KYTHERA
DECIPHERING HEALTHCARE

Clinical Notes

Unlocking the Rich Clinical Narratives Behind Patient Journeys with Unstructured Clinical Notes

Kythera Labs develops data technologies to help life science teams uncover deeper clinical context - enabling faster, more confident insights and supporting the construction of complete, longitudinal patient journeys across complex healthcare data domains.

Kythera's Clinical Notes Data at a Glance

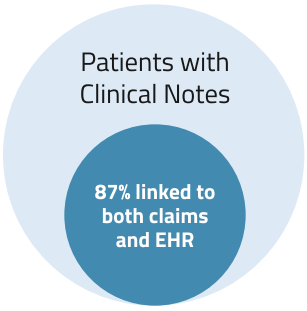
3,317,891,435	Unstructured encounter notes
35,391,676	Unique patients represented
50,569,507	Unique patient-provider encounters captured
30,937,597	Patients linked with structured claims
35,391,666	Patients linked with structured EHR

Behind every patient journey is a story told in the clinical notes from healthcare providers. Kythera's Clinical Notes capture these hidden details at scale.

Kythera's Clinical Notes capture the unstructured details recorded by clinicians - observations, impressions, and narratives from patient encounters. This dataset complements structured medical claims and Electronic Health Records (EHR) data, providing a more complete, contextualized view of patient care.

When linked with Kythera's structured data, such as claims and EHR, clinical notes illuminate diagnostic clues and relationships to accelerate understanding of disease origins, progression, and treatment. The result is rich insight into real-word patient journeys and a stronger foundation for research and analysis across data modalities and therapeutic areas.

KYTHERA

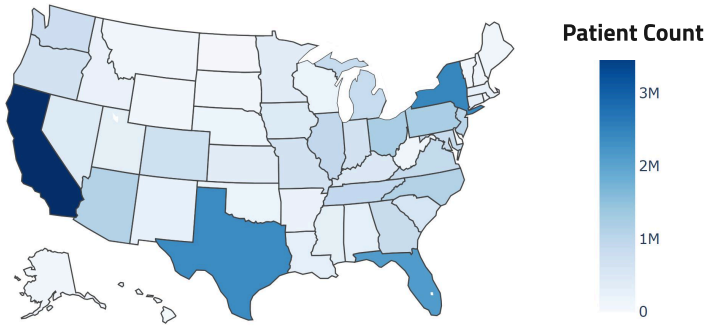


High linkage between clinical notes and structured notes unlocks deeper clinical context. With 87% of patients with clinical notes linked to both structured claims and EHR, Kythera's Clinical Notes reliably connect clinical narratives to coded diagnoses, procedures, drugs and events - all essential for comprehensive downstream analysis.

Longitudinal Depth & Broad Geographic Coverage

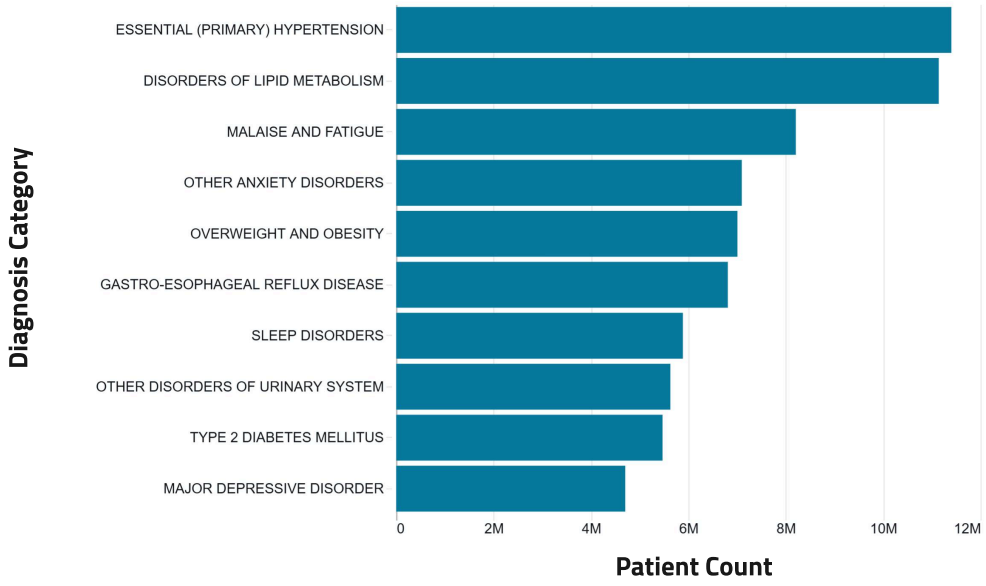
Kythera's Clinical Notes provide over a decade of coverage (2010 - present), enabling researchers to follow patients over time and reconstruct detailed real-world journeys of disease onset, progression, and care pathways, especially when combined with structured claims and EHR data.

Clinical encounters are drawn from all major U.S. states, providing geographically diverse visibility into patterns of diagnosis, care, and outcomes.



Wide Therapeutic Scope

The linkage with Kythera's Claims data, each clinical note can be associated with diagnostic categories. Kythera's Clinical Notes cover patients across 785 diagnostic categories, enabling analyses by therapeutic area and exploration of comorbidities, disease overlaps, and highly specific cohort definitions for complex and rare conditions.



What Makes Kythera's Clinical Notes Distinctive

The value of Kythera's Clinical Notes is in what can be extracted from unstructured text. Each note captures a clinician's reasoning, including presenting symptoms, diagnostic thought processes, and treatment decisions. This provides insight that structured fields and claims data cannot deliver. For most diseases, diagnostic delays and subtle symptom patterns are common, so this level of context enables more accurate patient identification and more comprehensive real-world analyses. It allows researchers to see the "why" behind the data, not just the "what." The below illustration highlights clinical entities in **GREEN** and Kythera's entity classification in **ORANGE**.

Comments: <PERSON> is an 83-year-old AGE male GENDER with past medical history significant for lymphoma CONDITION currently under chemotherapy TREATMENT, diabetes CONDITION, neuropathy CONDITION, bladder cancer CONDITION, and esophagitis CONDITION who presents to the emergency department in the afternoon hours of <DATE_TIME> for concerns of bowel obstruction SYMPTOM. He underwent exploratory laparotomy PROCEDURE with right hemicolectomy PROCEDURE and ileocolic side-to-side anastomosis PROCEDURE bilateral internal tap blocks PROCEDURE, umbilical hernia repair without mesh PROCEDURE, and placement of Prevena incisional vacuum management system PROCEDURE later that evening with <PERSON>. He received a 5-day course of Zosyn DRUG, as he had leukocytosis CONDITION with WBC 31K LAB, which eventually normalized during his admission. In addition, he had a postoperative ileus CONDITION, which slowly resolved during his admission as well. He returns today for scheduled follow-up. Overall, patient reports that he is healing well. He does not require any analgesics at this time. He has not NEGATION noticed any redness, formation, or drainage from his midline wound SYMPTOM. He has no issues with his diet and is passing flatus and having bowel movements. He denies NEGATION fevers and chills SYMPTOM. He offers no additional concerns today.

Privacy-Preserving Technology: Safe Storage & Analysis

All Kythera Clinical Notes undergo rigorous de-identification and security measures such as:

Tokenization: Identifiers in each clinical note are converted into a unique token (using 17 patient-specific features) to remove Protected Health Information (PHI) while still enabling accurate linking of notes to the same patient across structured claims and EHR data. This tokenization approach enables bridging unstructured and structured data without exposing personal identifiers.

De-identification: Clinical notes are fully obfuscated to remove personally identifiable information in the note text (e.g. names, addresses, dates) in accordance with HIPAA Safe Harbor standards. This automated de-ID process ensures the narrative content is usable for analysis while stripping out sensitive details.

Expert Determination Review: Clinical notes are expertly determined in order to ensure that protected health information is fully obfuscated or removed before being used for downstream research purposes.

Secure Query Environment: Clinical notes are stored in a highly secure, access-controlled, collaborative environment where users can share, access and analyze data, but never remove/export the underlying data without prior approval by all parties. Programmatic search queries find relevant notes and insights, that are then de-identified, tokenized, and expertly determined.

Example Research Applications

- Early identification of undiagnosed or underdiagnosed patients
- Mapping diagnostic pathways and time to diagnosis
- Evaluating treatment response and disease progression
- Supporting safety and post-market outcome studies
- Assessing clinical trial feasibility and site selection
- Conducting real-world evidence and comparative effectiveness research

Why Kythera Labs

Kythera Labs combines deep healthcare data expertise, data science, and state-of-the-art technology to make complex unstructured information more accessible and useful. See how deeper visibility into unstructured clinical data can enhance research and strengthen cohort discovery. Contact us to schedule a conversation or request a demonstration tailored to your priorities.

Kythera Labs is a data and technology company that enables organizations to rapidly ingest, de-identify, standardize, and analyze healthcare data with scale and speed using the Wayfinder technology platform, pre-configured data pipelines, data science toolkits, and remastered data sets for competitive advantages across markets.

KYTHERA™
DECIPHERING HEALTHCARE

346 Main Street
Franklin, TN 37064
www.kytheralabs.com