Dear Workshop Participant,

The practical portion of our workshop is based upon real trial data. Following the workshop, if you are interested in gaining access to the trial data directly to continue exploring the SAS and R code provided, you can do so by applying for access using the following the steps below:

Clinical trial data accessed through Project Data Sphere: <https://data.projectdatasphere.org/projectdatasphere/html/home>

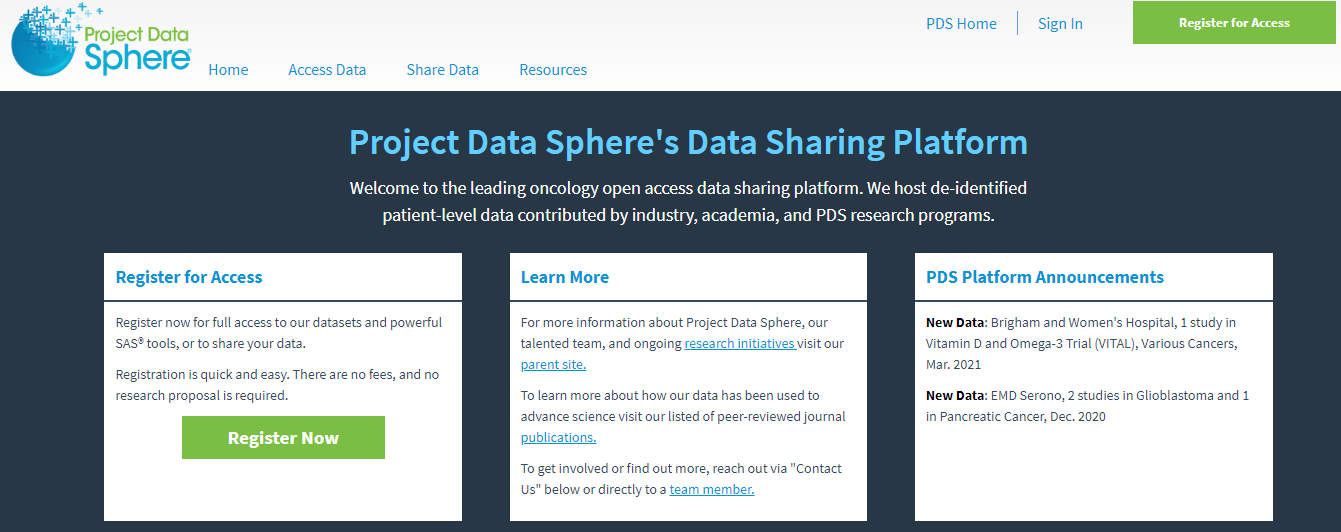
**Data Source: Project Data Sphere Clinical Trial Data**

# **Applying for Access to Project Data Sphere (Trial Data)**

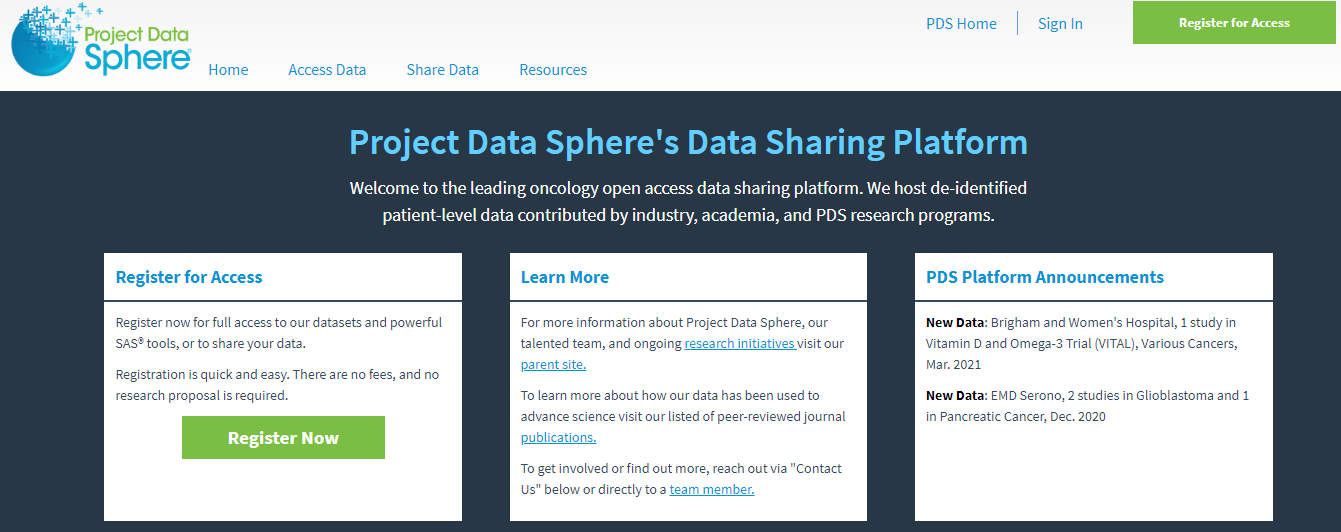
Please note: it may take up to 7 business days to process the data access requests

## **STEP 1: Applying for Project Data Sphere Access.**

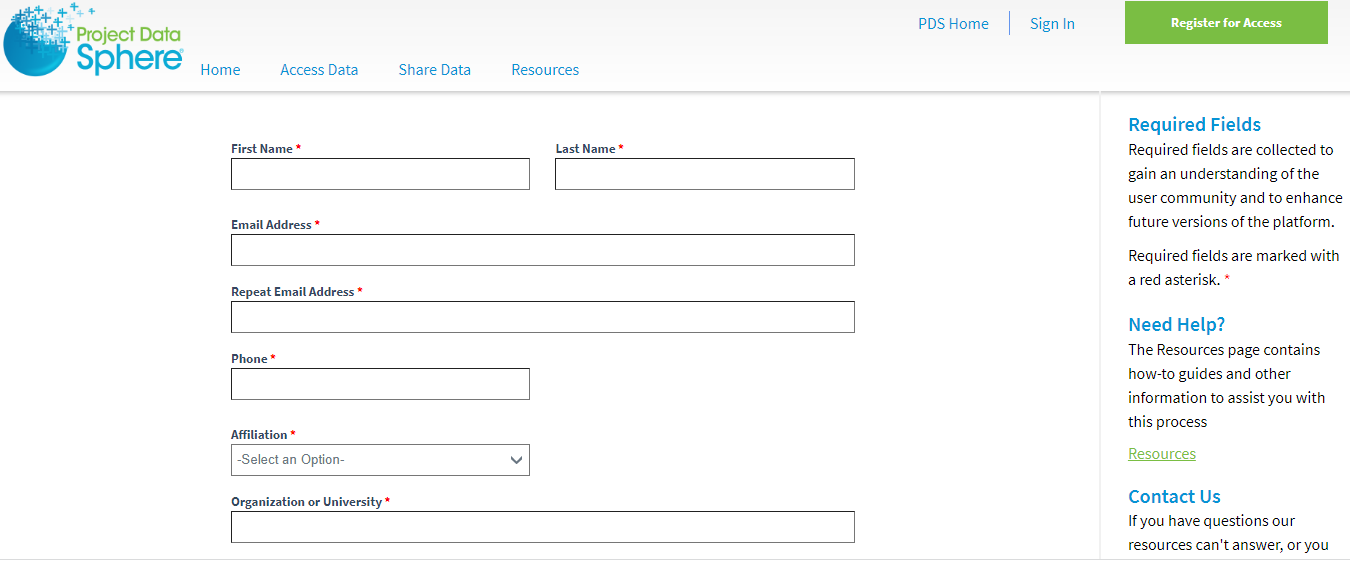
* **1a.** First, visit <https://data.projectdatasphere.org>. This will bring up the following webpage (**Note:** Do not use Firefox as a browser for this process – it is not likely to work):



* **1b.** Next, select “register for access” in the upper right-hand corner or the box on the left.



This will bring up the following page.



* **1c.** Complete all of the required fields. **PLEASE AVOID USING AUTO-FILL FEATURES – THIS MAY CAUSE PROBLEMS ON THE WEBSITE.**

To streamline the process, the “Research description” and “Research goals” information are provided below.

Research description: “A treatment’s efficacy in a clinical trial can trial differs from its effectiveness in clinical practice. Methods for generalizability and transportability can help bridge this gap.  I plan to use data from Project Data Sphere to combine RCT and simulated target population data to visualize differences in patient characteristics between the trial and target populations and to evaluate therapy effectiveness for a specific target population.”

Research goals: “I was a participant in a workshop on Nov 8, 2022 titled “Transporting treatment effects from randomized trials to real-world target populations” offered through Karolinska Institute in Stockholm, Sweden. During this workshop, I (1) learned about the foundations of epidemiologic methods for generalizability and transportability, (2) used visual tools to characterize trial and target populations and inform variable selection for transport models, and (3) applied this knowledge to an example of a trial for chemotherapy in metastatic colorectal cancer. I would now like to access the trial data directly for my further personal educational use.”

* **1d.** Once finished, agree to the Data User Agreement and submit.
* **1e.** Following submission, please allow up to 7 days for the application to be accepted. Once the application is accepted, you will receive an email at the address provided in the application. It should look like this:

Graphical user interface, text, application, email

Description automatically generated

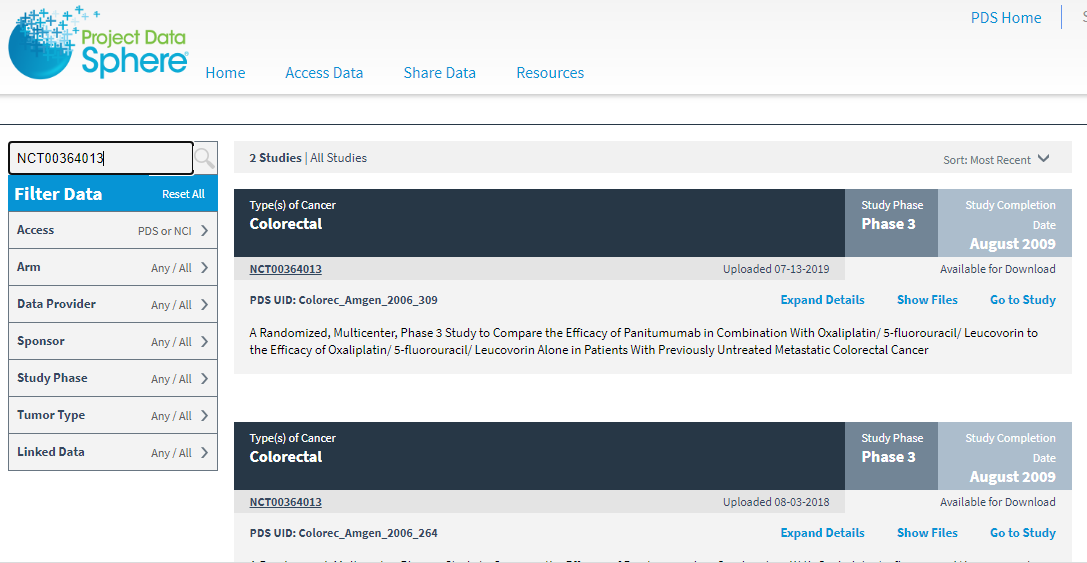
# **STEP 2: Downloading the Data.**

* **2a.** Return to <https://data.projectdatasphere.org> and sign in using the link in the upper right-hand corner.
* **2b.** Select “access data” from the banner along the top.

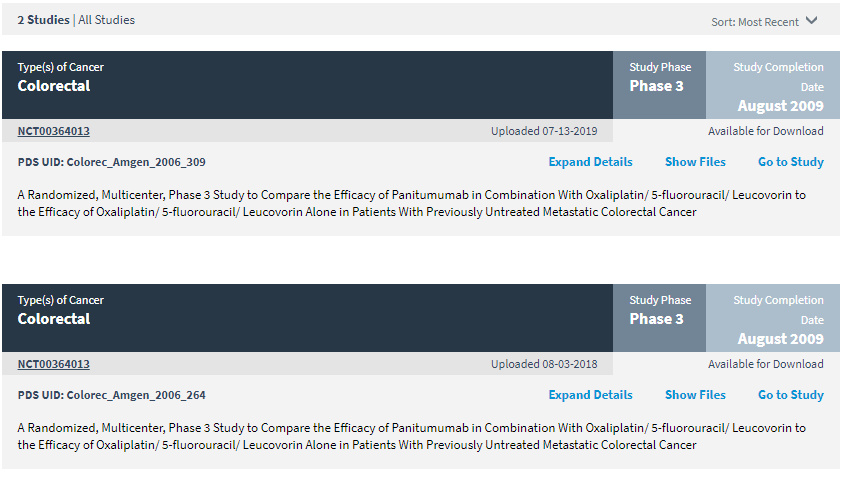
Graphical user interface, text, website

Description automatically generated

* **2c.** In the search bar, type the study ID: “NCT00364013”



* **2d.** This will bring up two results. Scroll down and select the second result. This study will have the following title “A Randomized, Multicenter, Phase 3 Study to Compare the Efficacy of Panitumumab in Combination…” with the “PDS UID: Colorec\_Amgen\_2006\_309”



* **2e.** Press “Expand details” and “Show files”. Read through the trial summary. Then scroll down to the files.

Graphical user interface, text, application

Description automatically generated

* **2f.** Only certain sections of the dataset must be downloaded for use with the provided SAS and R code. Scroll down to the bottom of the page, under “PDS\_DSA\_20050203.zip Contents”

Graphical user interface, application

Description automatically generated

* **2g.** Select “download” for the two data files called: “adsl\_pds2019.sas7bdat” and “biomark\_pds2019.sas7bdat”. Please note that this data is in SAS format. If using R, conversion would be required.

Graphical user interface, application

Description automatically generated