

# The Mini-Sentinel Distributed Database

## Year 1 Accomplishments

Lesley H. Curtis  
Duke University

January 12, 2011

# Agenda

- Creating the Mini-Sentinel Common Data Model
- Developing the Mini-Sentinel Distributed Database
- Building the infrastructure
- Generating useful information

# Creating the Mini-Sentinel Common Data Model

- Develop guiding principles
- Review existing common data models
- Draft and revise specifications

## Guiding Principles (selected)

- Data Partners have the best understanding of their data and its uses; valid use and interpretation of findings requires input from the Data Partners.
- Distributed programs should be executed without site-specific modification after appropriate testing.
- The Mini-Sentinel Common Data Model accommodates all requirements of Mini-Sentinel data activities and may change to meet FDA objectives.

# Review of Existing Common Data Models: Lessons Learned

- It's feasible for multiple Data Partners to assemble patient-level files according to a common data structure.
- Data Partners can retain complete control of their data while working toward common objectives.
- It's necessary to evaluate carefully all coding schemes used by each Data Partner to ensure that variability is understood and addressed.
- Analytical imperatives can be met using a distributed model.

# Development of Common Data Model

- Straw-man common data model
  - Minimal transformation to maintain granularity
  - Leverage prior experience
- Data Partner review and comment
  - Can your site implement these specifications?
  - Are definitions of tables and variables specific enough?
  - Are important data elements not included?
  - Are the requirements consistent with your expectations?
- FDA review and comment

# Mini-Sentinel Common Data Model v1.0

- Describes populations with administrative and claims data
  - Has well-defined person-time for which medically-attended events are known
- Data areas
  - Enrollment
  - Demographics
  - Outpatient pharmacy dispensing
  - Utilization (encounters, diagnoses, procedures)
  - Mortality (death and cause of death)

# Developing the Mini-Sentinel Distributed Database

- Each Data Partner translated local source data to the common data model structure and format and documented the process in a detailed report.
- Questions and issues were discussed on weekly teleconferences.
- Transformed data were characterized using standard programs developed by the Mini-Sentinel Operations Center.



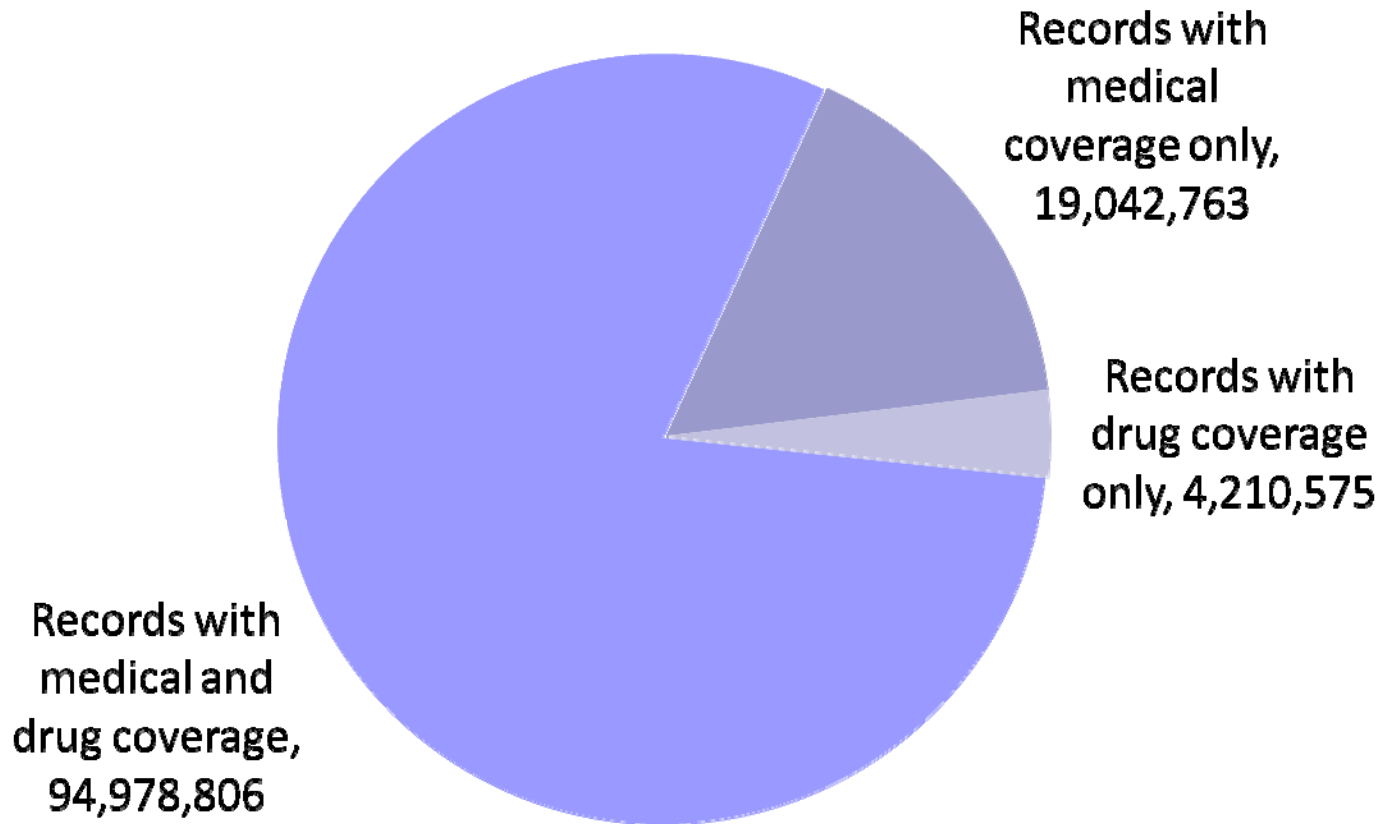
# Characterization of the Mini-Sentinel Distributed Database

- Overall, the Mini-Sentinel Distributed Database spans from 2000-2010
  - Most HMORN and Kaiser sites have data beginning in 2000
  - HealthCore has data going back to 2004
  - Humana has data going back to 2006

\*As of 7 Jan 2011

# Data Characterization: Enrollment\*

**Total Records in Enrollment Table: 118,232,144**



\* As of 7 Jan 2011

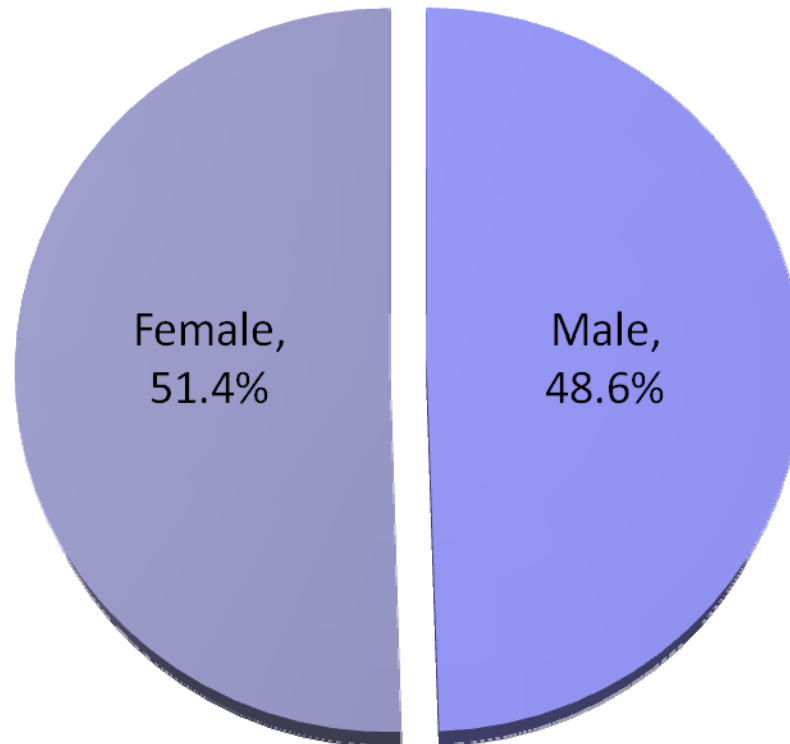
# Data Characterization: Enrollment\*

Unique members	71,152,385
Current <sup>†</sup> unique members with medical <i>and</i> drug coverage	22,482,689
Total person-years of observation time	167,295,216
Average person-months of observation time per member	28.2

\* As of 7 Jan 2011

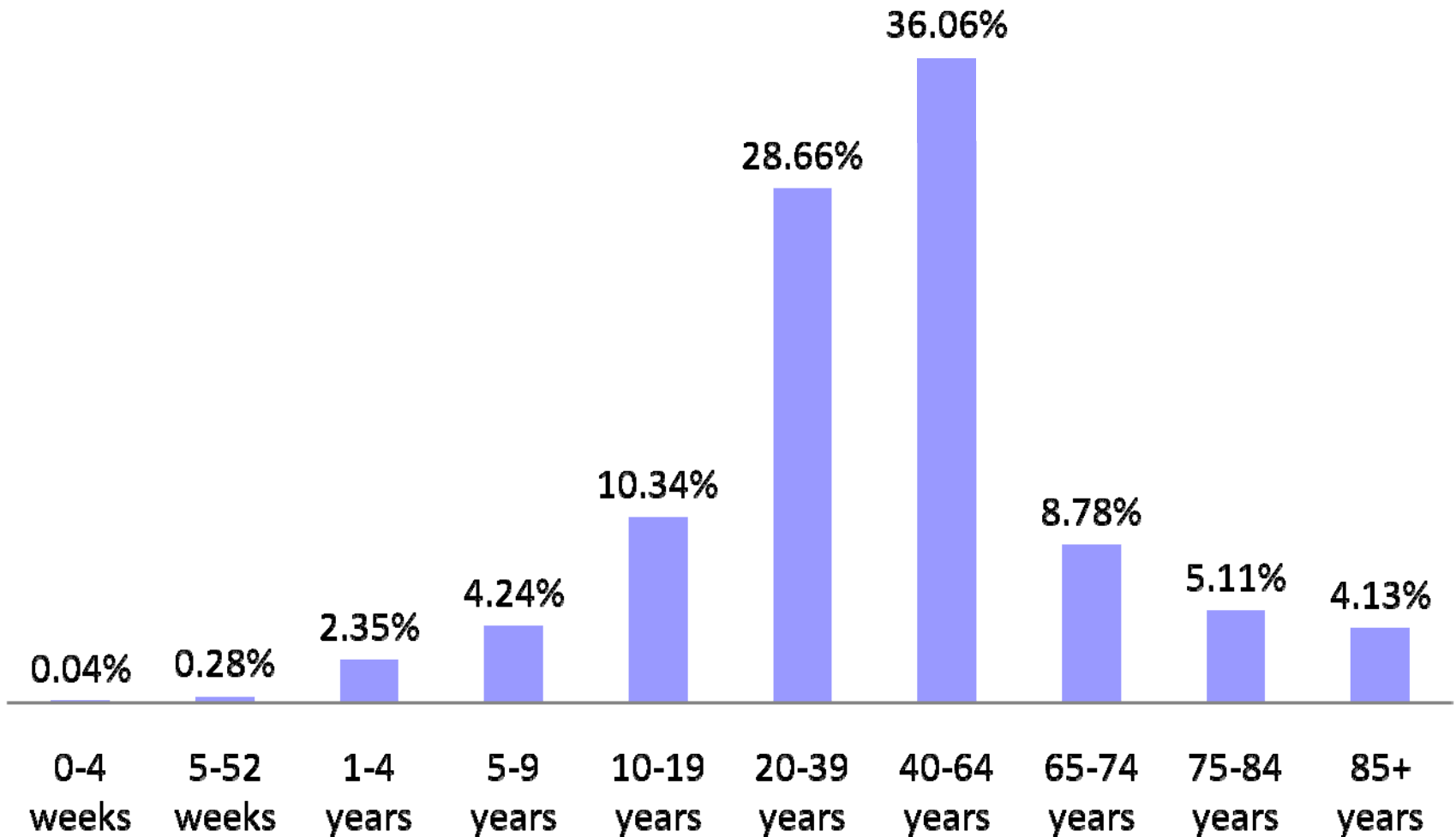
<sup>†</sup>Total number of unique members enrolled in the month of January 2009

# Data Characterization: Sex\*



\* As of 7 Jan 2011

# Data Characterization: Age \*

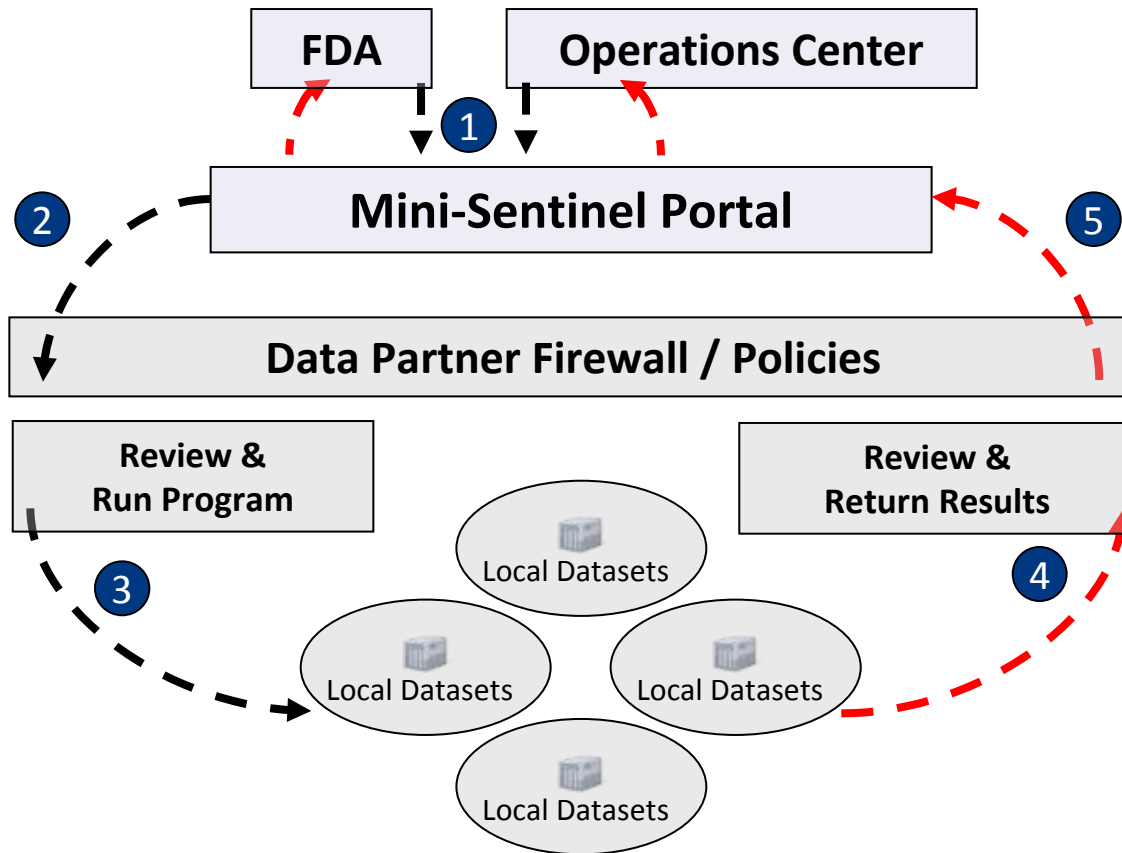


\* As of 7 Jan 2011

# Building the MS Infrastructure

- Standard programs to characterize and check quality of the Mini-Sentinel Distributed Database
- Formal assessment of Data Partners' technical environments
- Preparation for quarterly refresh cycles
- Empirical assessment of data latency
- Secure web portal for distributed analyses

# Mini-Sentinel Distributed Analysis



- 1- Query (an executable program) is submitted by FDA or Operations Center to the Mini-Sentinel Portal
- 2- Data Partners retrieve the query on the Distributed Querying Portal
- 3- Data partners review query and perform analysis locally by executing the distributed program
- 4- Data partners review results
- 5- Data partners return results to Distributed Querying Portal for review by FDA and/or Operations Center

# Current Modular Programs

1. Drug exposure for a specific period
  - Incident and prevalent users combined
2. Drug exposure with a specific condition
  - Incident and prevalent users combined
  - Condition can precede and/or follow
3. Outcomes following first drug exposure
  - May restrict to people with pre-existing diagnoses
  - Outcomes defined by diagnoses and/or procedures
4. Concomitant exposure to multiple drugs
  - Incident and prevalent users combined
  - May restrict to people with pre-existing conditions



Thank you