



Closing the Insight Gaps Left by Structured EMR Datasets

Realizing the Benefits of Unstructured Data from Medical Transcription Records

An Amplity Health White Paper | May 2022



While structured datasets from EMR systems serve an important business need and provide analytical value, they leave large gaps related to patient characteristics and treatment decisions. Fortunately, dictated medical transcription records (unstructured data), when mined effectively using leading digital processes, provide a much richer on-the-ground view of patient care at a level of detail unavailable from any other source.

Introduction

Pharma-company commercialization teams are always looking for reliable ways to capture and analyze patient and prescriber data. Information regarding patient care is important for patient documentation, of course, but it may be even more crucial for secondary use by drug manufacturers. In fact, in many respects, actionable patient- and prescriber-related data insights are the lifeblood of pharma marketing.

Historically, many life science companies have relied on large repositories of structured data from Electronic Medical Records (EMR) or claims datasets to identify patients and the providers who treat them. Structured data includes distinctly defined data types often tied to reimbursement and billable events. Common examples include procedures completed and medications prescribed, along with simple demographic information such as age, sex, height, weight, vital signs, etc., all of which are easily analyzed. Unfortunately, while structured datasets derived from EMR systems or claims provide value at scale and for certain analyses, they are incomplete in their capture of the care of the patient.

The Problem with Structured EMR-based Datasets

EMR companies' systems and platforms are typically purpose-built to collect and manage structured data. As a result, prescribers face the unenviable task of documenting patient-care content in a dizzying maze of EMR dropdown boxes and tiny notes fields. The limitations inherent in these tightly structured layouts open the door for major gaps in data capture.

Furthermore, for EMRs the task of properly de-identifying any unstructured data they do capture is complicated by substantial regulatory obstacles, including HIPAA challenges. They face significant barriers with respect to technology adoption, documentation options, and time. Not surprisingly, finding an EMR that provides de-identified fully unstructured data can be a fruitless endeavor.

In the end, the information generated from EMR datasets offers quantity, but leaves large qualitative gaps related to patient characteristics and treatment decisions.

Benefits of Unstructured Data from Medical Transcription Records

Notwithstanding the limitations of structured data from EMR, a treasure trove of unstructured qualitative data is available today, making it possible to learn about patient experiences straight from prescribers' recorded experiences. This data set includes insights into patient characteristics and treatment decisions at a level of detail unavailable from any other source.



Dictated medical transcription records (unstructured data) can be mined effectively using leading digital processes. This provides a much clearer on-the-ground view of patient care than structured EMR-based datasets, offering insight into symptoms, genomic data, biomarkers, reasons for treatment pattern (start, continue, discontinue, switch), and more.

And there are additional advantages. First, the only requirement for a prescriber is that they speak; no annoying dropdowns or tiny cells to fill in. HCPs can easily document all aspects of patient care in ways that structured EMRs simply do not allow. Second, prescribers are already producing these records as part of their standard documentation process as they dictate aspects of patient care. And third, recent technological advancements such as natural language processing (NLP), text mining, etc., allow for pinpoint extraction of specific, relevant information. The resulting datasets offer rich, qualitative descriptions of patient care.



How We Can Help

Amplity Insights is a provider of medical data sets mined from unstructured text transcriptions of patient physician interactions. Ours is the only database containing full-text transcripts of dictated physician notes for physician-patient encounters, providing data, analytics, and insights not available from any other single source. Amplity Insights offers a direct view into the treatment of patients and the rationale of providers while remaining fully HIPAA compliant. Our dataset covers the United States and includes >150,000 individual multi-specialty healthcare providers and 50 million patient records. What's more, we're adding 2 million new records every month. To learn more, visit <https://www.amplity.com/insights/>.

Conclusion

Life science companies typically rely on large repositories of structured EMR data to identify patients and the providers who treat them. Unfortunately, while structured datasets derived from EMR systems provide value at scale and for certain analyses, they are incomplete in their capture of the care of the patient.

Dictated medical transcription records (unstructured data), when mined effectively using cutting-edge digital processes such as natural language processing (NLP), text mining, etc., provide a much richer on-the-ground view of patient care than structured EMR-based datasets.

Using medical data sets mined from unstructured text transcriptions of patient-physician interactions, Amplity Insights brings a new perspective and a much broader understanding of how patients are diagnosed and treated.





Meet the Author:

Joe Turner

As Head of Insights, Joe oversees all Insights Operations activities, from data ingestion to project execution, while also being actively involved in all product development, vision, and strategic initiatives. Joe has spent 15+ years in the healthcare space. For nearly half of that time he built and led a successful medical transcription company, eventually becoming aware that there was unique secondary value contained within the unstructured records the transcription process generated. In the second half of his career he has focused on productizing that value to the benefit of clients and patients. Outside of work, Joe loves to spend time with his wife, two boys, and his extended family and friends in both the US and Argentina. He enjoys cooking, learning, playing music, and the outdoors.