

# **The Emerging Primary Nature Of The Secondary Use Of Health Data:**

Lessons Learned at the National Center for Data to Health

Christopher G. Chute, MD DrPH  
Bloomberg Distinguished Professor of Health Informatics  
Professor of Medicine, Public Health, and Nursing  
Chief Health Research Information Officer  
Deputy Director, Institute for Clinical and Translational Research  
Johns Hopkins University, Baltimore, MD, USA

**JHCRN Trends in Clinical Research**

**Baltimore, 8 Nov 2018**

# Secondary Use of EHR Clinical Data

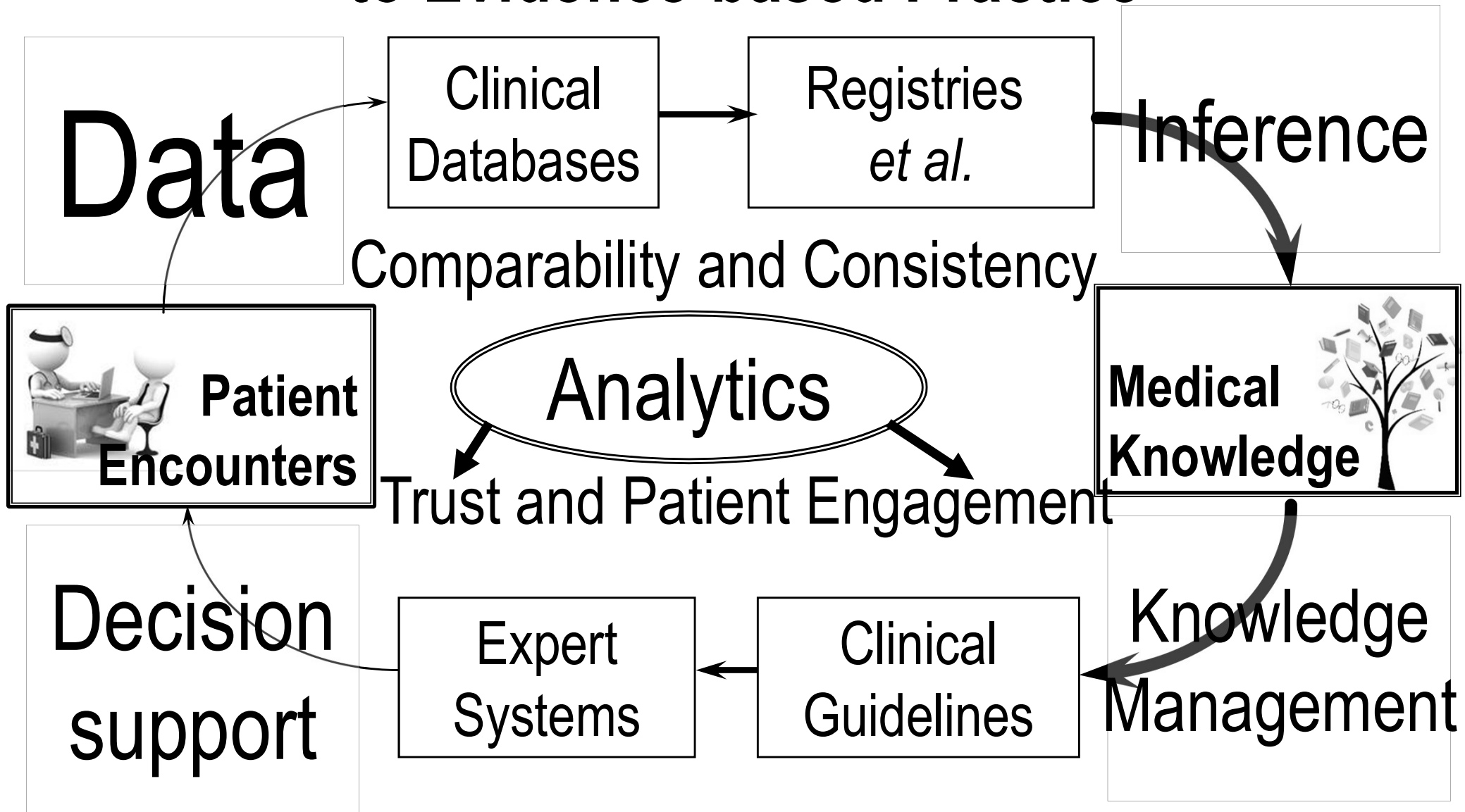
## Primary Use:

- Documentation of a patient's findings, conditions, and interventions for clinical care
- Records used one at a time

## Secondary Use

- Aggregation of many records for outcomes research and best evidence discovery
- Generalizations across many records

# From Practice-based Evidence to Evidence-based Practice



## Foundations for Learning Health System

# **Information Beyond Practice**

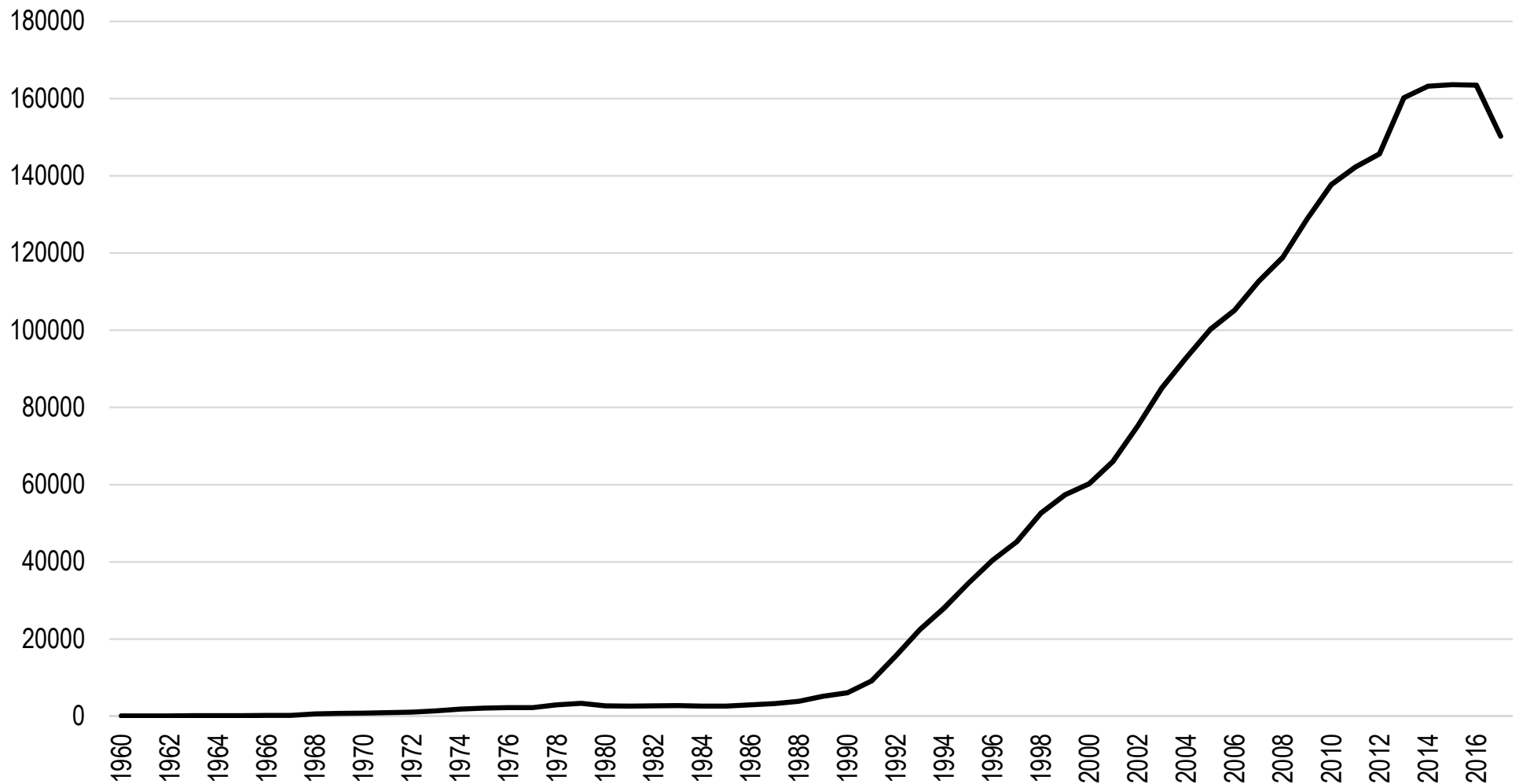
## **Secondary Re-use as Primary A Interest**

- Data Collected for Clinical Care Forms the Basis for Patient Experience Repositories
- The Importance of a Well Characterized, High Quality Patient Experience Repository May Exceed the Value of the Primary Information Many Fold

1994

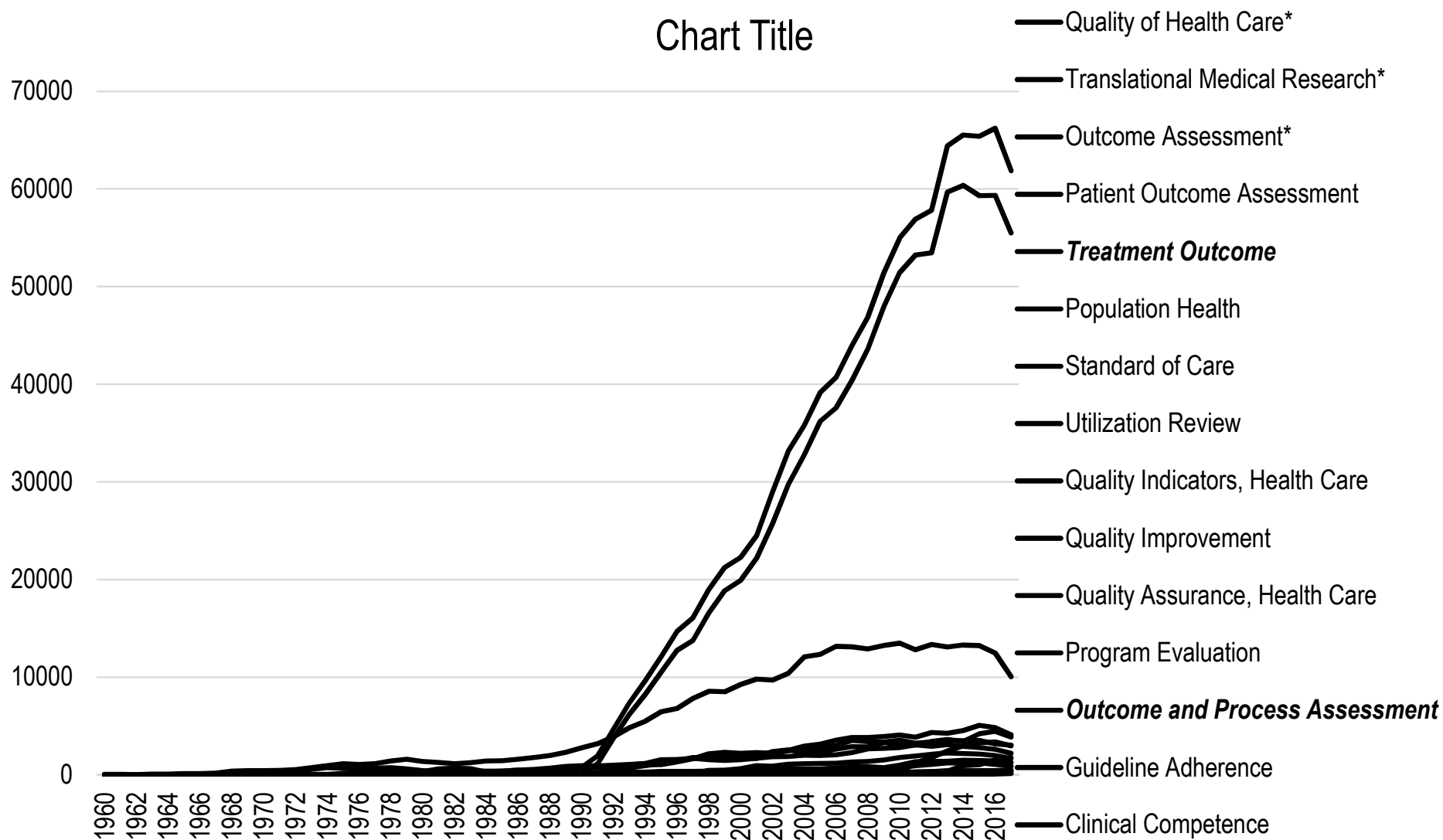
- Secondary use
- Contract with primary
- Bias of consent
  - Reproductive health
  - Sexually transmitted diseases
  - Mental illness

# Secondary Use PubMed Listings Total across all MeSH Categories





# Secondary Use PubMed Listings Total across all MeSH Categories



# CTSA Center for Disease to Health CD2H





# CD2H: Data-to-Health Coordinating Center for the CTSA Program

## *Accelerating Informatics Innovation to Advance Translational Research*

### CD2H Goals



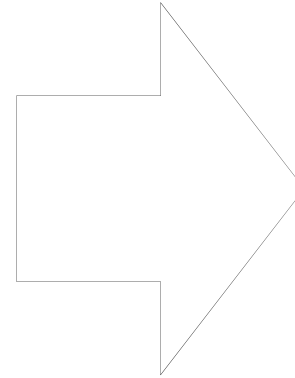
Make Data Easier to Share and Re-use



Make Tools More Accessible and Interoperable



Leverage Expertise and Foster a More Collaborative CTSA Culture



**Better  
translation of  
research and  
improved patient  
care**

# CTSA Aspiration

- Create a federated network of academic medical centers who can conduct “secondary use” (translational research) studies at scale
  - Approximately 65 CTSA funded “hubs”
- Realize the vision and opportunities of Big Data
- Enable novel analytic algorithms, including machine learning
  - Methods require more data than any single academic medical center can generate

# Data Sharing

- Federation is prevailing analytic design
  - No organization surrenders data
  - Analyses is distributed across organizations
- Significant challenges around common data model
  - TriNetX, PCORI, OMOP/OHDSI, ACT, etc.
  - Great promise in emerging clinical data standards



# Consent for Secondary Data Use

## Civic Rights vs. Societal Duties

- American College of Medical Informatics debate
  - November 6, 2018; San Francisco
- **Resolved: Biomedical researchers should have access to de-identified patient data without being required to obtain consent.**
- Pro: studies with significant non-participation are at best biased, and may be invalid
- Con: patients should have the right to determine the use of information about themselves

# What is the Real Risk to Participants?

- Inadvertent disclosure of medical data can be catastrophic for those affected
- Can research be conducted in a manner that mitigates or eliminates such risk?
- Good data practices
  - De-identify and anonymize to degree practical
  - Encryption at rest
  - Prohibit all raw data transfer from secure source locations (in situ analytics)

# Is “research” a Special Case

- Secondary Use analytics occurs at high volumes for quality improvement purposes
- Society appears to have limited concerns about such analytics
- It is only when one intends to share findings (publish) from quality improvement studies that “red flags” are raised
- There is a profound asymmetry in societal attitudes and fears around secondary use

# Where is this going?

- Secondary use has emerged as hugely valuable and critical to the vision of Learning Health Systems
- Challenges exist for data sharing
  - Required to achieve sufficient numbers for large-scale data science analytic techniques
- Balancing the needs of society with the rights to privacy is a pressing issue