Clarify

# Claims Data Enrichment:

A Game-Changer for Healthcare Analytics

# Table of contents

# Introduction

Turning claims data into actionable insights presents a variety of challenges. Often, teams spend too much time trying to answer simple business questions, too much money processing data, and too much effort cleaning and ensuring quality. Thankfully, there are ways to get from data to insights faster, more cheaply, and with greater precision by partnering with the right vendor.

Best-in-class claims processing techniques utilize machine learning and other algorithms to create the best possible "single source of truth" to inform business decisions.

At Clarify, we designed a unique six-step data enrichment process to make claims data more complete, more robust, and, ultimately, more useful.

Through the process of enrichment, Clarify combines multiple claims sources into a single standardized source, eliminates duplicates, and enhances raw claims with additional contextual information. By adding this context, which includes medical and diagnostic codes, patient demographics, and physician reference data, organizations can access a more comprehensive and detailed view of patient care. An enriched claims database unlocks enhanced analytical capabilities, including predictive modeling, risk stratification, and population health management.

In this paper, we will describe Clarify's data enrichment process and provide examples of how enriched claims can be utilized to deliver timely insights and drive immediate return on investment for your organization.

# Costly barriers to harnessing the power of claims data

Healthcare organizations frequently face the choice of either outsourcing the process of transforming raw data into useful insights or bringing that work in-house. As you consider your options, it is important to understand the challenges and costs involved. Here are some important barriers that your organization would avoid by partnering with Clarify Health for data refinement:

INEFFICIENT WORKFLOWS
Handling a large volume and variety of claims necessitates data integration from various sources, insurance databases, and external vendors. Even internally, healthcare organizations frequently rely on multiple software platforms with distinct data formats, coding standards, and terminologies.

HIGH COST OF ACQUIRING EXPERTISE
Claims enrichment requires specialized knowledge and expertise in medical coding, billing regulations, and reimbursement methods. For example, creating a comprehensive claims record requires an in-depth understanding of healthcare coding systems like ICD- 10 and CPT to navigate intricate data transformations. However, acquiring and retaining such skilled professionals can be challenging and costly, including training, certifications, and ongoing education to stay abreast of industry standards. These costs can range from several hundred thousand to several million dollars per year, depending on the size of your organization.

INVESTMENT IN INFRASTRUCTURE
Establishing and maintaining the infrastructure required for in-house claim enrichment entails substantial financial investments. This includes hefty upfront costs associated with multiple data licenses, infrastructure set-up, and data engineering that could total as much as $10 million. Furthermore, organizations need to allocate at least several hundred thousand dollars for maintenance annually.

INADEQUATE QUALITY ASSURANCE AND RISK MITIGATION
Ensuring the accuracy, completeness, and integrity of enriched claims data is pivotal for analytics and compliance. In-house teams must implement stringent quality control measures, conduct regular audits, and employ risk mitigation strategies to minimize errors, discrepancies, and potential financial losses. These activities demand additional resources, time, and expertise, further adding to the complexity and cost of in-house claims enrichment.

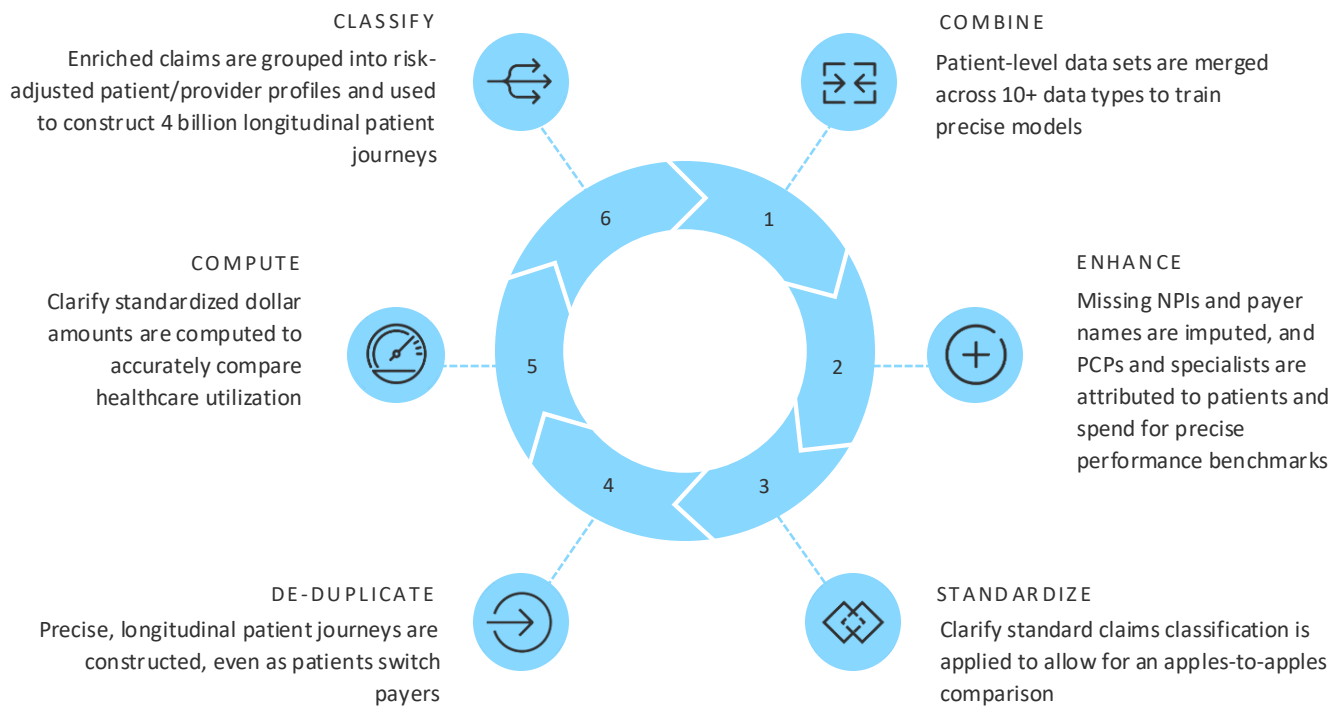Lack of holistic visibility into patient needs
Even with enriched claims, a comprehensive understanding of patient demographics, medical histories, and treatment patterns may still be lacking. To address these gaps, it is crucial to link enriched healthcare claims with patient-level data, such as social and behavioral determinants of health (SBDoH) and provider reference data. This linkage helps establish a more holistic picture of the patient, enabling organizations to gain deeper insights and make more informed decisions.

# Clarify's approach to data cleaning and enrichment

The Clarify Atlas Platform® houses the healthcare industry's largest collection of patient journeys, delivering the most precise and actionable insights to payers, providers, and life sciences companies. It maps over four billion patient journeys across over 300M lives, leverages best-in-class grouper technology, and drives 20B+ AI-powered predictions, answering healthcare's most complex business questions in an instant. Leveraging patient journeys to enrich price transparency data provides a cleaner data set for unparalleled access to rate intelligence that is truly meaningful.

## 6-step process to turn raw data into insights

**CLASSIFY**
Enriched claims are grouped into risk-adjusted patient/provider profiles and used to construct 4 billion longitudinal patient journeys

**COMBINE**
Patient-level data sets are merged across 10+ data types to train precise models

**COMPUTE**
Clarify standardized dollar amounts are computed to accurately compare healthcare utilization

**ENHANCE**
Missing NPIs and payer names are imputed, and PCPs and specialists are attributed to patients and spend for precise performance benchmarks

**DE-DUPLICATE**
Precise, longitudinal patient journeys are constructed, even as patients switch payers

**STANDARDIZE**
Clarify standard claims classification is applied to allow for an apples-to-apples comparison

Clarify's platform brings together data across more than 300 million unique individuals. However, what sets the platform apart is its highly efficient, patented, and automated process for data enrichment, as shown above. Clarify's enrichment process follows these steps:

### 1. COMBINE

Claims data is derived from over ten sources, including direct sourcing from payers, clearinghouses, and other firms aggregating claims data. To ensure a comprehensive patient journey, we employ scoring algorithms that compare data across sources, selecting the highest quality version of each claim for inclusion in the patient's timeline of events.

## 2. ENHANCE

The process incorporates several steps to enhance the value of the data. We create a secure Clarify Patient ID to connect data across all assets, allowing for linkage across sources. Inpatient claims, whether commercial or Medicare, are assigned an MS-DRG for comparative purposes. We identify and impute missing provider and patient values, such as operating physician or location and zip code information, by analyzing data across the entire patient journey. Missing or incorrect dates are adjusted or imputed based on claims history. We leverage our large data lake and complex logic to impute missing fields, leading to more accurate and robust data. Additionally, our clinical informatics team defines service classifications for all claim types, ensuring clinically appropriate utilization and analysis. Our algorithms also associate patients with primary care physicians and appropriate specialists based on claims analysis.

## 3. STANDARDIZE

To facilitate comparisons across commercial and Medicare populations, we have developed a standardized claims data model and standardized field values. This standardized approach allows for effective comparisons of post-acute, inpatient, outpatient, professional, and other claims across different data sources and payers.

## 4. DE-DUPLICATE

We invest significant time and effort into analyzing each patient's journey and eliminating duplicate services. This involves breaking down claims into line-level details to eliminate duplicate events and examining multi-day stays to avoid double counting. Our data includes linked claims data on over 300 million lives annually, integrated prescription data on 270 million lives annually, and is enriched with social and behavioral data, as well as published price transparency data from payers and providers.

## 5. COMPUTE

We have developed a proprietary, standardized claim payment amount that eliminates differences in payments caused by payer variations or geographical disparities. Similar to CMS's methodology, our Clarify Standardized Dollar amounts account for all medical services performed, making them suitable for all patient populations. These standardized amounts serve as the foundation for our risk-adjusted benchmarks, allowing us to compare resource utilization across providers nationwide.

## 6. CLASSIFY

We organize longitudinal patient data into clinically relevant standard units of analysis called Clarify Care Groupings (CCGs). Our configurable engine is at the heart of our business, allowing for flexible aggregation of claims. With our engine, you can analyze patient journeys, including post-surgical experiences, and perform time period analysis to support metrics such as PMPY/PMPM and care utilization. Referral pattern analysis is also supported, enabling you to identify connections between different events. Furthermore, we build risk-adjusted benchmarks, enabling fair comparisons of provider performance across similar patient populations.

# Enriched claims data in action

By starting with enriched claims data, healthcare organizations can access insights with speed, precision, and less resource investment. Below are examples of how health plans, healthcare providers, and pharmaceutical companies leverage enriched claims analytics to enhance their operations, drive value, and improve outcomes for their patients.

### PATIENT JOURNEY INSIGHTS

Identify critical junctures in a patient's journey, evaluate care quality, assess referral patterns, and optimize resource allocation for providers. These enriched claims analyses lead to the identification of care gaps, assessment of treatment efficacy, and enhanced member satisfaction. Claims can also help organizations gain a deeper understanding of patient needs and preferences, enabling targeted interventions and more effective commercialization efforts.

### TELEHEALTH PROGRAM IMPACT

Evaluate the impact of virtual care on health outcomes, patient satisfaction, resource utilization, and cost-effectiveness. Through longitudinal analysis, stakeholders can uncover trends and associations over time, gaining insight into the benefits and limitations of virtual care. This analysis ensures fair comparisons between virtual care and traditional care models.

### SHARE OF WALLET INTELLIGENCE

Gain valuable insight into consumers' healthcare service usage and preferences. This knowledge enables organizations to identify targeted marketing opportunities and enhance service offerings, ultimately increasing their share of healthcare spend.

# Enriched claims data in action

**PROCEDURE PERFORMANCE AND QUALITY**

Enable predictive modeling and risk stratification for procedure outcomes. By developing predictive models to assess the likelihood of successful outcomes for specific procedures, stakeholders can identify patients at higher risk of complications or poor outcomes. This knowledge facilitates personalized interventions, pre-procedure optimization, and informed decision-making to improve overall patient care.

**OPTIMIZATION OF PATIENT SERVICES**

Uncover utilization patterns, revealing areas of overutilization or underutilization. These insights lead to informed decisions about resource allocation, improved access to care, and reduced waiting times for patients. With the right insights, organizations can enhance care coordination and management by identifying high-risk patients who require proactive interventions and personalized care plans

**ENHANCEMENT OF CARE MANAGEMENT PROGRAMS**

Empower analysts to evaluate the effectiveness of treatment programs and measure their impact on overall healthcare costs. By combining enriched claims with other sources like electronic health records and patient-reported outcomes, stakeholders can make evidence-based decisions, shape care guidelines, and optimize care management initiatives.

# Large health system leverages enriched claims to improve referral analytics, leading to $14M in revenue growth in one year

■ BACKGROUND

A large not-for-profit health system was looking to increase in-network volume and optimize physician referral patterns but lacked insights across all settings of care—PCPs, specialists, and ancillary care. By identifying high-volume specialists for engagement and acquisition and having more data-driven conversations with physicians, they increased referrals from employed and independent physicians.

■ SOLUTION

The team needed more visibility into their outpatient flow and downstream referral activity to understand how their network penetration varied across the care continuum.

By leveraging linked, cleansed, and enriched data from Clarify, the health system gained more visibility into the needs of patients in their market and referral patterns across physicians, and greater confidence in the accuracy of the insights.

> **$14M revenue growth in one year**

■ RESULTS & HIGHLIGHTS

By identifying high-volume specialists for engagement and acquisition and having more data-driven conversations with physicians, the health system increased referrals from employed and independent physicians across all service lines by 5% in one year, generating $14M in revenue.

| ↑↑↑ | 5% | 2% |
|---|---|---|
| **PHYSICIAN ALIGNMENT** | **VOLUME GROWTH (OVERALL)** | **VOLUME GROWTH (PCPs)** |
| The business development team was able to have detailed, data-driven conversations with physicians that led to greater alignment and positive behavior change. | Among employed and independent physicians, and across all service lines, the system increased referral volume retention by 5% from Q1 2021 to Q1 2022. | Among employed and independent PCPs, the system increased net new referral volume by 2%, resulting in $3.5M of new revenue. |

# Conclusion

Enriched claims analytics have a wide range of applications in healthcare, including share of wallet analyses, disease progression analysis, and procedure performance and quality benchmarking. With domain expertise, advanced analytics techniques, and comprehensive data integration capabilities, Clarify empowers healthcare organizations to unlock the full potential of claims data, generating significant ROI for customers on their data and analytics investments.

# Clarify

# About Clarify

Clarify Health Solutions® is a healthcare data and analytics company trusted by some of the most established organizations in healthcare, including providers, payers, tech and services, and life sciences. The Clarify Atlas Platform® is the foundation, leveraging the industry's largest and most robust dataset to map over 300M+ lives to deliver 20B+ AI-powered predictions to surface actionable insights with unparalleled speed and precision.

To learn more about how Clarify's data and analytics solutions can help your organization answer critical business questions about provider performance, visit clarifyhealth.com.