

Barriers and Challenges: Data Collection and Merging Datasets

Moderator:

Keith Marsolo, PhD – Co-chair, EHR Core

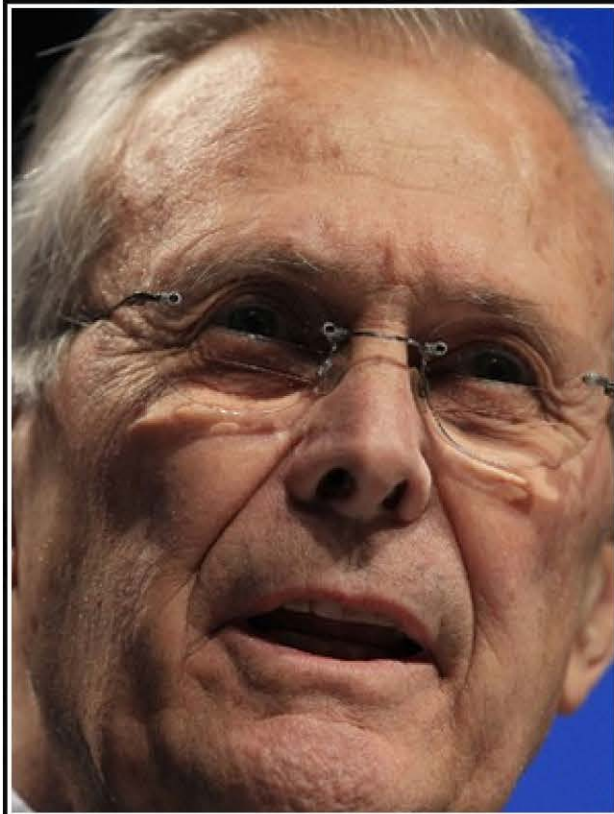
Department of Population Health Sciences
Duke University School of Medicine



**NIH PRAGMATIC TRIALS
COLLABORATORY**

Rethinking Clinical Trials®

Collecting and aggregating data from diverse data sets ... involves many expected and unexpected challenges



There are known knowns; there are things we know we know. We also know there are known unknowns; that is to say we know there are some things we do not know. But there are also unknown unknowns - the ones we don't know we don't know.

— *Donald Rumsfeld* —

AZ QUOTES

Running trials across multiple health systems, care settings and EHRs – what could go wrong?

Table 2. Reported challenges related to using EHR data in pragmatic trials

Survey Question: (Yes/No) Did you experience any challenges related to:	# Projects reporting: <i>n</i> = 20
IT staff turnover	11 (55%)
Integrating data from heterogeneous systems	11 (55%)
Using EHRs to support intervention delivery or health system change	10 (50%)
Missing data	8 (40%)
Assessing the validity or accuracy of EHR data	7 (35%)
Acquiring data	7 (35%)
Use of free-text or narrative data from EHR	6 (30%)
Data cleaning	6 (30%)
Completeness of data across EHRs	6 (30%)
Interoperability	5 (25%)
Data quality assessment	5 (25%)
Requesting data (ie, knowing which data are available)	5 (25%)
Identification and use of clinical phenotype definitions	4 (20%)
Determining or negotiating the frequency of data pulls	4 (20%)
Gaining permission to use data	2 (10%)
Need to upload data or write to EHRs	1 (5%)



Sharing lessons learned to reduce the number of unknown unknowns

- Data & technology-related challenges, potential solutions & recommendations for future studies
- Panelists
 - Ted Melnick – EMBED
 - Holt Oliver – ICD-Pieces
 - Margaret Kuklinski – GGC4H
 - Andrea Cheville – NOHARM

**Let's make
better mistakes
tomorrow.**