Draft Proposal for a Rockefeller University Training Program to use Electronic Health Records as a Research Tool

Manish Ponda, Lesley Curtis, Richard Platt, Jonathan Tobin, Barry Coller

Premise: Electronic health data can be a powerful research tool. This is true not only for epidemiologic studies, but also for providing corroborating clinical data for basic science discoveries.

Goal: Create a training program to enable T1 researchers to utilize electronic health databases to answer basic science research questions using clinical data.

1. Understanding how and why electronic data are captured and what it means for research

2. Defining a query: translating a basic science discovery into a tractable clinical question

3. Utilizing and incorporating existing knowledge through effective literature searches

4. **Understanding human subjects protection and HIPAA regulations**

5. **The NIH Collaboratory, PCORNet, FDA Mini- Sentinel and NYC-CDRN**

6. **The structure of EHR databases**

7. **Principles and practice of distributed research data networks**

8. Identifying a database: parameters available; representative population; data quality

9. Constructing a query: working with clinicians and other stakeholders, epidemiologists and data specialists; inclusion/exclusion criteria

10. Choosing the appropriate descriptive and inferential statistical methods

11. Data Integrity: sources of error/threats to validity; importance of data refresh; internal measures of validity; longitudinal completeness

12. Teamwork and Leadership: collaborating with external data managers to design, develop and execute studies

13. Bioethical considerations in searching patient databases

14. Effective ways to communicate results and study output: tables; graphics; presentation formats

15. Case studies