Dartmouth Cancer Center - Pathology Shared Resource

Data Type

Types and amount of scientific data expected to be generated in the project: ***Summarize the types and estimated amount of scientific data expected to be generated in the project.***

Describe data in general terms that address the type and amount/size of scientific data expected to be collected and used in the project (e.g., 256-channel EEG data and fMRI images from ~50 research participants). Descriptions may indicate the data modality (e.g., imaging, genomic, mobile, survey), level of aggregation (e.g., individual, aggregated, summarized), and/or the degree of data processing that has occurred (i.e., how raw or processed the data will be)

PSR generates imaging data associated with various histological stains including but not limited to hematoxylin and eosin (H&E), fluorescent antibody labelling (immunofluorescence) and semi-quantitative immunohistochemistry. Images are acquired on Aperio whole-slide scanners and provided to users in either svs or TIF format. Image sizes are typically 100s of MBs per slide.

Scientific data that will be preserved and shared, and the rationale for doing so: ***Describe which scientific data from the project will be preserved and shared and provide the rationale for this decision.***

Image files and associated metadata are preserved. These are the only sources of data generated for the client.

Metadata, other relevant data, and associated documentation: Briefly list the metadata, other relevant data, and any associated documentation (e.g., study protocols and data collection instruments) that will be made accessible to facilitate interpretation of the scientific data.

Standard Operating Procedures (SOPs) for the processing and generation of data are maintained by the PSR and are available to users upon request.

Related Tools, Software and/or Code

State whether specialized tools, software, and/or code are needed to access or manipulate shared scientific data, and if so, provide the name(s) of the needed tool(s) and software and specify how they can be accessed.

All image files generated by the PSR are provided in, or can be converted to, commonly used open source formats that can be analyzed using freely-avalialble tools such as ImageJ.

Standards

State what common data standards will be applied to the scientific data and associated metadata to enable interoperability of datasets and resources, and provide the name(s) of the data standards that will be applied and describe how these data standards will be applied to the scientific data generated by the research proposed in this project. If applicable, indicate that no consensus standards exist

svs and TIF files can be easily converted to OME-TIF, an open source file format widely used in the imaging field.

Data Preservation, Access, and Associated Timelines

Repository where scientific data and metadata will be archived: Provide the name of the repository(ies) where scientific data and metadata arising from the project will be archived.

Image files are stored shared with users via ShareFile, a secure cloud-based platform accessible to anyone at Dartmouth, and externally via a sharable link.

How scientific data will be findable and identifiable: Describe how the scientific data will be findable and identifiable, i.e., via a persistent unique identifier or other standard indexing tools.

All images receive a unique sample identifier that can be traced to the order placed by the user. Whole slide images include the slide label, which contains all information provided by the user for that specimen.

When and how long the scientific data will be made available: Describe when the scientific data will be made available to other users (i.e., no later than time of an associated publication or end of the performance period, whichever comes first) and for how long data will be available.

Data are stored on ShareFile and local servers for 3 years.

Access, Distribution, or Reuse Considerations

Factors affecting subsequent access, distribution, or reuse of scientific data: NIH expects that in drafting Plans, researchers maximize the appropriate sharing of scientific data. Describe and justify any applicable factors or data use limitations affecting subsequent access, distribution, or reuse of scientific data related to informed consent, privacy and confidentiality protections, and any other considerations that may limit the extent of data sharing.

There are no anticipated factors or limitations that will affect the access, distribution or reuse of the scientific data generated by the proposal

Whether access to scientific data will be controlled: State whether access to the scientific data will be controlled (i.e., made available by a data repository only after approval).

Data stored on ShareFile is only made available to the user requesting the service and can only be shared with others upon request.

Protections for privacy, rights, and confidentiality of human research participants: If generating scientific data derived from humans, describe how the privacy, rights, and confidentiality of human research participants will be protected (e.g., through de-identification, Certificates of Confidentiality, and other protective measures).

Oversight of Data Management and Sharing

Describe how compliance with this Plan will be monitored and managed, frequency of oversight, and by whom at your institution (e.g., titles, roles).

Primary oversight of this plan is provided by Dr. Tsongalis, Director of the Pathology Shared Resource. Additional oversight is provided by the Dartmouth Cancer Center's Associate Director for Shared Resources, Dr. Kolling.