

Using Metrics to inform the Return on Investment (ROI): A Public Funder's Perspective

**Sean Coady, Division of Cardiovascular Sciences,
National Heart, Lung, and Blood Institute**

**NIH Virtual Workshop on Data Metrics
February 19, 2020**

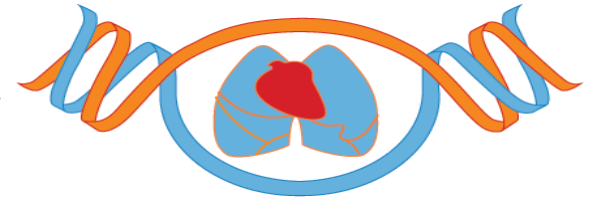


NHLBI Data/Materials Sharing



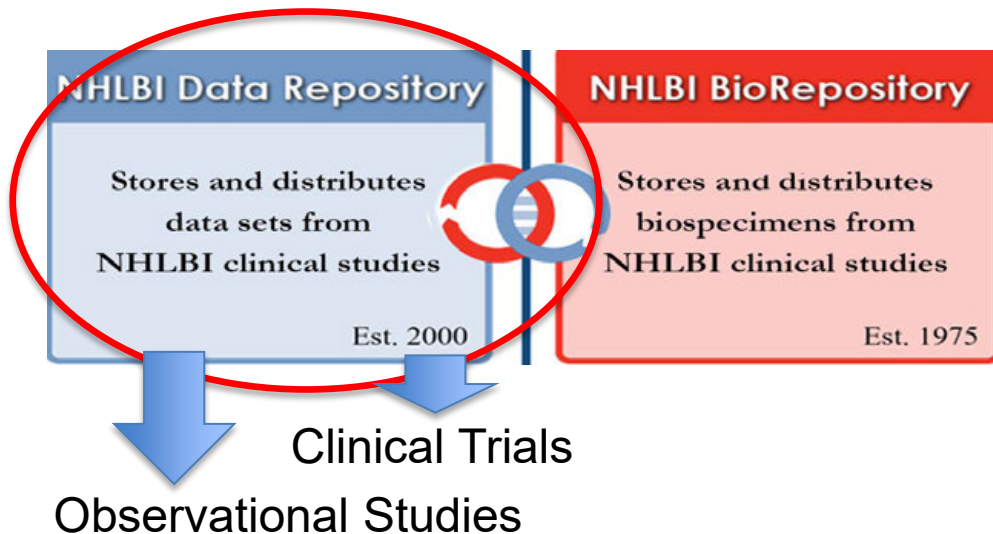
database of Genotypes and Phenotypes (dbGaP), GWAS, microRNA, gene expression profiling

Trans-Omics for Precision Medicine (TOPMed); WGS, metabolic profiles, protein and RNA expression, DNA methylation



BioData Catalyst

Storage, Toolspace, Access and analytics for big data Empowerment



BioLogic specimen & data repositories
INformation Coordinating Center (BioLINCC);
biospecimens & phenotypic data

Background – Data Repository

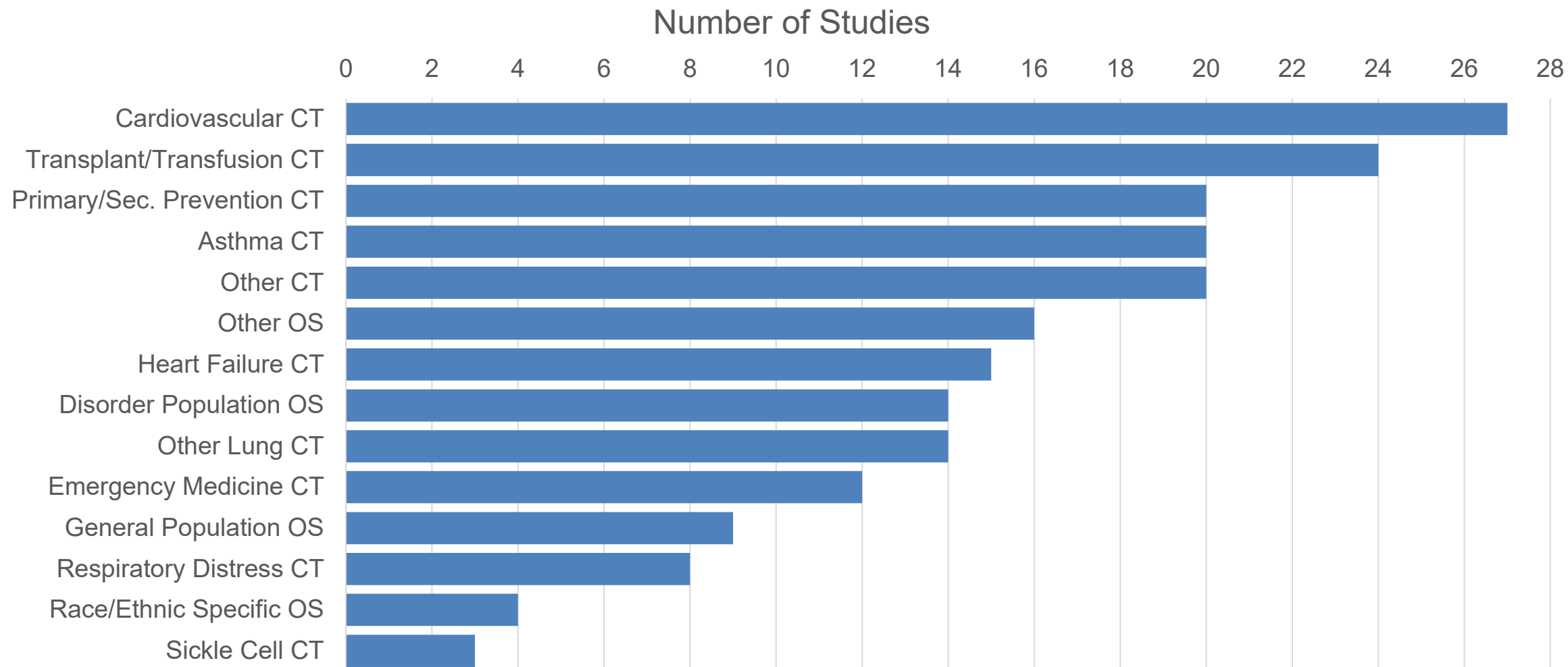
- 1999 NHLBI IRB approves data repository protocol. Website opens in 2000 (NIH IRB # 12-H-N198)
- Purpose:
In order to take full advantage of NHLBI supported clinical trials and epidemiologic studies and maximize their research value, data should be made available, under appropriate terms and conditions, to the largest possible number of qualified investigators in a timely manner.

Background BioLINCC

BioLINCC established in 2009 to coordinate activities of Data Repository and Biorepository

The screenshot shows the BioLINCC website interface. At the top left is the NIH logo and the text "National Heart, Lung, and Blood Institute". To the right is a search bar with a "Search" button. Below this is a navigation bar with links: "Home", "Biospecimen and Data Resources", "Procedures and Forms", "Build/Submit New Collection", "Contact Us", and "My BioLINCC". A large banner features the BioLINCC logo and the text "Web Access to NHLBI Biospecimens and Data". Below the banner are four circular icons with links: "Learn about the program: The BioLINCC Handbook (PDF - 2.2 MB)", "Request data and/or biospecimens from Open Collections", "View the NHLBI Biorepository Guide to Building Biospecimen Collections", and "View the NHLBI Biorepository video: NHLBI Biospecimen and Data Repository Program: Advancing Medical Research". On the right side, there is a "Featured News - 4 Items" section with a scrollable list of news items, including "New Study: SPRINT (data and specimens)" and "New Study: AsthmaNet-SIENA (data)". Below this is a "Recent News" section with "New Study: AsthmaNet-STICS (data)". At the bottom left, there is a "Search for Study Datasets and/or Biospecimens" section with a search input field and a "Search" button. Below the search field is a note: "Visit the study search page or enter keywords of interest (e.g. 'Heart Disease') to search for studies with available resources."

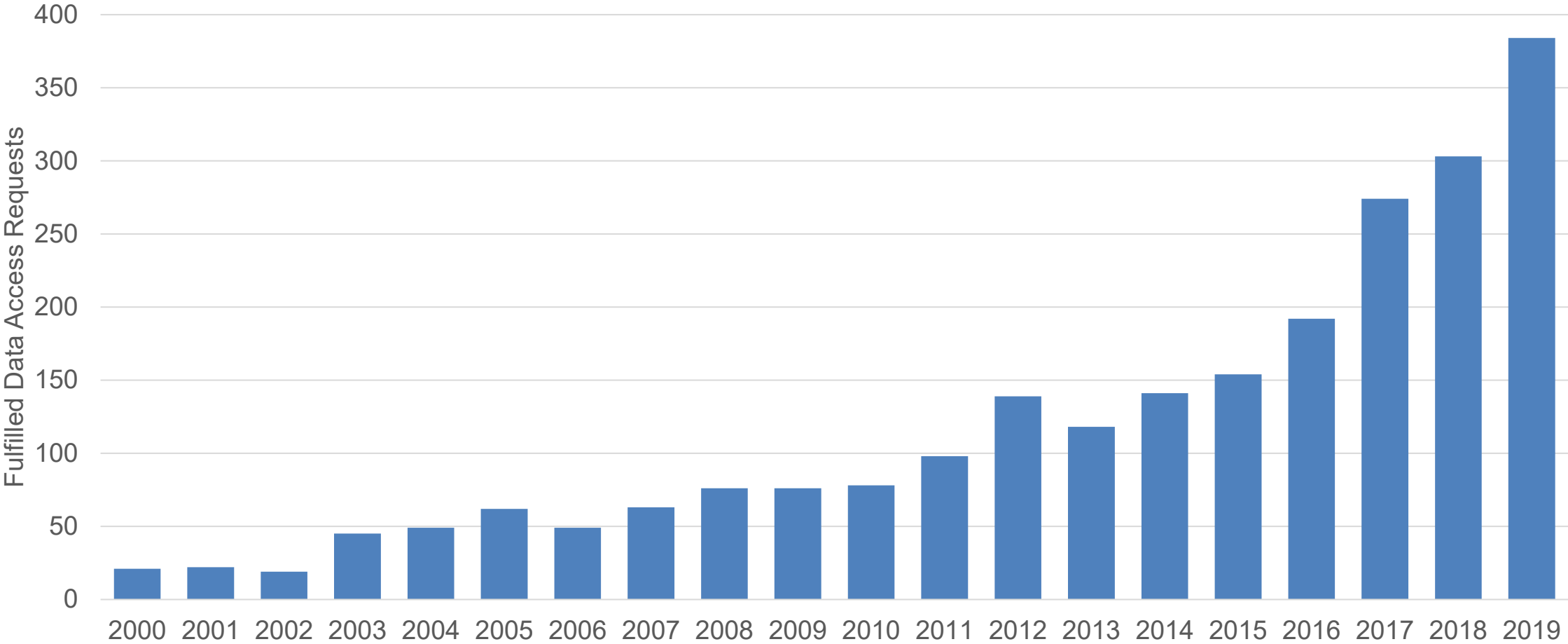
Current controlled access portfolio: 50 Observational studies (778,732 participants), 157 trials (460,816 participants)



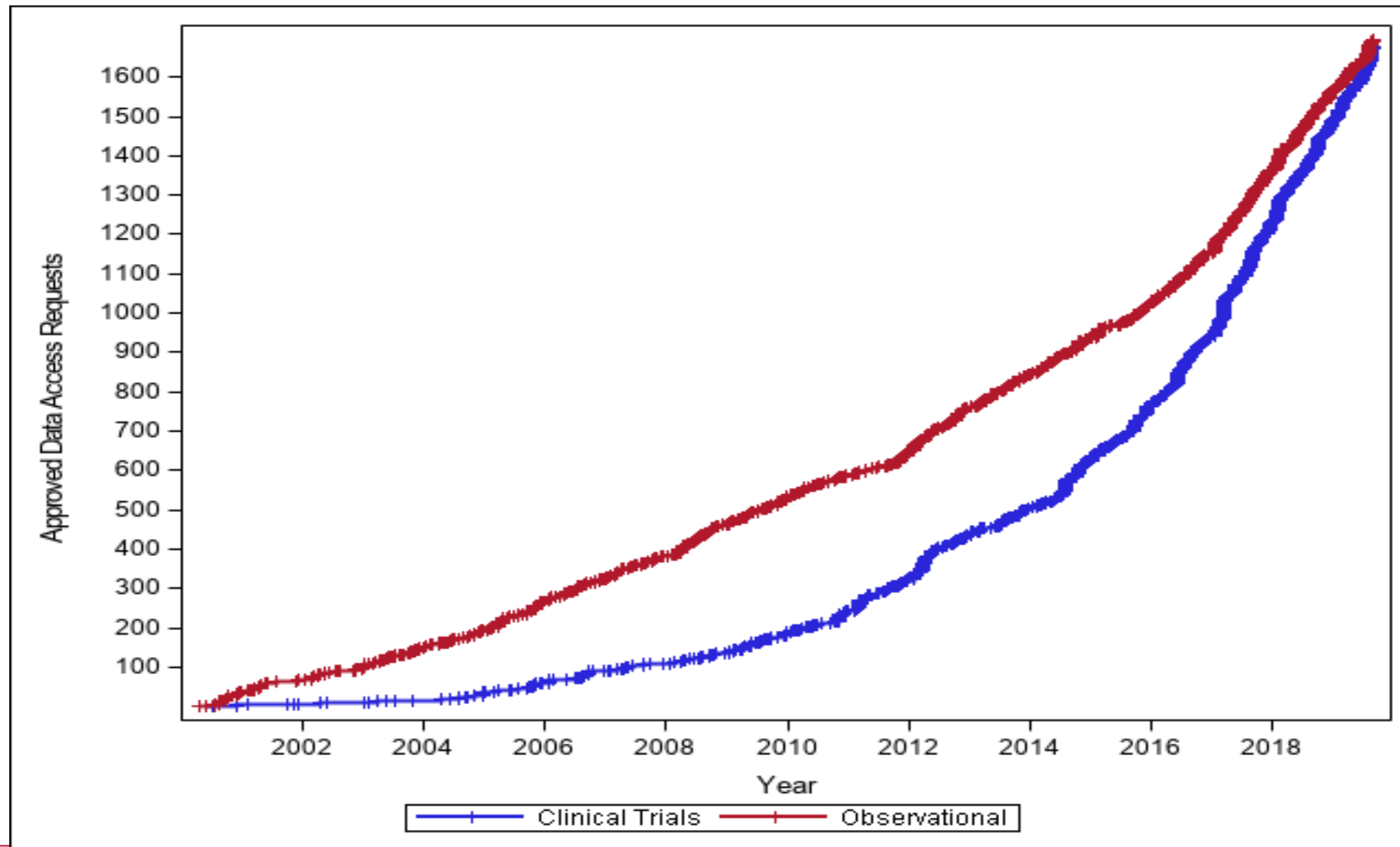
Data Repository Metrics

- Utilization Metrics
- Outcome Metrics
- Effort Metrics
 - Number of interactions and time to complete data access request
 - Support encounters
 - Monitoring

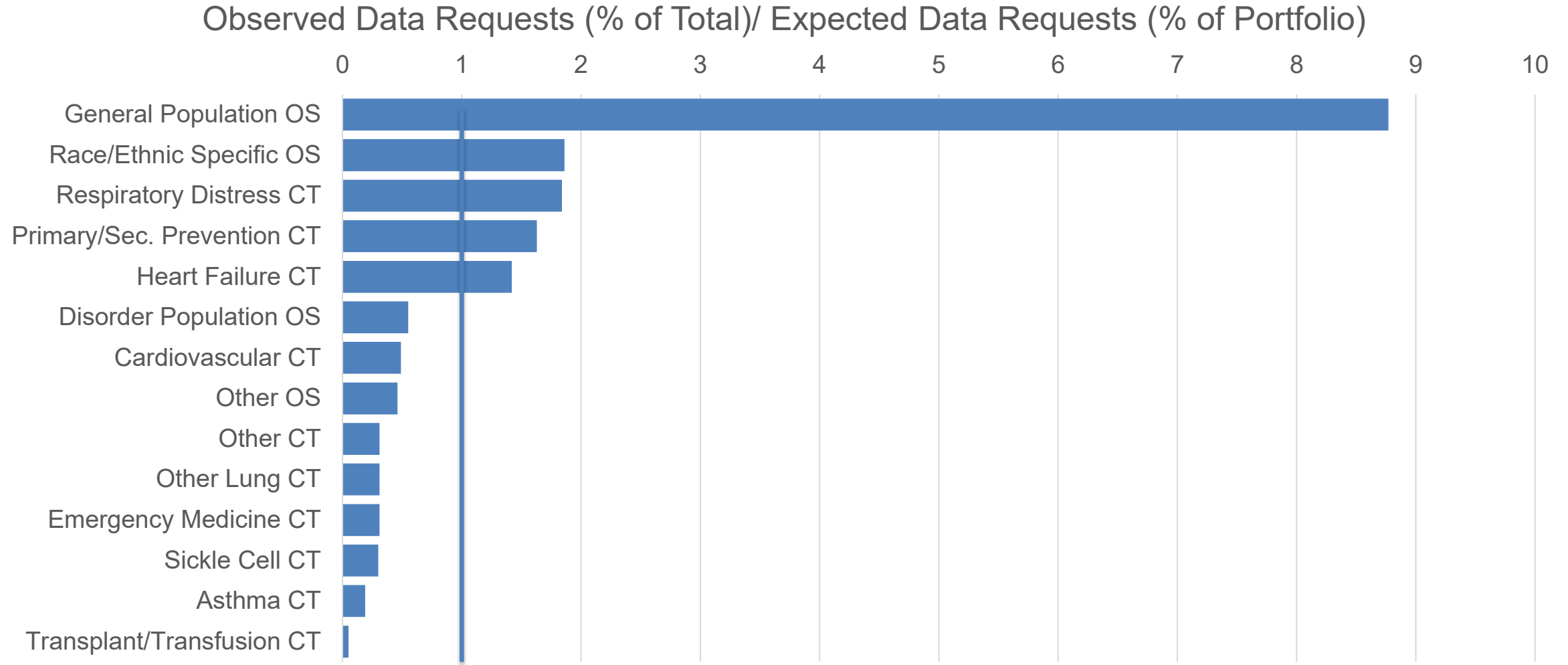
Utilization Metrics: Completed data access requests



Utilization Metrics: Cumulative data access requests for trial and observational study data (thru, 2019)

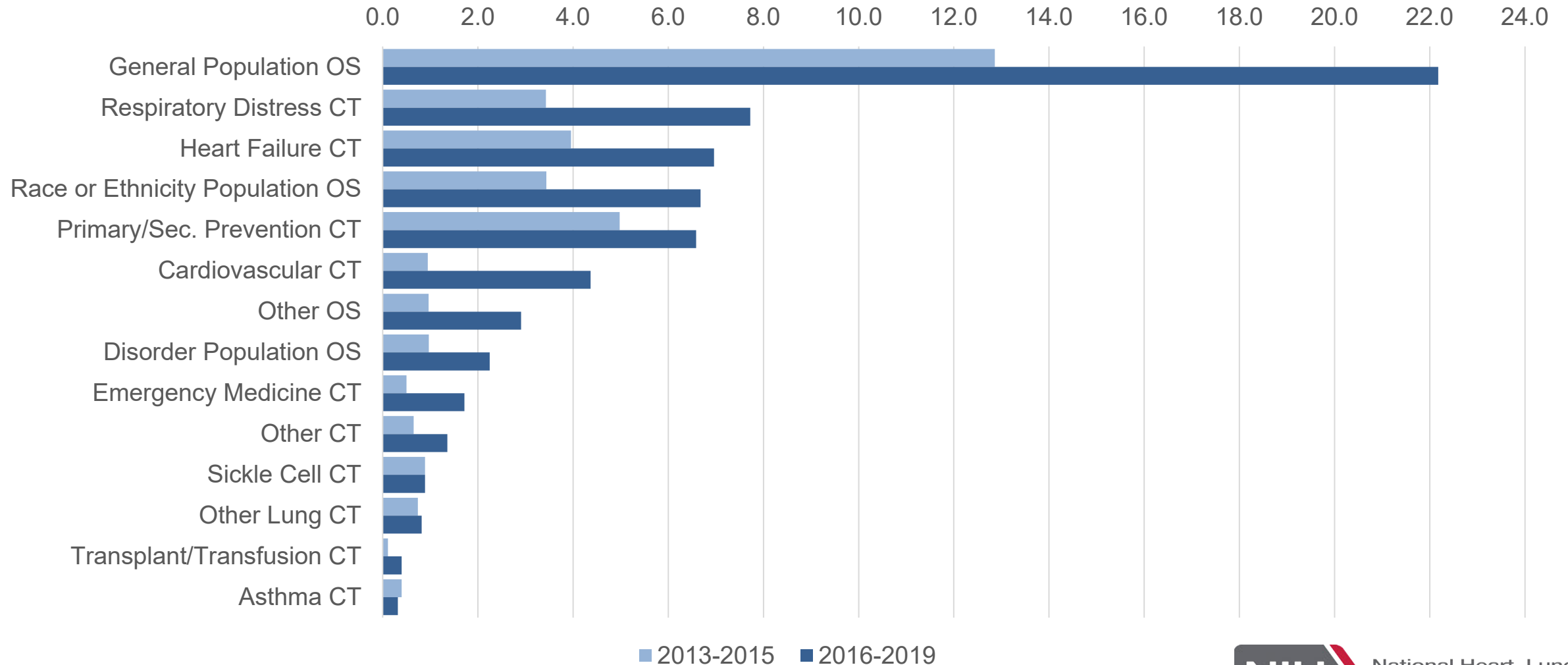


Utilization Metrics: Observed / Expected Utilization by study type

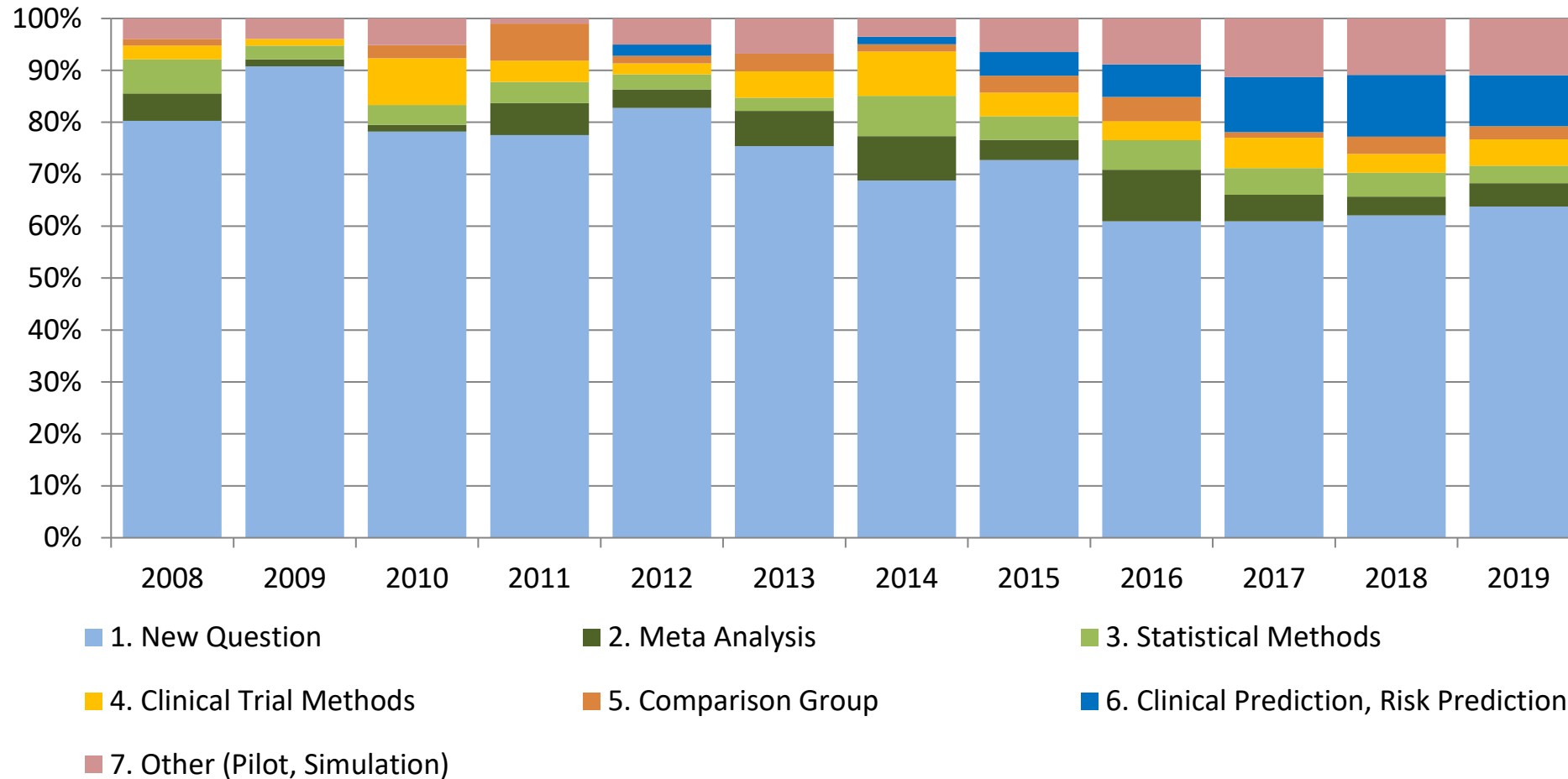


Utilization Metrics: Trends in Utilization by Study Type

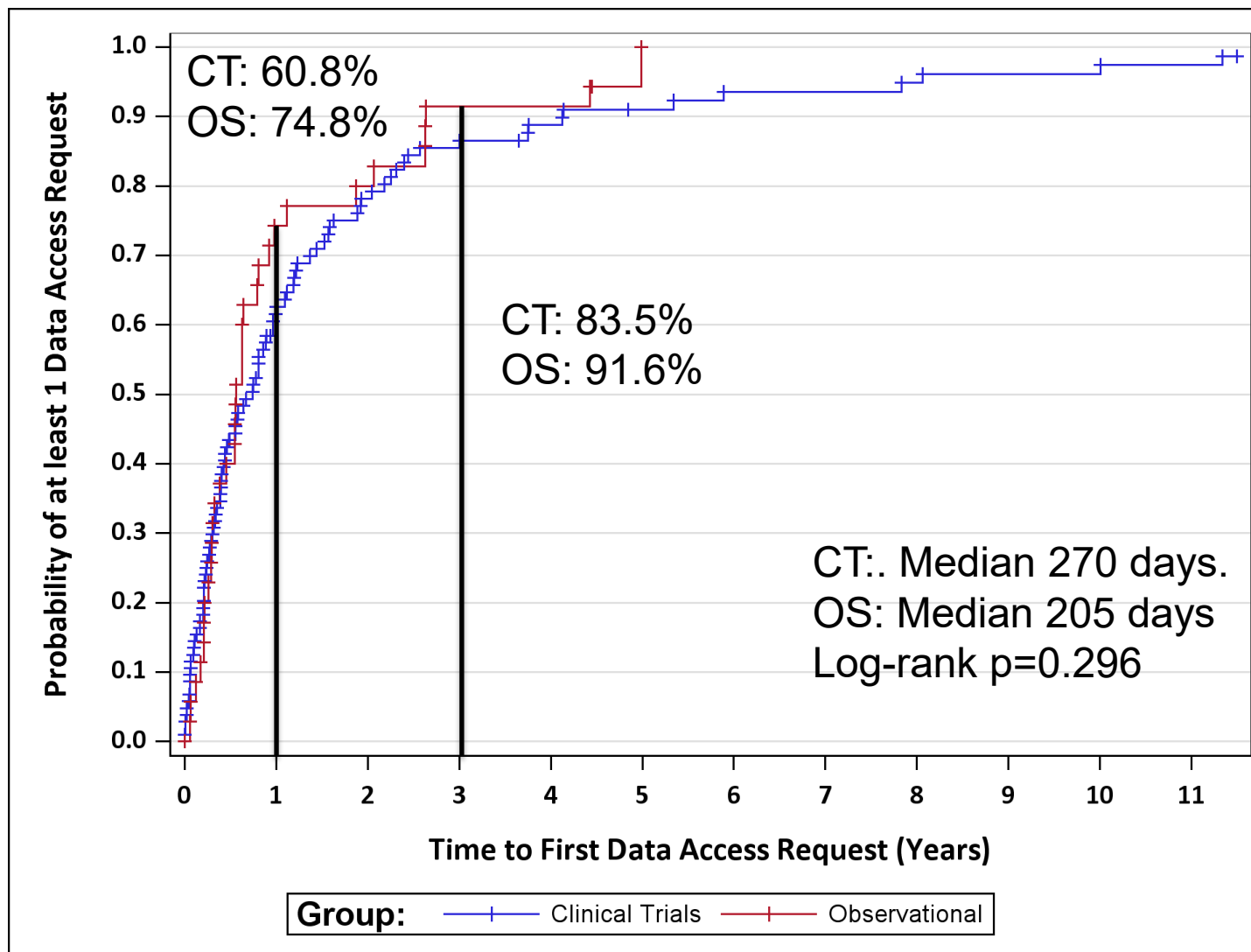
Average number of data access requests per study per year



Utilization Metrics: Primary reason for data access request

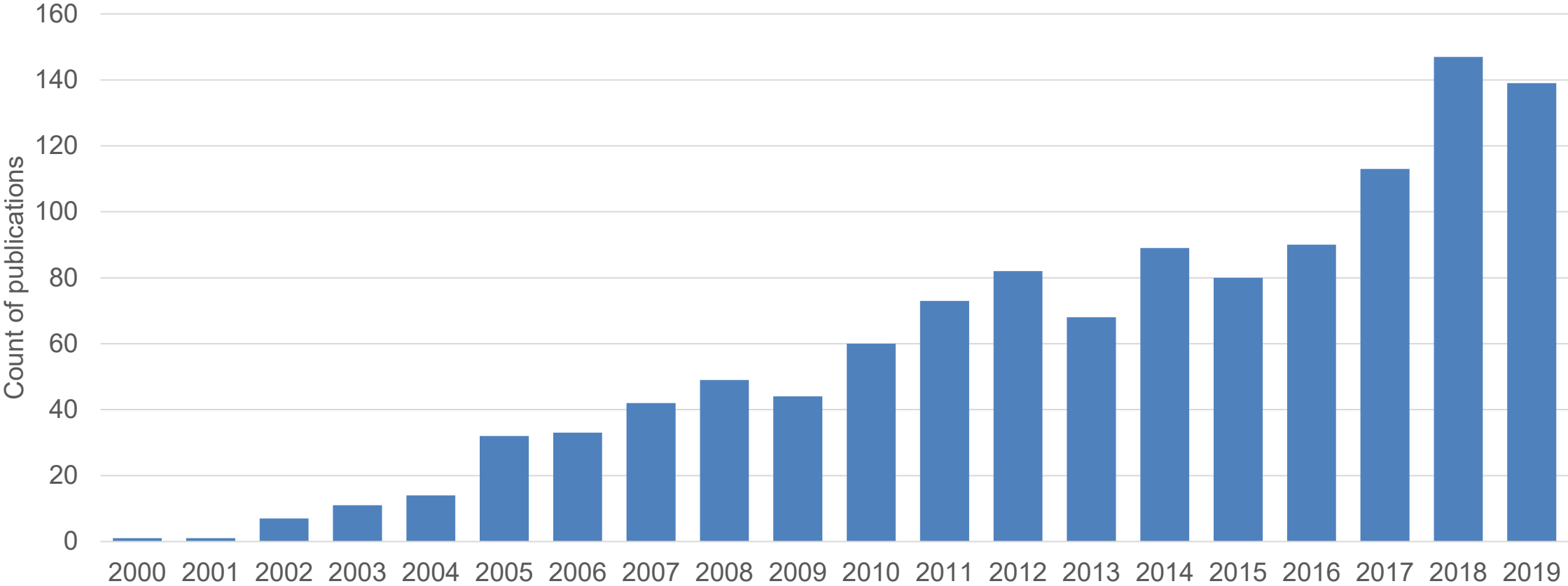


Utilization Metrics: Time from availability of study to first access request

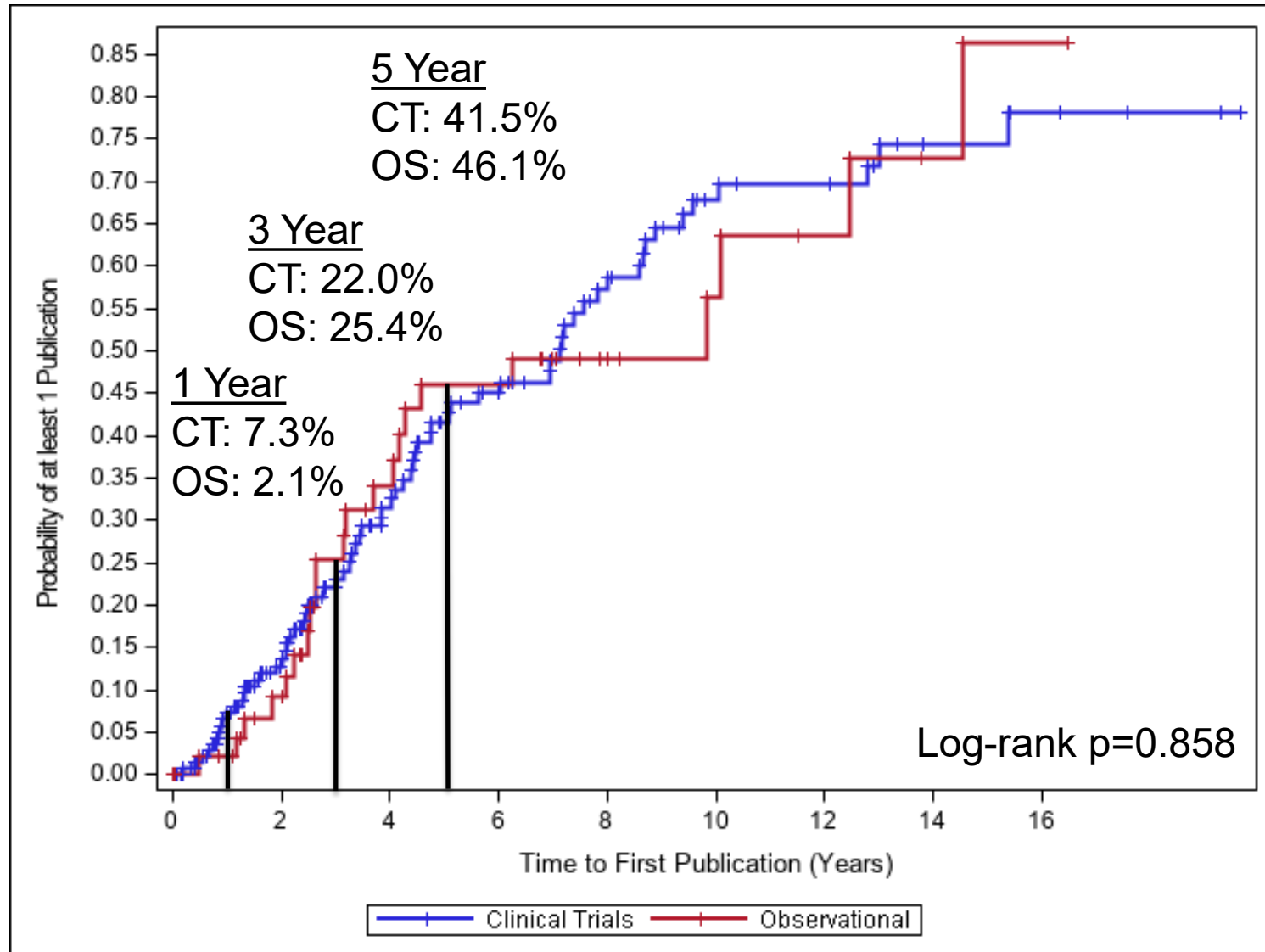


Outcome Metrics: Publications

Number of publications per calendar year



Outcome Metrics: Time to “Incident” publication by type of study (through 2019)

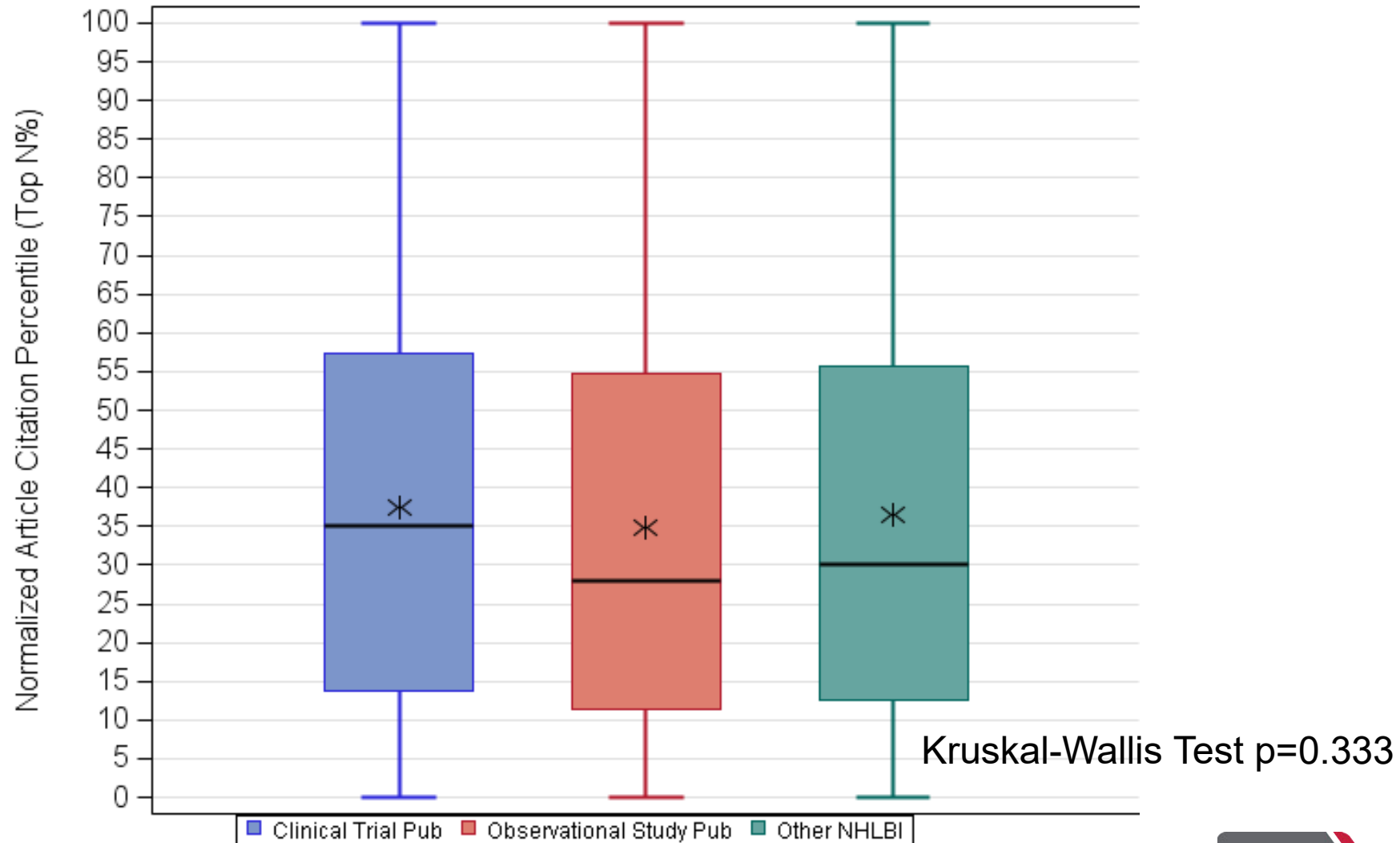


1179 publications
596 Included data from 1 or more trials

69 Trials 1+ publications

506 investigators have published at least once. Mean number of publications per investigator publishing=3.0

Outcome Metrics: Citation Percentile (Top N%) publications using repository trial data, observational data and 10% random sample of all NHLBI supported articles (Articles thru May 2015)



Outcome Metrics: Workforce training

- 23% of completed data access requests indicate primary user has 0-5 years of research experience
- QVR search of NIH grant applications found only two training applications (1 K01, 1 K99) mentioned BioLINCC

Messages from the metrics

- Demand for secondary use of data from clinical studies continues to steadily increase
- Growing demand for data from clinical trials
- Metrics can suggest gap areas
- Data from repositories fulfill a range of purposes
- Unclear if publication rates are low
- Citation metrics suggest quality of publications utilizing repository data are similar to publications supported directly by the Institute
- Need to better assess role of repositories in training new investigators