Sentinel

Strengths and Limits of Claims and EHR-based Data Sources

Sentinel Community Building and Outreach Center (CBOC)

Sentinel's Use of Healthcare Claims and EHR-based Data Sources

One of the main focuses of the Sentinel System's Five-Year strategy (2019-2023) is to incorporate the use of electronic health record (EHR) data, supplemented with claims data, to support public health surveillance. The strategy also aims to incorporate emerging data science innovations, expand the type of regulatory questions that could be addressed, and enhance the methods tools to run queries.

SENTINEL IS BUILT UPON CLAIMS DATA

- The Food and Drug Administration Amendments Act (FDAAA) of 2007 mandated the FDA establish a post-market risk identification and analysis with access to disparate data sources including at least 100 million lives. In response, the FDA created the **Sentinel Initiative** and its **active risk identification and analysis** (ARIA) system. The FDA uses this national system based on real-world data (RWD) to conduct post-market safety surveillance.

- Sentinel was designed to leverage claims data, supplemented by EHR data available through data partners with integrated delivery systems. The approach provides Sentinel with a **rich source of longitudinal data** while safeguarding patient privacy.

RWE DATA ENTERPRISE ENHANCED USE OF EHRs

- To advance the use of Real-World Evidence (RWE), in response to the 21st Century Cures Act ("Cures Act"), the FDA created the **FDA-Catalyst** program with the goal of **establishing standards for high-quality RWD and evaluating RWE applications**.

- The FDA is dedicated to developing more **efficient mechanisms to access data from electronic healthcare records**, as laid out in the Sentinel System Five-Year Strategy.

LEVERAGING CLAIMS AND EHRs TO IMPROVE CAPABILITIES

- The Sentinel Innovation Center (IC) was established in 2019 to advance objectives in **Sentinel's Five-Year strategy** and increase **Sentinel's ARIA** capabilities.

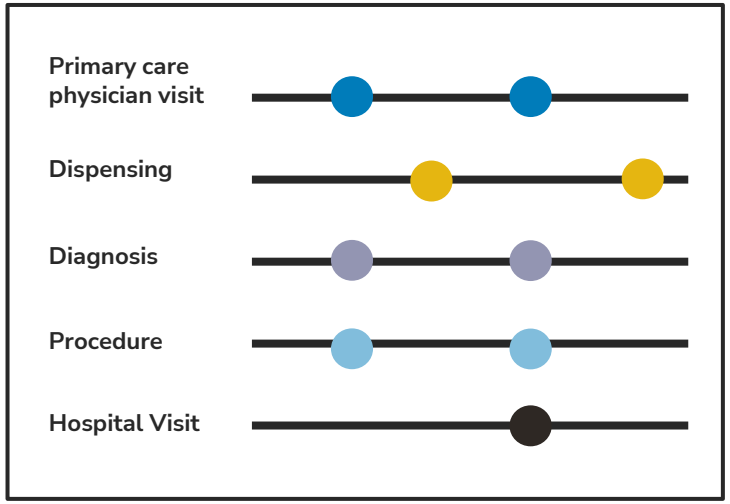
- FDA's vision for the Sentinel System is to use both **EHR and claims data sources for medical product safety surveillance** and expand the utility of real-world data for regulatory decision-making.

Differences in Claims and EHRs

Claims and EHRs differ in the depth and breadth of patient information captured.

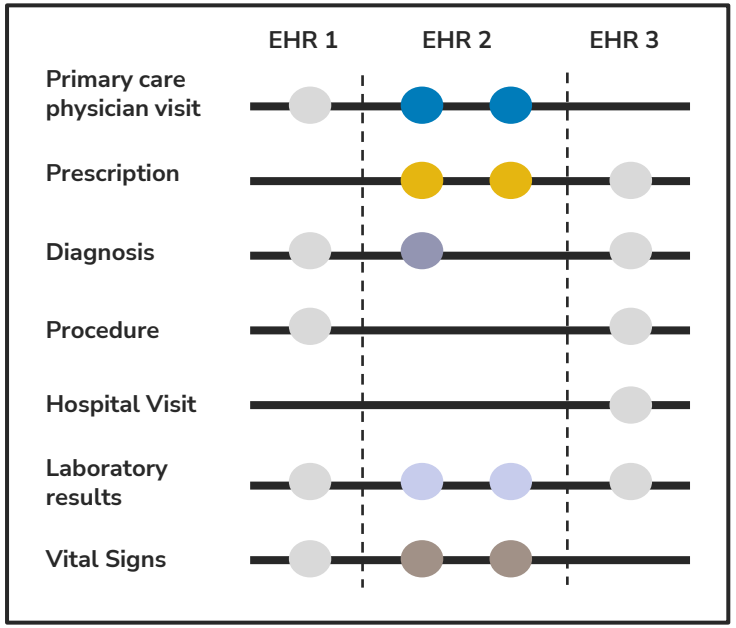
Healthcare Claims

- Claims data are generated from **billing codes** that physicians, pharmacies, hospitals, and other **healthcare providers submit to payers for reimbursement** (e.g., insurance companies, Medicare).
- Claims capture medical services provided **over time and across different provider settings for each individual patient,** making it possible to study causal relationships between drugs and health outcomes.
- Claims data provides **comprehensive data across all encounters**



Electronic Healthcare Records

- EHRs are electronic versions of a **patient's medical record**, that is maintained over time by a provider. EHR data may also include, but is not limited to, clinical data relevant to that person's care under a particular provider.
- EHRs provide **detailed clinical data for a single encounter.**



Considerations for Using Claims Data

Claims have been the primary source of data to study post-market safety due to reliability and ability to capture longitudinal information. However, using claims data alone may limit the ability to capture unstructured and detailed clinical information.

Considerations for Using EHR Data

EHRs can provide key clinical information documented by providers during a patient's healthcare visit. EHR data alone may have limited ability to follow patients across time or across different healthcare systems.

Benefits of Supplementing Claims Data with EHR data in Sentinel

Claims and EHRs each have benefits in providing comprehensive data on a patient's health status. Therefore, post market safety studies will benefit from surveillance systems that integrate both data sources and leverage the strengths of each.

Supplementing Claims data with EHR data can:

Strengthen Sentinel's ARIA capabilities supporting post market safety studies

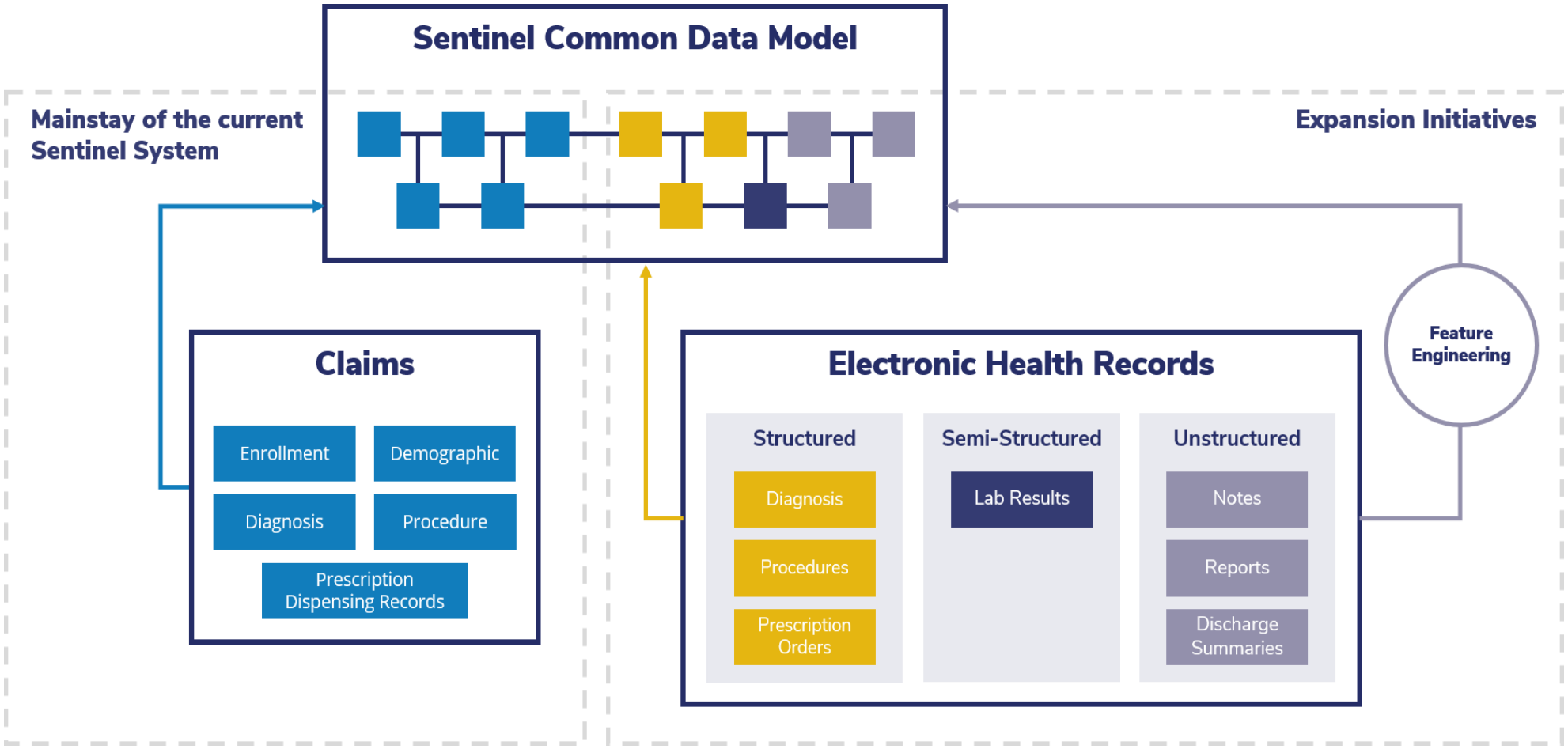
- EHR data can help increase ARIA sufficiency in the following areas (among others):
  - Ability to identify **health outcomes of interest**, including supporting validation of code-based algorithms
  - Detailed medical history and physical exam data
  - Expanding signal identification capabilities
- Sentinel will evaluate the extent to which natural language processing and machine learning approaches can be applied to EHR and claims data to leverage their respective strengths.

Improve Computable Phenotyping and Natural Language Processing (NLP)

- FDA identified a specific need to expand access to and use of EHRs to address a lack of valid and robust computable phenotypes for many **health outcomes of interest**.
- The FDA uses Sentinel to address these topics in a series of projects, including exploring the use of EHR data to better characterize key variables of interest.
  - Incorporating EHR data can enhance Sentinel's ability to capture in-patient exposures (leveraging hospital EHR systems) or adjust for confounding with additional covariates

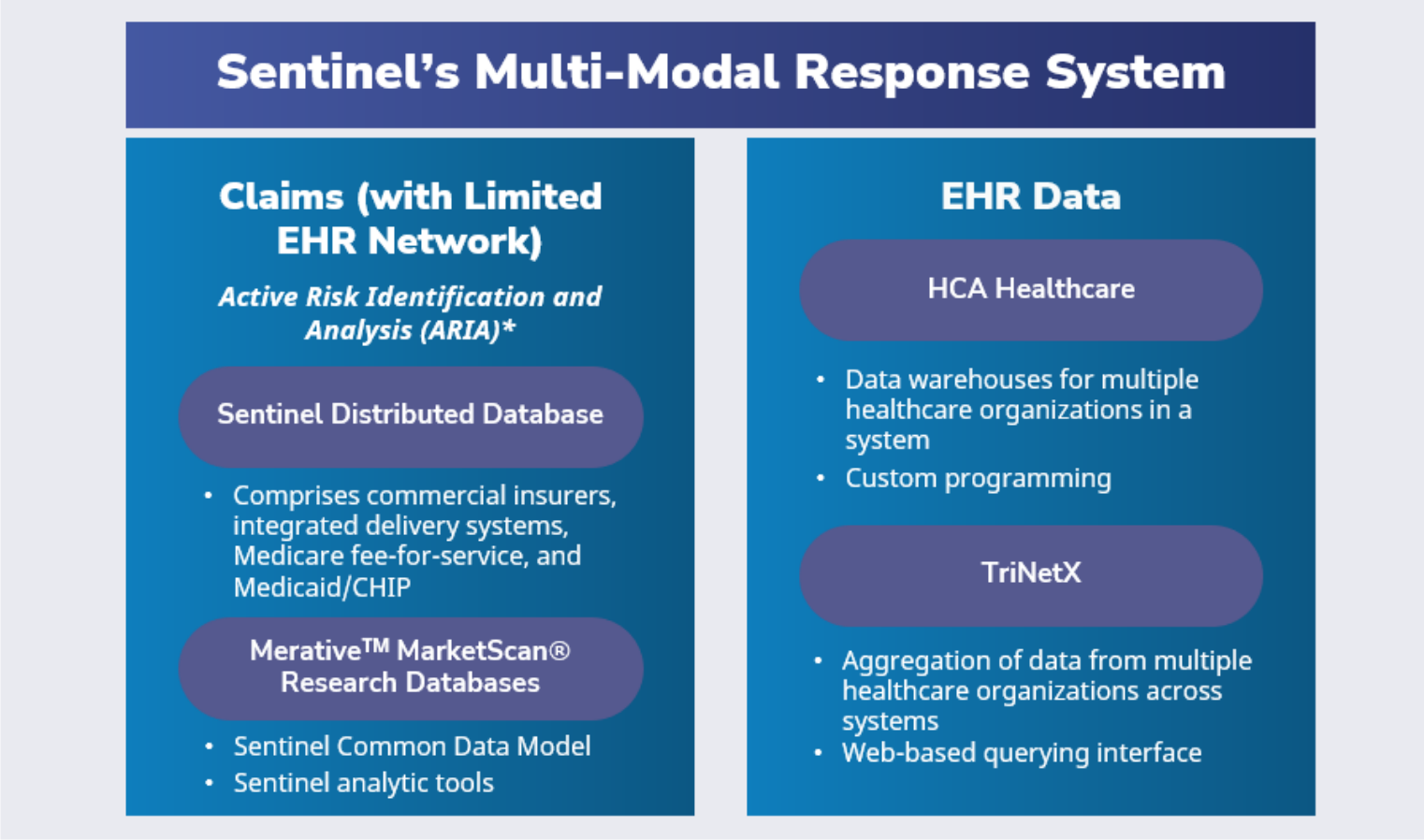
Integrating EHR Data into the Sentinel System

Integrating granular clinical patient-level EHR data into the Sentinel System positions FDA to better assess medical product safety using richer clinical data than are generally not available in insurance claims data.



Sentinel's Multi-Modal Response System

Sentinel's ARIA is complemented by pre-existing EHR data systems, enabling the selection of a data source that best fits a drug safety question of interest. The Active Risk Identification and Analysis (ARIA) system is the largest and most developed component of Sentinel. ARIA uses state-of-the-art analysis tools and a distributed database of standardized claims and claims linked with electronic health records (EHR) data to monitor the safety of medications.



*__Note__: *The Active Risk Identification and Analysis (ARIA) System is comprised of the Sentinel Distributed Database, the Sentinel Common Data Model, and Sentinel analytic tools.*

Case Study of Both Claims and EHR Data Use for COVID-19 Surveillance

The Sentinel System affords the unique opportunity to create large, well-characterized, national cohorts of patients diagnosed with COVID-19. Analyses of these cohorts has addressed key knowledge gaps that improve the understanding of the medical history, characteristics and outcomes of COVID-19 patients.

THE PROBLEM

with existing COVID-19 studies using limited data sources

- Early in the pandemic, published COVID-19 studies evaluated predominantly hospitalized patients, examined few determinants and outcomes of COVID-19, and frequently did not adjust for important confounding variables.
  - As a result, there were **major knowledge gaps on the epidemiology of COVID-19**, especially in certain subgroups.
  - **Population-based studies** evaluating the epidemiology of COVID-19 were therefore needed to provide insights into its mechanisms and consequences and inform the development of interventions to reduce the risk of adverse outcomes.
- "Further, even with direct access to such data, there can be **challenges with identifying certain types of care such as oxygen use**. Mechanical ventilation or supplemental oxygen use, indicators of disease severity, are likely underestimated if only procedure codes are relied upon, necessitating supplemental access to nursing documentation, diagnosis codes, or other data sources for more complete ascertainment.

THE APPROACH

to solving the problem by using both Claims and EHR data

USE CLAIMS DATA

In studying COVID-19, **medical history is an important data element** to track pre-existing conditions and chronic conditions.

- In Claims-based data sources, **enrollment requirements** can be selected to **ensure capture of medically attended events** during specified periods before or after COVID-19 diagnosis.
- EHR-based data sources do not include enrollment data so uniform look-back periods to capture medical history and comorbidities cannot be captured consistently.

USE EHR DATA

As COVID-19 primarily affects the respiratory system, capturing and **monitoring a patient's vital signs** are critical in understanding their respiratory rate, heart rate, and blood pressure.

- In Claims-based data sources, vital signs are not available as they do not appear on bills to the payor, and therefore cannot be used to measure basic bodily functions of a patient.
- In EHR-based data, **vital signs are captured** by providers and physicians, and therefore can **provide a patient's physiological conditions.**

Stay Engaged with Sentinel

https://www.sentinelinitiative.org/engage-sentinel

Newsletter

Receive information about the latest Sentinel updates, assessment results, event information and changes to the Sentinel System.

Public Training

Request trainings on Sentinel data, tools, or other areas of interest.

RSS Feed

Stay informed about the latest updates to the Sentinel website by subscribing to our RSS feed.