

## The NCTN/NCORP Data Archive: Improving the Impact of Clinical Trials Through Data Sharing

Written by Ying Huang

At the [inaugural symposium hosted by National Cancer Institute \(NCI\)'s Office of Data Sharing \(ODS\)](#) in October 2023, Dr. Margaret Mooney, Associate Director of the Cancer Therapy Evaluation Program (CTEP), Division of Cancer Treatment and Diagnosis (DCTD) at the NCI presented successes and challenges of sharing cancer clinical trial data, including National Clinical Trials Network (NCTN)/ NCI Community Oncology Research Program (NCORP) Data Archive. Dr. Mooney's presentation has sparked great interest in sharing clinical trial data and broadening the impact of NCTN/NCORP Data Archive for secondary use.

The [NCTN/NCORP Data Archive](#) is a centralized, controlled-access database, for storing and sharing patient level data generated from clinical trials supported by the NCTN and the NCORP. It started accepting datasets since 2017, with the goal of making clinical trial datasets available in a timely manner, under appropriate terms and conditions, to researchers who wish to analyze the data in secondary studies to enhance the public health benefit of the original work.

Over the decade, The NCTN/NCORP Data Archive has made strides in making cancer research data broadly available to the communities. The NCTN/NCORP Data Archive has pioneered the patient level data sharing efforts that made 134,537 individual trial participants' data available from 100 trials of NCTN groups and 25 trials of NCORP groups. The clinical trial infrastructure that these datasets are generated from over 2000 sites with an annual enrollment of 17,000-20,000 participants. Current publicly available datasets focus on datasets associated with primary publication with additional selected non-primary publications (published after Jan 2015) of clinical trials. The data sharing efforts pioneered by NCI through NCTN/NCORP Data Archive have expanded from phase 3 trials to randomized phase 2 trials, as well as facilitating prompt data sharing.

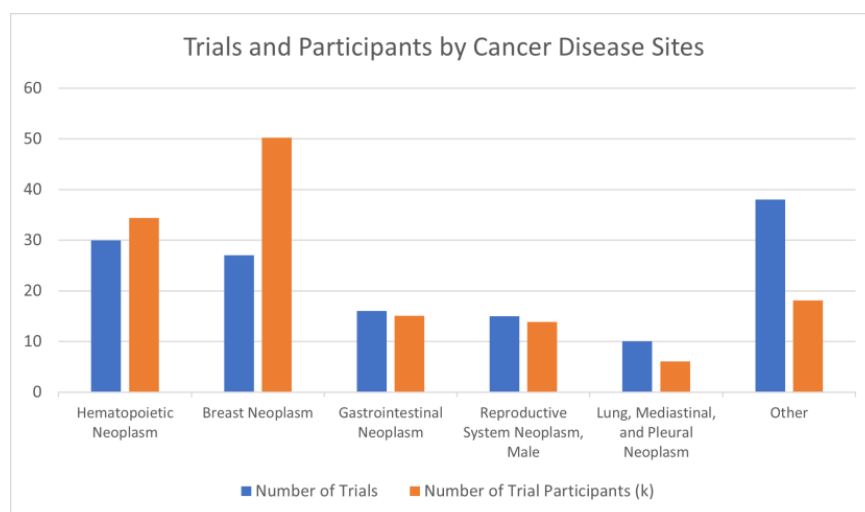
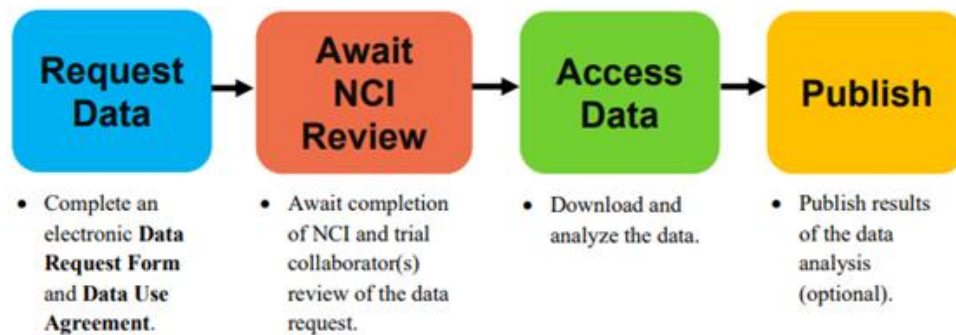
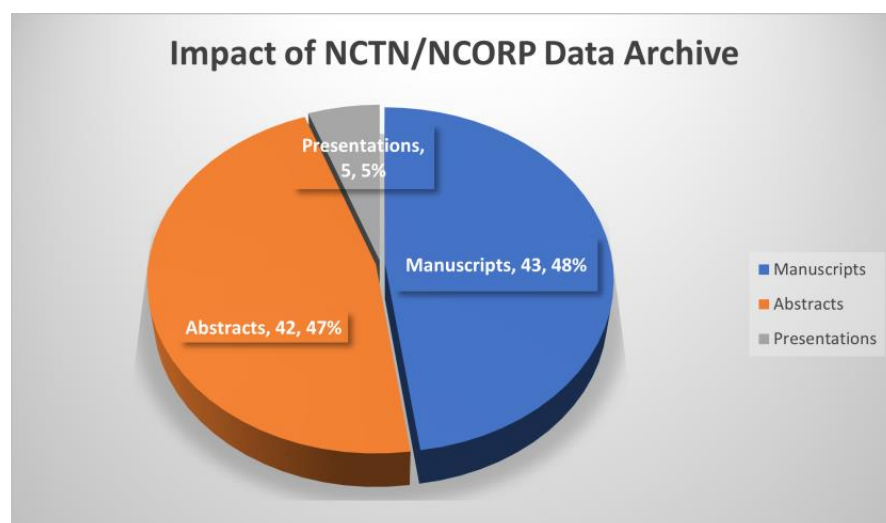


Figure: The NCTN/NCORP Data Archive hosts datasets from over 100 clinical trials including trials focusing on hematopoietic, breast, gastrointestinal, male reproductive system and lung neoplasm. The breast neoplasm clinical trials contain the greatest number of participants followed by hematopoietic.

NCTN/NCORP Data Archive has enhanced the clinical trials data collection in database of Genotypes and Phenotypes (dbGaP), Proteomic Data Commons (PDC), The Surveillance, Epidemiology, and End Results (SEER) with essential individual-level information regarding diagnosis, treatment, and outcomes. For selected trials, imaging data are also available via [The Cancer Imaging Archive \(TCIA\)](#). Data users could request datasets including individual-level, deidentified clinical dataset, data dictionary, and limited metadata fields from one or multiple clinical trials via registering as NCTN/NCORP Data Archive user and submitting online Data Request Form (see application process flow chart below, from [Usage Guide](#)). The review process prioritizes the necessary documentation of the data access application, rather than conducting in a scientific evaluation of the secondary research plan.



The NCTN/NCORP Data Archive has become a valuable resource not only for scientists in the cancer research field but also for interdisciplinary advancement, such as computer science. It has over two thousand registered users from various entities including academics, industries, and government, of which nearly 40% registered users are international data users mostly from Europe and Asia. Data shared via NCTN/NCORP Data Archive has so far resulted in total of 90 publications and presentations, while 56 have submitted to NCI in review to be released.



The NCTN/NCORP Data Archive plays an important role in sharing cancer data from individuals participating in clinical trials. The structured and well annotated patient-level data available

through archive serves as a vital resource for scientists and organizations dedicated to preventing and curing cancer.