceConnectionID</td>
</tr>
<tr>
<td>CatalystAdmin</td>
<td>TableBASE</td>
<td>varchar(255)</td>
<td></td>
<td></td>
<td>SourceTableNM</td>
<td>SourceTableNM</td>
</tr>
<tr>
<td>CatalystAdmin</td>
<td>TableBASE</td>
<td>varchar(255)</td>
<td></td>
<td></td>
<td>TableNameNM</td>
<td>TableNameNM</td>
</tr>
<tr>
<td>CatalystAdmin</td>
<td>TableBASE</td>
<td>longvarchar</td>
<td></td>
<td></td>
<td>DescriptionTXT</td>
<td>DescriptionTXT</td>
</tr>
<tr>
<td>CatalystAdmin</td>
<td>TableBASE</td>
<td>longvarchar</td>
<td></td>
<td></td>
<td>SourceQueueTXT</td>
<td>SourceQueueTXT</td>
</tr>
</tbody>
</table>
Patient privacy legal requirements require the patient’s permission to share the data between the provider and the payer. Some patients may have reservations about giving their payers access their clinical data, or they may have concerns about the security of their personal health information.

Providers have different EMRs and relationships with multiple payers with specific data requirements, which makes abstracting and sharing data problematic.

2 WAYS TO OVERCOME THE CHALLENGES TO BE ABLE TO USE BOTH CLINICAL AND CLAIMS DATA FOR POPULATION HEALTH MANAGEMENT

The task to create a systematized, consistent method to match codes with cohorts is complex, but not insurmountable. The following two strategies will help providers and payers address these complexities to make it easier to combine claims and clinical data and create a superior population health management initiative:

- **Map outpatient codes to clinical care process families.** Analytics vendors and other organizations can undertake a significant effort to map the various outpatient codes to clinical care process families (such as diabetes and heart failure). For example, the vendor would look at skilled nursing facilities and analyze the billing codes they commonly use. They would then map different groupings of these codes to the care process family. This type of sophisticated mapping would provide a consistent methodology for assigning patients to a disease cohort regardless of care setting. Health Catalyst is an example of a vendor that’s in the process of developing these types of mapping efforts.

![Health Catalyst’s Population Health Management Mapping Project](image)
Select flexible and scalable technology. ACOs and other organizations must select a technology infrastructure to manage their populations that is flexible enough to integrate and map new data sources quickly, on an as-needed basis. Because population health management and accountable care are in their infancy, the types of data sources required and use cases for the data will often change. Many analytics solutions targeted at the population health market today lack this flexibility. Organizations should look for technology that is adaptable enough to handle rapidly evolving needs. They should also ensure it is sophisticated enough to effectively leverage both clinical and claims data to accurately define and manage disease cohorts.

Consider this example of the need for flexible technology. When a diabetic patient visits a doctor in the outpatient setting for a foot ulcer, the doctor may or may not document the diagnosis of “diabetes.” Instead, the diagnosis may simply be “foot ulcer.” By using a flexible analytics solution, however, analysts can easily categorize this patient in the diabetic population cohort even if the patient is seen for the primary complaint of a foot ulcer.

Correcting these issues may sound simple, but without a flexible solution, it isn’t. The coming years promise more changes to the healthcare industry (new research, new technologies, etc.), not fewer, so the more flexibility built into the technology, the better a job the organization can do to create population health management solutions that deliver the expected results.

CONCLUSION

Achieving population health management is a critical goal for the healthcare industry. But the traditional way of using either claims data by itself or clinical data by itself won’t yield the valuable insights necessary to understand the data. Instead, payers and providers need to work together to overcome the challenges to combine claims and clinical data. Once the challenges are overcome, a more complete view of patients from across all care settings will allow for more accurate definitions of disease cohorts, identification of variations in care and waste, and a better measurement of the effectiveness of treatment. When coupled with the flexibility to adjust to changes in our rapidly growing understanding of what population health actually entails, data aggregation and analysis become a powerful tool for managing population health.
REFERENCES


Bobbi Brown, Vice President, Financial Engagement

Bobbi Brown started her healthcare career at Intermountain Healthcare before moving to Sutter Health and, later, Kaiser Permanente, where she served as Vice President of Financial Planning and Performance. Ms. Brown also acted as team lead on the installation of planning software at Ascension as Vice President of Financial Engagement. Ms. Brown regularly writes and teaches on topics including value-based purchasing, payer-provider collaboration, revenue cycle management, and Medicare reimbursements.

Luke Skelley, Vice President, Sales

Luke Skelley joined Health Catalyst in May 2013. His clinical nursing background was in critical care and he managed organ and tissue donation programs for 13 years. Mr. Skelley’s background includes orthopedic products, severity adjusted outcomes software for Medstat Group, clinical pathway development software for CareScience, and prospective payment methodologies for Optum Insight. He was recruited by Navinet for a new line of clinical products for physician offices before they were acquired by Lumeris.
ABOUT HEALTH CATALYST

Health Catalyst is a mission-driven data warehousing and analytics company that helps healthcare organizations of all sizes perform the clinical, financial, and operational reporting and analysis needed for population health and accountable care. Our proven enterprise data warehouse (EDW) and analytics platform helps improve quality, add efficiency and lower costs in support of more than 30 million patients for organizations ranging from the largest U.S. health system to forward-thinking physician practices. Faster and more agile than data warehouses from other industries, the Health Catalyst Late-Binding™ EDW has been heralded by KLAS as a “newer and more effective way to approach EDW.”

For more information, visit www.healthcatalyst.com, and follow us on Twitter, LinkedIn, Google+ and Facebook.