Research Opportunity Announcement
OTA-21-015D
REsearching COVID to Enhance Recovery (RECOVER) Initiative:
Data Repositories

The NIH is soliciting applications for Data Repositories that will serve as integral partners with the PASC research community to rapidly and flexibly deploy, manage, and grow a robust, secure digital infrastructure that can meet near-term and long-term needs of the program. This Research Opportunity Announcement (ROA), OTA-21-015D, is for the following: **Digital Health Data Repository, Imaging Data Repository, Pathology Data Repository** and **Clinical and Observational Data Repository** components of the Initiative.

Introduction

Recovery from SARS-CoV-2 infection is extremely variable with many having a quick resolution of their symptoms, while for other individuals there are important post-acute sequelae. Reported symptoms among persons who have been infected with SARS-CoV-2 range from mild to incapacitating, may persist after recovery from acute disease, may involve multiple organs and systems, and can adversely affect overall quality of life. In some cases, the onset of Post-Acute Sequelae of COVID-19 (PASC) symptoms and findings may be reported during the acute infection and can continue after the resolution of the acute phase. The magnitude of the public health impact of these sequelae is currently unknown but likely to be large, given the numbers of individuals across the age spectrum who have been and will be infected with SARS-CoV-2. It is a public health priority that we better understand and develop strategies to prevent and treat the post-acute sequelae of SARS-CoV-2 infection (PASC) and that these strategies enable rapid innovation, evolution, and adaptation as more is learned about PASC and its potential impact on public health.

The goal of the trans-NIH RECOVER Initiative is to rapidly improve our understanding of the recovery after SARS-CoV-2 infection and to prevent and treat PASC. Toward these ends, the Initiative is designed to address these fundamental scientific questions:

- What is the clinical spectrum of and biology underlying recovery from acute SARS-CoV-2 infection over time?
- For those patients who do not fully recover, what is the incidence/prevalence, natural history, clinical spectrum, and underlying biology of this condition? Are there distinct phenotypes of patients who have prolonged symptoms or other sequelae?
- To what extent does SARS-CoV-2 infection initiate or promote the pathogenesis of conditions or findings that evolve over time to cause organ dysfunction or increase the risk of developing other disorders?

The Initiative is designed to be a collaborative and an inclusive approach for rapidly advancing our understanding of the recovery process and the epidemiology (including

incidence/prevalence) and natural history (including duration) of PASC. Studies conducted will characterize: the clinical spectrum of recovery from SARS-CoV-2 infection, including the subset of patients who have symptoms of disease beyond the standard course; the individual, clinical, and contextual factors that contribute to the duration, types of symptoms, and severity of disease; phenotypes of patients who have prolonged symptoms or other sequelae; the impact of treatments for acute COVID-19 or post-acute symptoms on the duration and severity of symptoms; and factors that impact the outcomes in patients infected by SARS-CoV-2.

At the heart of the Initiative is the rapid launch of the SARS-CoV-2 Recovery Cohort and SARS-CoV-2 Recovery Cohort Investigator Consortium.

The **SARS-CoV-2 Recovery Cohort** is a collaborative meta-cohort that will leverage ongoing fit-for-purpose cohorts, as well as new cohort studies, to chart the process of recovery in diverse adult and pediatric populations. This will include cohort studies of patients acutely infected with SARS-CoV-2 (acute cohort), as well as cohorts of persons suffering from post-acute symptoms (post-acute cohort), along with appropriate control participants. The RECOVER Initiative will emphasize inclusive participation and leverage a variety of clinical platforms, including large-scale electronic health record (EHR)/health systems-based cohort studies; large and long-standing longitudinal studies; COVID-19 clinical trials/networks; and COVID-19 clinics, registries, and observational studies. These will be augmented by utilization of mobile and digital health strategies for participant recruitment, data collection, and follow-up. These SARS-CoV-2 Recovery Cohort studies will characterize PASC symptoms and findings and their trajectory over time and across the lifespan. They will include investigator-initiated studies taking a variety of approaches to probe for evidence of tissue injury or organ system dysfunction or other conditions (e.g., immunologic, pulmonary, cardiac, neurologic/cognitive, metabolic, mental health). Some may focus on special populations including children, older adults, new mothers, or those with relevant comorbidities. Diversity in study populations will be necessary to generalize findings to the U.S. population affected by SARS-CoV-2 infection. Toward this end, RECOVER Initiative investigators are encouraged to collaborate where feasible with other relevant NIH initiatives (e.g., Rapid Acceleration of Diagnostics-Underserved Populations (RADx-UP), Community Engagement Alliance (CEAL) Against COVID-19 Disparities).

Given the heterogeneity of symptoms and findings involving multiple tissues and systems, understanding PASC will require a multidisciplinary approach. Toward this end, all study investigators under this initiative (OTA 21-015D) will work together in a **SARS-CoV-2 Recovery Cohort Investigator Consortium** with the goal of immediately launching a multi-disciplinary collaboration to conduct rapid systematic screening and follow-up evaluations of SARS-CoV-2 infected individuals, to provide a resource for in-depth multi-disciplinary phenotyping, and to pool data and share biospecimens and data across studies. After the award, Consortium investigators will be convened to rapidly develop a streamlined set of common core protocol elements (specific hypotheses, design elements, screening evaluations, exams, lab tests, functional assessments, imaging, digital health measures, patient-reported outcomes, etc.) and to provide a collaborative for multi-disciplinary phenotyping. Consortium investigators may also propose site- or study-specific hypotheses that, due to specific expertise or technology constraints, may only be possible in subsets of the collaborative as sub-studies or ancillary

studies. Successful applicants will be expected to participate in collaborative protocol development and implementation and responsible data sharing in a timely fashion. All data is expected to be shared under General Research Use.

Importantly, the Initiative also will leverage **EHR- and other Real-World Data-based Approaches** to provide data and information on the incidence/prevalence of post-acute sequelae, PASC symptoms, imaging, and lab test results to inform the definition of PASC; describe patient demographics; identify comorbidities; define health care utilization patterns; provide real world data for comparative effectiveness studies, as well as reducing time and scope of potential clinical trial design and implementation; and inform PASC clinical characterization through health systems-based patient surveys. (See https://www.fda.gov/science-research/science-and-research-special-topics/real-world-evidence for a description of real-world data.)

Exploratory clinical trials testing strategies to treat symptoms and prevent progression of SARS-CoV-2 infection to PASC are also a critical part of this initiative and will be the subject of subsequent solicitations.

Also, critically important to understanding the pathology associated with PASC will be **Autopsy Cohort Studies** that will include in-depth histopathologic analysis of brain and other organs and tissues to identify tissue injury due to SARS-CoV-2 infection and/or its sequelae that lead or contribute to PASC.

RECOVER Initiative Structure

The goals of this initiative will be realized by multiple components working closely together to support the three research study areas (SARS-Cov2-Recovery Cohort Studies, Autopsy Cohort Studies, and EHR- and Other Real-World Data-based Studies) described above. The components include but are not limited to:

- A **Clinical Science Core** to coordinate the SARS-CoV-2 Recovery Cohort investigator consortium; track regulatory documents, facilitate clinical protocol development, implementation, monitoring, and data analysis; foster the use of common data elements across groups; oversee collection and quality control of data; promote multi-disciplinary collaboration; and foster community engagement.
- A **Data Resource Core** to coordinate and oversee the PASC data management strategies and processes including implementation of data submission software, data harmonization and quality control, foster the use of common data elements across groups, monitor the PASC Biorepository to create uniform biospecimen reporting and data submission processes and workflows, provide analytical tools and statistical support to the Clinical Science Core and awardees, and to create an interoperable cloud-based data analytics platform (workbench) that brings together the PASC data repositories capabilities, unique assets, and PASC data. In this regard, the Data Resource Core (DRC) provides a seamless and interoperable data ecosystem for both the PASC investigators within the consortium as well as providing the larger research community

an interface to publicly available, controlled access data that is deposited within the PASC data repositories.

- A **PASC Biorepository Core** to receive, manage, and make available a diverse range of biospecimens derived from PASC Consortium studies.
- A **Mobile Health Platform** to develop customized mobile apps to collect PASC digital health measures in the Clinical Recovery Cohort studies.
- **Data Repositories** that**,** with the oversight and coordination of the DRC**,** leverage existing and well-established robust capabilities to harmonize, manage, deploy, and make PASC and related data available to the DRC and to the research community.
- An **Administrative Coordinating Center** to provide central coordination and oversight.

The Clinical Science Core, Data Resource Core, and Biorepository Core are the subjects of a separate but related ROA OTA-21-015A. The three research study areas (Clinical Recovery Cohort Studies, Autopsy Cohort Studies, and EHR- and Other Real-World Data-based Studies) are also the subject of a separate but related ROA OTA-21-015B. Applicants are strongly encouraged to review these related ROAs in detail and to be familiar with their contents. Additional ROAs may be issued in the future as needed.

Consortium investigators will be expected to develop, implement, and participate in a collaborative governance structure that includes community representatives and affected persons. All study investigators are expected to rapidly and appropriately share data and biospecimens and to consent participants for general research use of their data, other medical information, and biospecimens, as feasible.

This initiative supports NIH's longstanding commitment to making the results, underlying scientific data, and outputs of NIH-funded research available to the public through effective and efficient FAIR (Findable, Accessible, Interoperable, and Reusable) data sharing practices and policies. Consortium investigators will make research data and biospecimens available through the Clinical Science Core, Biorepository Core, and Data Resource Core at agreed upon milestones and upon completion of their study. Permission to access controlled-access PASC datasets that are released to the public will require approval through NIH Data Access Committees via the database of Genotypes and Phenotypes (dbGaP) Data Access Requests or a similar NIH-governed mechanism. Authorization based on these data access approvals will be managed throughout the PASC data ecosystem and through integration with the NIH Researcher Auth Service. Secondary users from the researcher community will agree not to distribute controlled-access datasets and will acknowledge use of PASC datasets through citations in manuscripts and presentations.

PASC Data Strategy

The overall vision for PASC data is a hub-and-spoke model as described below, with the Data Repositories serving as the spokes surrounding the central coordinating hub, which is the PASC Data Resource Core (DRC). The DRC, which was described in a previous ROA (see OTA-21-015A), will play a critically important and central role in coordinating and integrating data from the PASC Data Repositories (described below).

The DRC will serve as the data hub for the RECOVER Initiative. It will lead PASC-related data management and cross-initiative harmonization for greater interoperability, integration, and sharing by coordinating activities of the Data Repositories and fostering relationships and synergy between various repositories and other elements of the Consortium. The DRC will lead assessments and decision-making processes regarding adoption and implementation of data collection standards, a core set of PASC common data elements (CDEs) and tools, and harmonization procedures including mapping to common data models (e.g., OMOP), the Data Repositories, the Clinical Science Core, and PASC Investigator Consortium to maximize comparability across the program and measurement modes (e.g., web, mobile app, in-person) for longitudinal research and evaluation of program impact.

In addition, to increase maximal interoperability of the data resulting from PASC research, the DRC will assist PASC-supported research projects with data harmonization, as needed, and will oversee and coordinate data ingestion to the appropriate Data Repository. The DRC will also manage linking data across the supported Data Repositories and will provide data informatics tools to monitor investigator consortium progress and performance. The DRC will also oversee and facilitate the necessary mapping to common data models (e.g., OMOP), will provide any additional and overall data quality control, data curation, and analyses that will better prepare data across all sites for ingestion and integration into the PASC Data Repositories. Lastly, the DRC will oversee and lead an effort with the PASC Data Repositories to ensure all PASC data is indexed and findable across the landscape of repositories. It is expected that data generated from this initiative will be consented for General Research Use and shared in a timely fashion. To this end the DRC will provide a master index of all PASC data and metadata and maintain clear provenance of the data. By using the master index and search functions, researchers will be able to easily find and use relevant data **within and across the PASC Data Repositories.** In addition researchers will be able to **aggregate data and analytical tools across the PASC program through an interoperable DRC-led cloud-based workbench.** In this vision, data is made available in a manner that maximizes the number of different researcher use cases and enables researchers to find, access, and co-analyze multiple datasets across the PASC data repository ecosystem (Figure 1).

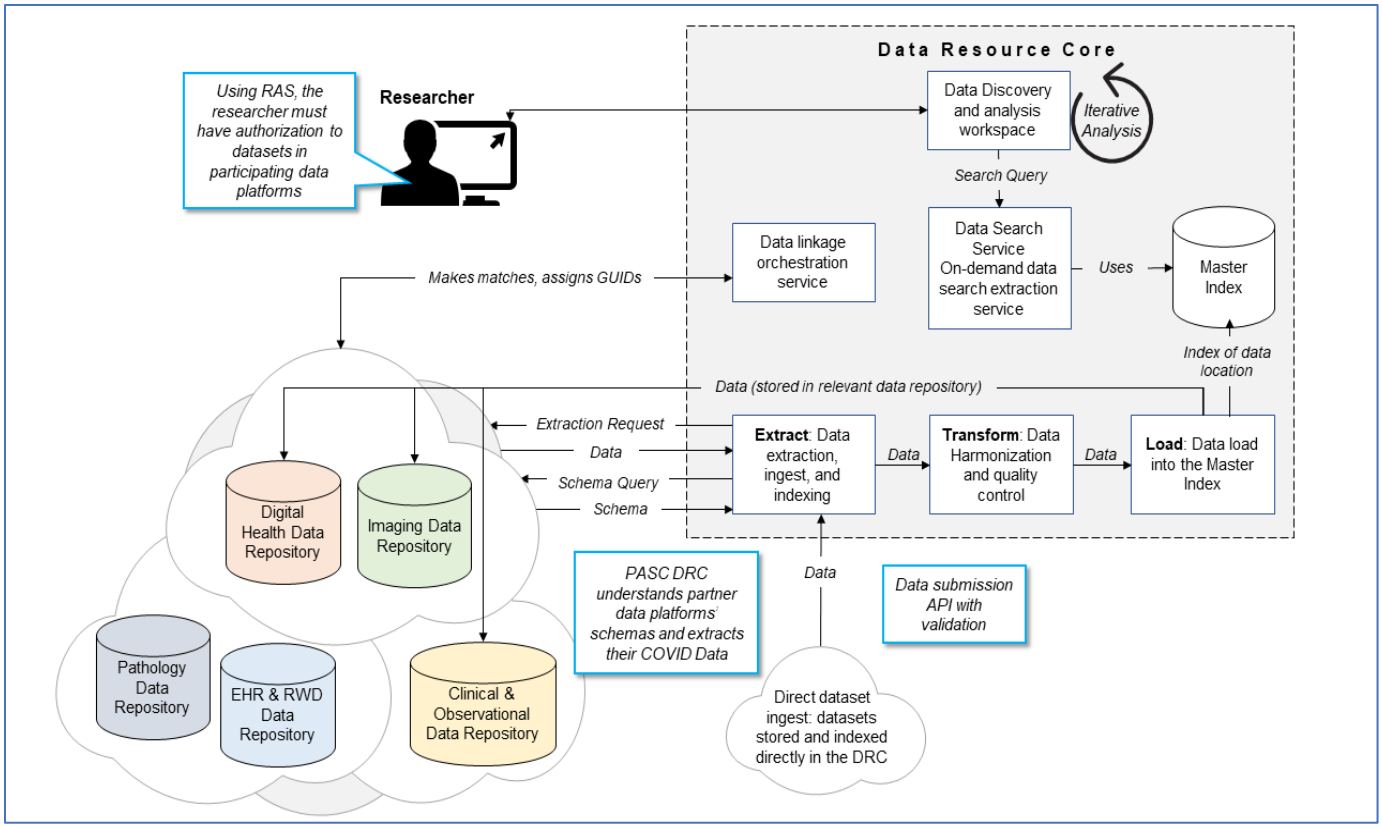


Figure 1. PASC Data Strategy

The flow of this diagram begins with the "**Researcher**" depicted in the top left. The researcher obtains access to the requisite data using the NIH **Researcher Authentication Service (RAS),** then accesses their own PASC DRC "**data discovery and analysis workspace**". The researcher can perform iterative analysis on PASC data within this workspace. The "**data discovery and analysis workspace**" uses software built by the PASC DRC to perform **search queries** for PASC data. The PASC DRC "**data search service**" and "**on-demand data search extraction service**" use a "**master index**" to find and acquire disparate PASC data. The purpose of the **master index** is to centrally record the location of PASC data that's spread across multiple repositories. The PASC DRC, as the "hub" in a "hub and spoke" model, continually receives and deidentifies new PASC data over time from clinical sites and other PASC supported efforts via an **API**. **PASC Data Repositories** may play a central role in data de-identification and harmonization. It is anticipated that the DRC may be required to perform an **Extract, Transform, Load (ETL) function** for any additional data harmonization, and will load data to the appropriate data repository. The DRC also records the index (i.e., the data's final location) in the PASC DRC master index. The PASC DRC may also routinely query **data repositories** to learn about new PASC data that is available and how that data is structured (i.e., its **schema**), as shown with the box labeled "**Extract: Data extraction, ingest, and indexing**". The "**master index**" is updated upon discovering new PASC data in a repository.

Objectives

With this ROA, NIH is soliciting applications for the following: **Digital Health Data Repository, Imaging Data Repository, Pathology Data Repository** and **Clinical and Observational Data Repository** for the RECOVER Initiative. NIH's vision for these data repositories is that they work closely together as integral partners with the PASC research communities to rapidly and flexibly deploy, manage and grow a robust, secure digital infrastructure that can meet near-term and

long-term needs of the program. In this vision, the PASC data repositories are the spokes in a hub and spoke model, with the PASC DRC acting as the coordination hub. The data repositories are expected to be capable of handling and harmonizing data and information, making data and its derivatives readily available to researchers, and having the appropriate security controls and procedures for secure data access and analysis.

In this model, the Data Resource Core serves as a data hub in the RECOVER Initiative andwill provide oversight and the appropriate expertise in PASC-related data collection, harmonization, integration, and sharing as well as proven relationships and expertise with clinical consortiums and networks and capacity. The DRC will work with the consortium to standardize collection of clinical data (i.e., incorporating new and reusing existing COVID-19 common data elements and case report forms) as well as oversee the incorporation of less structured data (i.e., imaging reports, results of specialized tests) to maximize comparability across datasets and measurement modes (e.g., web, mobile app, in-person) for longitudinal research and evaluation of program impact; assist consortium research projects with linking these data, via implementing GUIDS and a laboratory information management system (LIMS) for the corresponding biospecimens and data from other sources; oversee and perform higher levels of quality control, data curation, and analyses, and provide data informatics tools to monitor/track investigator consortium progress and performance; facilitate and oversee the necessary data harmonization, mapping to common data models (e.g., OMOP), and will be responsible for preparation of data for ingestion into the data repositories for secondary data reuse.

Importantly, to create maximal discoverability of PASC data and to enable a broad view and cross-repository search, the DRC will maintain a “master index” including the metadata of all PASC data held in the DRC and the repositories. The ability to aggregate research subjects across PASC will be further enabled through an implementation privacy-preserving Record Linkages (PPRL). This is complementary to the use of GUIDs and once a PPRL ‘match has been made’ a GUID can be assigned. In addition, unique dataset IDs may be provided by individual repositories but should be harmonized to a common GUID implementation by the DRC. These activities are intended to provide for cross repository search and enable construction of aggregated, cross-repository datasets for secondary analysis.

As spokes in this model, it is anticipated that the PASC Data Repositories will work with the DRC and provide capabilities in primary data harmonization, curation, quality control and data ingestion for secondary data reuse. For further interoperability, it is anticipated that the PASC Data Repositories will make available to the DRC their analytics tools, platforms, or workbenches for federated analysis. It is also anticipated that the DRC will work with the PASC repositories and create the necessary and sharable application programming interfaces (APIs), data aggregation applications, and workflows, that are easily deployable within the DRC or the PASC repositories in a secure and sharable cloud environment. In this way, data obtained from the RECOVER Initiative will be made easily integrated and FAIR regardless of where the data resides across PASC repositories. It is therefore important for awardees to nimbly adapt to the changing landscape of studies, measures, and the needs of this Initiative.

The following are requirements for all awardees of PACS Data Repositories. These requirements outline expected capabilities that would be in-place, or easily configured, within a Data Repository at the time of award.

Requirements Common to all Data Repositories

A. Data Curation, Metadata, Provenance and Sustainability
   1. Curate and map data to ensure datasets are accompanied by metadata and paradata to enable discovery, reuse, and citation of datasets while maintaining provenance of the data.
   2. Assure complete and accurate high-quality data with the necessary data harmonization procedures, QA/QC, data release expectations and timelines for data access to secondary data users, and adherence to any standards and common data models.
   3. Ensure all data received and stored are appropriately de-identified for PII and contains study- compliant required GUIDs and unique persistent identifier (PIDs), as assigned by the DRC. Additionally, Data Repositories will be required to implement privacy preserving record linkage ('hash') for PASC data.
   4. Provide clear capabilities and guidance on data access and data restrictions and processes for obtaining access to controlled-access data.
   5. Provide tools to track the provenance of all ingested data and metadata. Record the origin, version, chain of custody, and any modifications to submitted datasets and metadata.
   6. Develop QA/QC metrics and perform QA/QC on all ingested data to ensure the accuracy and integrity of datasets and metadata.
   7. Assign well-defined datasets a citable, (PID), such as a digital object identifier (DOI) or accession number.
   8. Provide long-term management of data, including maintaining integrity, authenticity, and availability of datasets. Provide documentation on policies for data retention.

B. Timely Data Ingestion and Sharing
   1. Work with the Data Resource Core for data ingestion.
   2. Support secure data ingestion through an Application Programming Interface (API), graphical interface, and command-line interfaces.
   3. Where appropriate, ensure that data submissions are clearly linked to appropriate consent groups and that associated Data Use Limitations are submitted and registered to a central controlled-access authorization authority (e.g., dbGaP) based on Institutional Certifications or other data sharing documentation.
   4. Work with the Data Resource Core and the Clinical Science Core to complete the appropriate data transfer agreements, institutional certifications, and/or other data sharing documentation and steps required for researchers to submit and share data through the repository (e.g., information required for dbGaP registration).
   5. Transparently provide citation and data use information for each dataset.

6. Ensure that data access and use is consistent with consortium agreements and/or NIH Data Access Committee approvals throughout the entire environment via integration with Researcher Auth Services (RAS).
7. Provide expertise in data systems to allow timely availability and/or transfer of data, management, quality control, and export to the DRC as needed for analyses and integration.
8. Accelerate PASC research through broad and rapid data sharing practices, including iterative and regular data releases to the Investigator Consortium and the wider research community.
9. Provide investigator support for data submission, including capturing the appropriate consent groups for secondary use (as described above), including data submission templates, instruction manuals, help desk support, application programming interfaces, and graphical and command line data submission tools. All tools should follow the least permissions model.

C. Data Management and Linkage
1. In collaboration with the Data Resource Core, support a data management model that will maintain the link between de-identified data submitted to different repositories (through directly assigned GUID at DRC level and additionally via third party Privacy Preserving Record Linkage), in support of interoperability across PASC repositories and coordinate linkages for a larger NIH data integration model.
2. Employ an established and transparent process for recording and tracking use of data to measure impact.
3. Collaborate with the Data Resource Core to support additional capacity to link data with PASC Biorepository Core, and from other sources, providing data informatics tools to monitor investigator consortium progress.

D. Collaboration
1. Support the PASC Clinical Science Core and the PASC Investigator Consortium in establishing a standardized and/or harmonized set of health measures for assessing the trajectory of acute COVID-19 and PASC over time.
2. Collaborate with the Data Resource Core to support scientific collaboration and data sharing with consortium members and secondary data users, enabling cross-Consortium collaboration and research by the wider research community through easy access and findability of data and rapid dissemination of findings.
3. Maximize the value of the collective data sets and enable interoperability with other NIH COVID-19 research resources by supporting studies and working with the PASC Data Resource Core (DRC) to make relevant data accessible to a cloud-based data "workbench." The DRC will oversee and lead on the interoperability efforts and it is expected that funded repositories will participate in the development of standardized APIs, indexing and search, and compatible and sharable analytical tools, workbenches, etc.

E. Tools and Resources

1. Provide search, visualization, and analytical tools and resources that address the needs of a wide variety of researchers with different expertise ranging from areas of computer science and bioinformatics but also clinical or experimentally-oriented user communities, such as biologists and healthcare practitioners. Interoperable tools should be designed, in collaboration with the DRC-led efforts, to integrate or federate with other tools across PASC repositories to create a seamless experience and enable co-analysis of multiple data types.
2. Provide an analytical environment where researchers can securely analyze multiple datasets, including the ability to upload their own data for co-analysis.
3. Provide a near real-time dashboard that allows for an easy access pathway for simple queries and key statistics of the data in the repository, while avoiding risk of re-identification.
4. Support "Bring your analytic" for external data analysis approaches.

F. Cloud and Technology
   1. Leverage cloud-based technology, where appropriate, to bring resources and tools to the data and enable complex and large-scale analyses of datasets across repositories.
   2. Foster re-use and reproducibility by developing and using open-source technology and architecture to the extent practicable. It is expected that all aspects of the repository infrastructure, including code for portals, tools, workflows, and/or pipelines created or used with support from this opportunity will be well-documented and shared with the wider scientific community, in a timely manner that would enable other researchers to replicate and build on for future research efforts, aligned with the open-source regime used to distribute software, where possible.
   3. Maintain a system architecture based on current cloud best-practices, including but not limited to microservices, purpose-built databases instead of monoliths, and least-permissions.

G. Documentation
   1. Develop tutorials, how-to guides, and other user-friendly documentation and resources to facilitate use of the repository and associated data and tools.
   2. Provide user-friendly documentation and examples for use of the data platform environment.

H. Privacy and security
   1. Have the appropriate security protocols and processes including repository managed systems can obtain and maintain an Authority to Operate (ATO) compliant with applicable regulations and standards, including Federal Information Security Management Act (FISMA) at the Moderate level and including cloud specific controls where appropriate.
   2. Adhere to HHS Information Security Program to ensure the integrity, confidentiality, and availability of Federal Information and the Federal Information system (FISMA moderate) including FedRAMP compliance.

3. Comply with all relevant Federal cyber security and privacy requirements including but not limited to FISMA, FedRAMP, and the Privacy Act and relevant National Institute of Standards and Technology (NIST) publications and guidance.
4. Comply with data protection and anonymization requirements of the RECOVER Initiative.
5. Prepare all required Security Assessment and Authorization (SA&A) documentation based on NIST guidance and gain Authority to Operate (ATO) approval from the appropriate Designated Authorizing Authority.
6. Integrate with the NIH Researcher Auth Service for authentication and authorization of controlled access data based on approvals from NIH Data Access Committees.
7. Implement contingency procedures for incidents and breaches, including audit and logging capabilities.

The following section outlines the requirements for awardees of the individual PASC Data Repositories.

Digital Health Data Repository

The NIH is soliciting proposals for adaptation and support of a scalable, configurable, and integrated Digital Health Data Repository to facilitate the collection, annotation, harmonization, curation, and sharing of digital health data collected via mobile apps and/or sensors by the PASC Investigator Consortium to augment existing clinical, EHR, and other real-world data. The Digital Health Data Repository will work collaboratively with the Mobile Health Platform to enable collection of PASC data from SARS-CoV-2 Recovery Cohort participants. Applicants are strongly encouraged to review the Mobile Health Platform ROA in detail and to be familiar with its contents (OTA-21-015C). The Digital Health Data Repository proposed for this initiative should have demonstrated experience and capabilities to support the following features. Applicants without existing capabilities in one or more listed areas below should describe how they will rapidly incorporate these capabilities into their platform.

The Digital Health Data Repository (DHDR) will include the following functionalities:

1) Data ingestion, Curation and Storage
    a. Provide integration capabilities to receive digital health data from SARS-CoV-2 Recovery Cohort studies that use the Mobile Health Platform app or their own app, potentially including development of an API to receive data if needed.
        i. DHDR staff will collaborate with staff from each Recovery Cohort study to ensure that all digital health data will be findable and accessible through the DHDR.
    b. Work with investigators from the Mobile Health Platform and Recovery Cohort studies to facilitate extraction and sharing of de-identified output measures (e.g., GPS-derived measures).
    c. Collaborate with the PASC Investigator Consortium (especially the Data Resource Core and Mobile Health Platform) to define a common data model (CDM), common data elements (CDEs), and minimum metadata and paradata standards for PASC

digital health data, including questionnaires, sensor data, and other relevant data types.

d. Leverage approaches or tools for data fusion to harmonize sensor data and health measures collected from different devices. This could, for example, enable comparison of heart rates collected from two different wearable devices at two different sampling rates.
e. Ensure the data are cleaned and usable by addressing issues of signal quality, missing data, erroneous recordings, etc., in all sensor data (data-janitoring).

2) Research data portal:
   a) Provide a user-experience-focused web application that enables the research community to browse, search, and access the de-identified data, metadata, paradata, and algorithms hosted and cataloged by the DHDR.
   b) Provide a standards-based API to provide access to the capabilities available within the web application, including enabling federated search of data across repositories (see discussion of the DRC Master Index above).

3) Support for analysis:
   a) Provide the ability to ingest, host, and deploy computational pipelines, workflows, and analyses over the data in the DHDR.
   b) Provide a core set of standard analysis tools, such as RStudio, SAS, Jupyter Notebooks, Galaxy, and/or other widely used analysis tools.
   c) Enable data hosted within the DHDR to be aggregated for analysis.
   d) Provide a means for researchers to share their own computational models, algorithms, and analysis results.
   e) Provide tools to track the provenance of algorithms, computational pipelines, and analysis results.
   f) If necessary, support a secure, collaborative sandbox environment that enables researchers to store, manage, compute on, and share their own data, analysis tools, and results—including the ability to combine their data with data hosted by the DHDR, as appropriate.

4) RECOVER Initiative engagement:
   a) Support the PASC Clinical Science Core and the PASC Consortium in establishing a standardized and/or harmonized set of digital health measures for assessing the trajectory of acute COVID-19 and PASC over time—including the data elements, temporality of assessment, and on-screen display of questions. These measures should include:
      - Core questions about the symptoms experienced by patients to chart recovery or worsening over time in symptoms and quality of life.
      - Sensor data from consumer wearable devices that can capture data and specific symptoms or clusters of symptoms that patients are experiencing (e.g., heart rate, respiration rate, blood oxygen, activity levels, cough, sleep patterns, actigraphy, temperature).

- Geospatial data, which will enable patient data to be linked with environmental and social context data (e.g., weather data, socioeconomic deprivation indices).
- Integration of data from COVID-19 testing, including serial at-home COVID-19 antigen testing to monitor possible reinfection.

Clinical Data Repository

The NIH is soliciting proposals for adaptation and support of a scalable, configurable, and integrated, cloud based Clinical Data Repository to facilitate the collection, annotation, harmonization, curation, and sharing of clinical data generated by structured observational studies and clinical trials including but not limited to clinical observations, lab result data, demographics, social determinants of health, genomic and other –omics data. The Clinical Data Repository proposed for this initiative should have demonstrated experience and capabilities to support the following features. Applicants without existing capabilities in one or more listed areas below should describe how they will rapidly incorporate these capabilities into their platform.

1) Data Ingestion, Curation and Storage:
   a) Collaborate with the PASC Data Resource Core to support scientific collaboration and common data elements, data harmonization, and data sharing, enabling cross-Consortium collaboration for easy access and findability of data and for rapid dissemination of findings.
   b) Utilize Common Data Elements (CDEs) as promulgated by the Data Resource Core and the Clinical Science Core. To the extent possible, the Clinical Data Repository should leverage CDEs generated through RADx and other COVID initiatives.

2) Research data portal:
   a) Provide a user-experience-focused web application that enables the research community to browse, search, and access the de-identified data, metadata, paradata, and algorithms hosted and cataloged by the DRC.
   b) Provide a standards-based API to provide access to the capabilities available within the web application, including enabling federated search of data across repositories (see discussion of the DRC Master Index above).

3) Support for Analysis:
   a) Provide standards-based APIs for access to data that respects authentication and authorization via the NIH Researcher Auth Service (RAS).
   b) Provide a library of workflows and analysis tools for common analytic techniques.
   c) Provide integration with data in other data repositories to enable multimodal analysis of data with data linkage based on GUIDs and PPRL, as appropriate (see discussion of the DRC Master Index above).
   d) Provide access to clinical, genotypic and -omics data using APIs that respect data access approvals.
   e) Provide collaborative workspaces for researchers that can collect data or pointers to data from across repositories for analysis using a variety of tools such as SAS, R,

Python, Jupyter notebooks, common –omic workflows such as GWAS and PheWAS, machine learning models, etc.

- Workspaces will ensure that all collaborators in the workspace have access based on approved Data Access Requests through integration with RAS.
- Workspaces will support reproducibility of analysis conducted within them while protecting access to data based on consents and associated NIH Data Access Committee approvals.
- Workspaces and their associated data sets will support the ability to mint a permanent unique identifier such as a DOI that can be used in citations and enable an interested party to rapidly reconstruct the analysis including data sets, analysis tools, workflows and processes including the versions of the above used in the analysis.
- Workspaces must support "ephemerality," i.e., that after an analysis has been completed, the workspace can be destroyed leaving only a secure archive suitable for reproducibility as described above.

Imaging Data Repository

The NIH is soliciting proposals aimed at leveraging, adapting, optimizing, expanding, and supporting a scalable, configurable, and integrated Imaging Data Repository to facilitate the collection, annotation, harmonization, curation, storage, and sharing of digital human imaging data across the data lifecycle. The repository should be capable of accommodating a diversity of image types and formats from diverse modalities (including, but not limited to radiology (radiography, computed tomography, PET, and magnetic resonance imaging), from a diversity of organs and tissues (including, but not limited to the cardiopulmonary, neurological, or abdominal systems) that will be collected by the SARS-CoV-2 Recovery Cohort studies and other PASC-related initiatives as appropriate. These images provide critical integration between observed and objectively measured clinical phenotypes.

Applicants for the solicited Imaging Data Repository should have demonstrated experience and capabilities to support the following features. Applicants without existing capabilities in one or more listed areas below should describe how they will rapidly incorporate these capabilities into their platform before data collection.

The Imaging Data Repository will include the following functionalities:

1) Data Ingestion, Curation and Storage
    a) Maintain and document an enterprise information management strategy for all types of imaging data, to ensure data integrity, availability, and confidentiality and to support dataset versioning and provenance.
    b) Provide for data de-identification, adhering to community standards and best-in-class approaches to identify and remove personal identifiable information, including but not limited to de-identification at image and image metadata level, and appropriate processes required to mitigate re-identification (for example, through facial reconstruction).

c) Ingest, curate, perform QA/QC, aggregate, and securely store the digital images generated using a wide range of imaging modalities and systems being collected by the SARS-CoV-2 Recovery Cohort studies and other PASC-related initiatives as appropriate, as well as appropriate comparison data.
d) Employ an extensible and versatile imaging data dictionary consistent with standards and best practices in the scientific community, to ensure that imaging data collected from different imaging modalities and different organ systems being collected by PASC studies and other cohorts can be co-analyzed and to provide appropriate context to secondary data users with respect to the participant informed consents through integration with Researcher Auth Services (RAS).
e) In addition to widely used formats such as DICOM, the Imaging Data Repository should provide support for submission and analysis of additional formats including but not limited to Brain Imaging Data Structure (BIDS)-formatted data (https://bids.neuroimaging.io/).
f) Facilitate the standardization and/or harmonization (i.e., usage of common data elements) of RECOVER Initiative imaging health data, consistent with standards and practices in the scientific community.
g) Provide means to link the imaging data to the other data collected by the clinical cohort studies, as appropriate, through GUID and PPRL, under DRC guidance.

2) Research data portal
   a) Provide a user-experience-focused web application that enables the research community to browse, search, and access the de-identified data, metadata, paradata, and algorithms hosted and cataloged by the DRC.
   b) Provide a standards-based API to provide access to the capabilities available within the web application, including enabling federated search of data across o repositories (see discussion of the DRC Master Index above).

3) Support for Analysis
   a) In addition to the search and cohort building resources common to all repositories, provide imaging-specific software tools such as DICOM viewers, tools for annotating images, imaging analysis workflows, as well as well-developed data science AI, and ML methods that facilitate automated and semi-automated image pre-processing, all in an interoperable manner with the DRC workbench.
   b) Provide investigator support for data analysis, including instruction manuals, help desk support, data model documentation, application programming interfaces, and custom cloud-based workspaces in response to researcher requirements. The approach to analysis tool support should start with maximum flexibility and support development and implementation of innovative image processing and analysis tools in support of specific use cases/workflows.
   c) Enable the PASC Consortium and the greater research community to access imaging data through a secure, interoperable portal, including capabilities for data analytics that can be securely deployed in or federated with the DRC's workbench.

d) Ensure that the repository's database design and implementation can facilitate performant querying of imaging datasets and 99% availability for repository web applications, including but not limited to search tools and analysis tools.
e) Organize structured data and data objects in a way that mirrors data access permissions so that browsing and quick previews are possible.

Pathology Data Repository

The NIH is soliciting proposals for adaptation and support of a scalable, configurable, and integrated Pathology Data Repository. This repository will facilitate support of the collection, annotation, harmonization, curation, and sharing of pathology imaging data in the autopsy cohort as well as in the clinical cohorts. The scope of the data collected by the SARS-CoV-2 Recovery Cohort studies and other PASC-related studies will include but not be limited to histopathology, whole slide imaging (WSI), immunofluorescence, multiplex Immunohistochemistry (mIHC), and electron microscopy (EM). These images provide critical integration between observed and measured clinical phenotypes with the opportunity to investigate tissue specimens for evidence of dysfunction at the organ, tissue, cell, and molecular levels.

Applicants for the solicited Pathology Data Repository should have demonstrated experience and capabilities to support the following features. Applicants without existing capabilities in one or more listed areas below should describe how they will rapidly incorporate these capabilities into their platform before data collection.

1) Data Ingestion, Curation and Storage
   a) Maintain and document an enterprise information management strategy for all types of histopathology imaging data, to ensure data integrity, availability, and confidentiality and to support dataset versioning and provenance.
   b) Provide for data de-identification, adhering to community standards and best-in-class approaches to identify and remove personal identifiable information, including but not limited to de-identification at image and image metadata level, and appropriate processes required to mitigate re-identification.
   c) Ingest, curate, perform QA/QC, aggregate, and securely store the digital images generated using a pathology imaging modalities and systems being collected by the SARS-CoV-2 Autopsy studies.
   d) Support a diversity of image formats, including the common WSI formats, as well as support submission of EM, immunofluorescence and other image types obtained from human samples.
   e) Provide, in conjunction with the DRC, linkage between the pathology data and other data collected by the clinical cohort studies, through PPRL and GUIDs.

2) Research data portal:
   a) Provide a user-experience-focused web application that enables the research community to browse, search, and access the de-identified data, metadata, paradata, and algorithms hosted and cataloged by the DRC.

b) Provide a standards-based API to provide access to the capabilities available within the web application, including enabling federated search of data across repositories (see discussion of the DRC Master Index above).

3) Support for Analysis
   a) Provide investigator support for data analysis, including instruction manuals, data model documentation, application programming interfaces, and custom cloud-based workspaces in response to researcher requirements.
   b) Enable the PASC Consortium and greater research community to access pathology imaging data and associated electronic reports through a secure, interoperable portal, including capabilities for data analytics that can be securely deployed in or federated with the DRC's workbench.
   c) Ensure that the repository's database design and implementation can facilitate performant querying of imaging datasets and 99% availability for repository web applications, including but not limited to search tools and analysis tools.

Special Award Terms

The complete terms and conditions of each OT Agreement or sub-agreement issued under this ROA are subject to negotiation and will be contained in the Agreement entered between the NIH and the Awardee. This Special Award Terms section is provided for informational purposes only to provide prospective applicants with an understanding of key expectations and terms that may differ from traditional NIH award mechanisms.

Lower Tier Agreements

With mutual consent of the Awardee and the NIH, the Data Resource Core will be expected to issue sub-awards to entities identified and approved by the NIH under this ROA associated with the RECOVER Initiative.

Negotiation

The NIH reserves the right to:

- Select for negotiation all, some, one, or none of the proposals received in response to this ROA;
- Segregate portions of resulting awards into components and their associated budget and/or milestones that differ from those that have been proposed;
- Accept proposals in their entirety or to select only portions of proposals for award;
- Fund projects in increments and/or with options for continued work at the end of one or more phases, which can consist of more than one milestone;
- Fund projects of two or more applicant entities as part of a reorganized, consolidated consortium operating under an article of collaboration, teaming arrangement, or other means acceptable to the NIH;
- Fund proposers as sub-awardees of a separate Coordinating Center entity to be established by the NIH;
- Request additional documentation (certifications, etc.); and

- Remove proposers from award consideration should the parties fail to reach a finalized, fully executed agreement, or the proposer fails to provide requested additional information in a timely manner.

Software sustainability and transition to another awardee

- A fundamental objective of this agreement is to ensure that the valuable data, tools, and resources managed and created by the PASC Data Repositories remain available without interruption to the research community in cases where the award is terminated by the NIH or if the awardee withdraws or otherwise can no longer manage the resources. Consistent with 45 C.F.R. 75.322, NIH will have unrestricted cost-free access and use of all such resources, including the right to transfer said resources to other NIH-funded and/or managed resource projects, at the NIH's sole discretion upon termination or expiration of this agreement.
- Open Source Technology: Capabilities and software developed as part of the PASC Data Repositories should be delivered under an open source model where practical. The repositories may use proprietary or commercial services, tools, and platforms so long as the requirements for data transparency and interoperability are maintained.
- At any time, NIH may request a detailed plan for transferring resources (e.g., portal, databases, software and other code) to NIH or another entity designated by NIH.

Data Governance

- In addition to enabling pre-release consortium access to the data, permission to access controlled-access PASC datasets that are released to the public will require approval through NIH Data Access Committees via dbGaP Data Access Requests or a similar NIH-governed mechanism.
- All data submitted and shared through the PASC Data Repositories are the property of the NIH and cannot be sold, redistributed, or otherwise transferred without the authorization of the NIH.

Authority

This Research Opportunity Announcement (ROA) is issued with the goal of establishing an "other transactions" agreement or sub-agreement pursuant to 42 U.S.C. § 282(n).

Eligibility

The following entities are eligible to apply under this ROA:

Higher Education Institutions

- Public/State Controlled Institutions of Higher Education
- Private Institutions of Higher Education

Nonprofits Other Than Institutions of Higher Education

- Nonprofits with 501(c)(3) IRS Status (Other than Institutions of Higher Education)
- Nonprofits without 501(c)(3) IRS Status (Other than Institutions of Higher Education)

For-Profit Organizations

- Small Businesses

- For-Profit Organizations (Other than Small Businesses)

Other

- Federally Funded Research and Development Centers (FFRDC) and University Affiliated Research Centers (UARC) are eligible to apply and/or participate as partnering organizations subject to any agency sponsor related requirements and restrictions on eligibility.

Submission and Contact Information

Interested applicants should send a letter of intent (1-2 page) outlining:

- How the proposed solution meets the needs of the PASC research communities with respect to fulfilling the requirements expected of the PASC Data Repository.
  - Applicants are expected to illustrate a commitment to PASC RECOVER vision and describe anticipated workflows and interactions with Biostatistics Center at Massachusetts General Hospital as they assume the role of the Data Resource Core

- The team's prior experience deploying their solution in a variety of research contexts and experience working as part of a multi-institutional collaboration.
  - Applicants are expected to provide 2-3 experiences in a collaborative network, serving as a data resource or data platform and community engagement resource and demonstrates ability to respond and act on community-driven feedback

- Technical proficiency with data management, privacy, security, accessibility, interoperability, providing a researcher platform for analytics, and adopting Consortium standards.
  - Applicants are expected to describe data ingestion, curation and storage in the Gigabyte to Petabyte range and describe their QA/QC capabilities.
  - Applicants are expected to describe their analytical capabilities, including ability to develop and deploy containerized workflows, standard APIs, platforms in the cloud and/or on-prem; and userbase that demonstrates ability to serve as a repository
  - Applicants are expected to describe de-identification capabilities, harmonization capabilities, and use of CDEs and/or other standards.
  - Applicants should acknowledge their ability to have or obtain an ATO, and adherence to NIST security standards
  - Applicants should acknowledge that they have a disaster recovery plan, transition plans and manual of operations

Letters of intent should be submitted to NHLBI_OTA@mail.nih.gov by **5pm EDT on July 16, 2021.**

If the Letter of Intent is favorably reviewed, the NHLBI OT Office will invite you to submit via eRA ASSIST a full proposal consistent with the format and requirements below. The OT office will provide instructions for submission at that time.

Inquiries can be submitted to NHLBI_OTA@mail.nih.gov, with financial and administrative questions addressed to Kevin Heath, NHLBI Agreements Officer and technical questions addressed to Audie Atienza, Scientific Program Lead. Prospective applicants are encouraged to contact RECOVER program officials before the application deadline, via the email address above, to ensure that proposals meet the goals of the Initiative.

Applicants are encouraged to register for the Technical Assistance Webinar to be held on July 15, 2021, at 5PM ET.

Full Proposal Format and Requirements

The full proposal should clearly and fully demonstrate the proposer's capabilities, knowledge, and experience. Applicants should provide a budget assuming an award term of four years to be funded on an annual basis. Proposals should include a Cover Page, an Executive Summary, a Timeline and Tasks for the first year, a Project Plan, and a Budget.

The Cover Page (1 page max) should include:

A. The proposal title. **Please ensure to reference the ROA# OTA-21-015D in the title of the proposal to ensure expedited processing.**
B. The applicant's:
   i) Legal entity name
   ii) Address and contact information
   iii) SAM # and expiration date
   iv) DUN # and expiration date
   v) EIN number
C. The name and contact information for the applicant's Principal Investigator (with eRA Commons account information)
D. List of key personnel with titles and affiliations
E. The name and contact information for the applicant's Business Official, the person authorized to negotiate and bind the applicant as a signatory to the Other Transaction agreement
F. The total cost proposed for each year

The Executive Summary (1 page max) should describe:

A. How the proposed solution meets the needs of the PASC Consortium with respect to fulfilling the requirements expected of the PASC Data Repository
B. The team's prior experience deploying their solution in a variety of research contexts and experience working as part of a multi-institutional collaboration

C. Technical proficiency with data privacy, security, accessibility, interoperability, and adopting Consortium standards

The Timeline and Tasks for the first year (1 page max) should address the following:

The applicant should provide a summary of major tasks to be accomplished in the first 12 months of the proposed project. Each major task should include:

- Timeline
- Milestones and benchmarks
- Deliverables
- Budget for the task (the sum of all the task budgets should match the total budget being requested for the first year)

The Project Plan (10 pages max) should include:

A. Technical Approach

The proposal should describe how the applicant will meet the specific responsibilities of the PASC Data Repository of interest section of this ROA. Applicants can apply for more than one PASC Data Repository, but the application will need to have clear sections for each data repository. All applicants must address each of the requirements in the section for **all PASC Data Repository applicants.** The proposer must demonstrate its understanding of the RECOVER Initiative and component(s) being proposed by clearly showing a grasp of the range and the complexity of the work. This section should include a detailed project plan that includes milestones and deliverables for each phase of the implementation, with a particular focus on early steps needed to rapidly stand-up key functions and capabilities. Proposers should demonstrate a conceptual understanding of the challenges specific to the tasks required in the ROA and suggestions for overcoming these. Applicants must address specific processes and procedures for how they will achieve the required integration with the other components and for resolving any areas of disagreement. Applicants should propose potential technical approaches for federating with other data systems and analytical tools across PASC, including the workbench that will be managed by the DRC, and the other Data Repositories.

B. Key Personnel Experience

Proposers should demonstrate experience of key personnel supporting the planning and implementation of activities described in the ROA. Please provide resumes describing key staff who will be assigned to manage performance and supervise the work for each task and subtask (as appropriate). These resumes will be reviewed to evaluate whether the individuals possess the required experience to perform the specific tasks. Resumes should be no more than three (3) pages in length and will not count toward the page limits.

C. Management/Staffing Plan

Proposals should detail how the applicant will provide the necessary project administration, organization, and staff to ensure quality control, compliance with ROA expectations, and necessary staffing adjustments. In addition, proposers must demonstrate the ability to simultaneously manage multiple tasks within set time periods.

D. Past Experience

Proposers should provide multiple examples of prior project experience serving as a data repository, as described in this ROA. Each example should include the total funding awarded and dates of award, contact information for a sponsor able to serve as a reference, and a brief description of the project itself, including how the project was analogous to the needs identified in this ROA with respect to the work being proposed. Applicants will need to demonstrate prior work with clinical consortia or networks AND competency associated with the work being proposed.

The Budget should address the following:

The Budget section must provide a realistic, fully justified annual budget and cost proposal for performing the work specified in the ROA over a period of 4 years. NIH prefers applicants use an SF424 template to complete the categorical budget, but this is not required. Budget information and any related administrative documentation shall not count toward the total proposal page limit.

The Budget should provide the overall expected cost for each of the following categories:

- Personnel
- Materials and supplies
- Equipment
- Travel
- Subawards/subcontracts/consultants
- Other direct costs
- Total cost (with indirect costs included)
- Proposed Cost Share contribution

Objective Review Process

Applications will be evaluated in two stages. The first stage will be based on technical evaluation of the written proposals and will be done by a committee of subject matter experts. In the second stage, a subset of applicants will be invited to participate in an interview. The purpose of the interview stage is to clarify details from the written application. The interviews will be conducted via a videoconferencing platform. NIH will not support any costs associated with these interviews.

All applications will be evaluated using the following criteria:

1. **Merit of Scientific, Technical, and Management Strategy**
   - Will the data repository be configurable to support multiple, concurrent cohort studies with different populations and outcomes supported by the RECOVER Initiative? Is the plan to handle data curation, QA/QC, including metadata and provenance, and sustainability adequate?
   - Are the approaches proposed, including demonstrated flexibility, likely to lead to a viable solution for the long-term vision of the RECOVER Initiative?

- Does the platform have the capability to support features as necessary including electronic consent management, data harmonization and integration? Does the repository have the capability to de-identify data and implement GUIDs in a consistent manner?
- Will the platform support multiple interfaces and configurations to enhance usability, accessibility, and engagement of the study participants?
- Is the plan to work with the DRC adequate? Are there standard APIs and containerized workflows?
- Will the infrastructure for the platform be operational shortly after the time of award? Is their security and system architecture plan adequate?
- Are there plans and adequate resources to collaborate with staff on study design, provide training and troubleshooting assistance? Are there adequate plans to generate "Information for Use" materials that will engage the diverse range of participants in the RECOVER Initiative?
- Are the proposed plans, methods, techniques, and procedures for the project sound, feasible, and valid?
- Does the plan propose a solution or have the flexibility to leverage commonly used measures, data and metadata standards, models and schemas, and propose to implement an approach that has a high level of interoperability, security, and privacy? Will data storage and transmission be compliant with government standards?
- Are anticipated risks and challenges associated with the approach adequately addressed?
- Is the timeline of deliverables realistic and feasible?

2. **Past Experience**
   - Do the PI and key personnel have a strong track-record of managing, integrating, or analyzing data as part of a consortium?
   - Is there strong evidence of the team's ability to integrate diverse sources of information, build high impact collaborations to that rapidly agree on and implement Consortium policies, and make a sustained contribution to a high-profile, complex program?
   - Does the past experience of the team demonstrate a capability to rapidly stand-up collaboration with other institutions and successfully complete studies according to timeline and study objectives?
   - Are there clear examples of how the team have validated and implemented a data repository and/or platform rapidly and demonstrated continued development that is responsive to the needs of a wide range of studies?

3. **Key Personnel and Staffing Plan**
   - Does the performance and duties of the PI demonstrate strong qualifications to lead the data repository?
   - Is the necessary expertise committed to the degree required for the research project to be successful as an Other Transaction award?
   - Will the proposed work be managed effectively, achieve its goals, and be flexible to evolving needs as an Other Transaction award?

4. **Consortium Engagement Plan**
    - To what degree do the proposed plans address the data management, consortium coordination, and community engagement needs of this ROA?
    - Is the proposed design and technical and engagement approaches likely to lead to effective harmonization of solutions, tools, and products by the other PASC components?
    - Are there sufficient resources, personnel time, and flexibility to efficiently work with the PASC Data Repository and the Data Resource Center?

The following will be considered when making funding decisions:

- Scientific and technical merit of as determined by objective review
- Availability of funds
- Relevance and complementarity of the application to RECOVER Initiative priorities
- Evidence that the applicants are committed to goals and policies of the Initiative including with regard to confidentiality, publications, sharing of information and resources, and collaboration
- Evidence of previous productive, cooperative, collaborative technology development taking into consideration the needs of end users
- Evidence that the application will contribute to the diversity of technical and intellectual approaches and to the overall goals of the Initiative.

Non-Responsive Applications

An application will be considered non-responsive to this ROA if it: 1) proposes a generic data repository or platform without consideration of specifically how the technology will be integrated with the PASC Recovery Cohort Studies, relevant data repositories, Clinical Science Core, and Data Resource Core; or 2) proposes a data repository that does not meet the minimum privacy and security standards specified

A note about eRA Registration

NIH uses the eRA Commons system to administer OT awards. If you are selected to participate you may need to submit additional information in eRA ASSIST and will need to be registered in eRA Commons, which can take some time to process – as many as several weeks in some cases. Therefore, if you are considering submitting a proposal and are not yet registered in eRA, it is highly recommended that you begin the process of registering your organization, Program Director/Principal Investigator (PD/PI) and Signing Official (SO) in eRA Commons as soon as possible to avoid possible award processing delays. To register, please follow the instructions via this website:
https://public.era.nih.gov/commons/public/registration/registrationInstructions.jsp.

1. Complete the online Institution Registration Form and click Submit.
2. The NIH database will send you an email with the link to confirm your email address.
3. Once your email address is verified, the NIH will review your request and let you know of the result via email.

4. If your request is denied, you will get an email notifying you of the reason.
5. If your request is approved, you will get an email with your Commons User ID and temporary password.
6. Log into Commons with the temporary password and the system will prompt you to change temporary password to a permanent one. Your SO will be prompted to electronically sign your registration request. (Please review your registration information carefully.)
7. Once your SO has electronically signed the request, your organization will be active in Commons and you may create and maintain additional accounts for your institution staff.

To complete the registration above, you may need to register for the following if you haven't done so already:

- Dun & Bradstreet Number (DUNS) – https://fedgov.dnb.com/webform/
- Employer Identification Number (EIN) – https://www.irs.gov/businesses/small-businesses-self-employed/apply-for-an-employer-identification-number-ein-online
- Small Business Administration (SBA) – https://www.sbir.gov/registration
- System for Award Management (SAM) – https://www.sam.gov/SAM/

RECOMMENDED SUBMISSION TIPS

1. **FILE FORMAT:** As a reminder, the eRA ASSIST submission system requires the use of non-fillable PDF forms. Please be sure the PDF forms you are using are not fillable. If you are unsure whether your document meets the submission requirement it is recommended to print the document using a PDF writer (using the Print option and selecting the PDF writer as your printer) or contact the eRA Helpdesk.
2. **ADDING ATTACHMENTS:** Utilizing the eRA ASSIST option to upload multiple attachments will create a linked Table of Contents in your OT application based on the attachment file names in the order they are uploaded. Submitting in this manner will provide greater efficiency in the review of your application. Users may Click on the Add Attachment button pictured below for each attachment to be added.

As an example, the OT Office therefore recommends uploading multiple attachments using the following naming conventions in the following recommended order. Please note, the Attachment document must be saved or renamed using the File Name prior to upload:

Attachment	File Name
Technical Plan	Technical Plan.pdf
List of Key Personnel and Biosketches	Key Personnel.pdf
Leadership Plan for Multi-PI applications (as applicable)	Leadership Plan.pdf
Milestones Plan (as applicable)	Milestones.pdf
Main Budget and justification	<Application Component>Budget.pdf
Subaward 1 Budgets and justification	Subaward_<SubName1>.pdf
Subaward 2 Budgets and justification	Subaward_<SubName2>.pdf
Letters of Support	Letters of Support.pdf

3. **ADDITIONAL ATTACHMENTS:** If your application consists of any additional section not covered in the chart above, the OT office recommends that you submit the sections individually using the process referenced in #2 above. **Please be clear in the naming of files to indicate the content of each attachment.**
4. **REVIEW SUBMISSION**: At any point while preparing your application, you can click the "Validate Application" button to check business rules as well as preview the application image. This will ensure that the attachments appear as expected. Instructions for performing the validation can be found on page 12 of this document.
5. **SUBMISSION WALKTHROUGH:** For a complete walkthrough of the submission process please see the eRA OTA Submissions Instruction Guide on the following pages of this document.

Instruction Guide for OTA Submissions

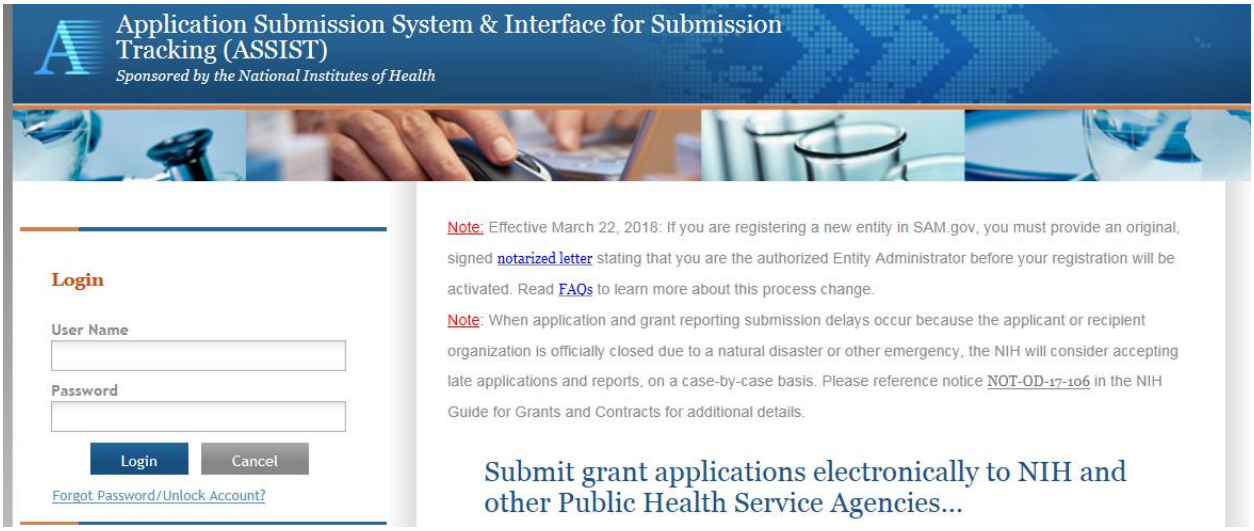
Use these instructions, together with the forms and information found in the funding opportunity announcement, to complete your application. The funding opportunity announcement (FOA) will include specific instruction and forms needed for your application submission. Remember that the FOA instructions always supersede these application instructions.

Prepare to Apply

NIH typically makes awards to organizations, not individuals. At the time of submission, the Program Director/Principal Investigator (PD/PI) and their organization must be registered at eRA Commons. In addition to the PD/PI, an individual with the role of Signing Official (SO) is needed. If an application is awarded, additional registrations (e.g. System Award Management) will be required.

OTA applications must be submitted using NIH's ASSIST web-based application submission system. Users can access ASSIST directly or through eRA Commons. To complete the application, users must have access to a browser, a pdf generator, and Adobe Reader software.

Log into ASSIST using eRA Commons credentials (username and password)



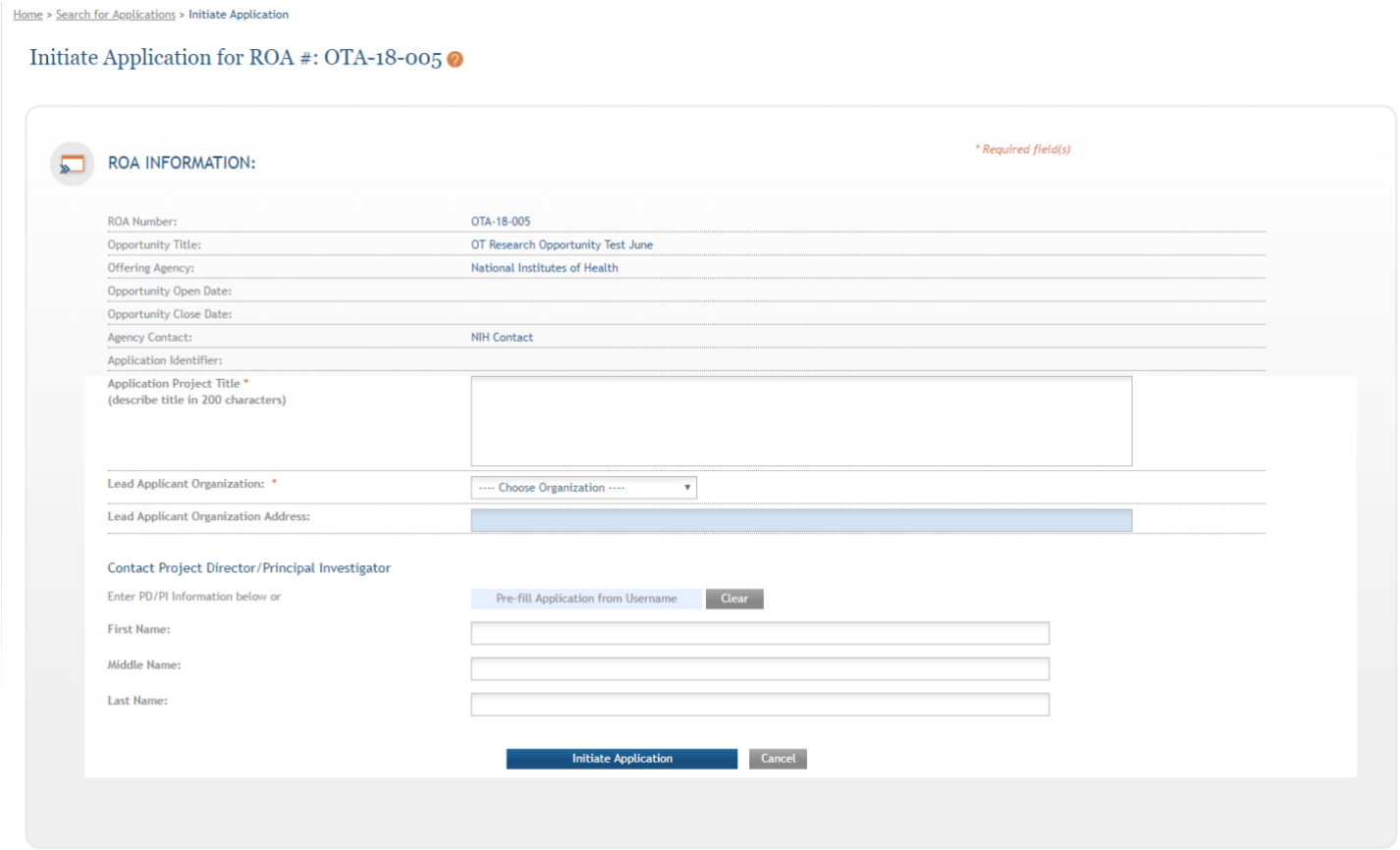
Initiating the Application

On the ASSIST Welcome screen, enter the OTA ROA number in the Funding Opportunity Announcement # field and then click 'Go'.

The Initiate Application screen contains several required elements: At a minimum the Application Project Title must be entered, and the Lead Applicant Organization must be selected from the drop-down menu.

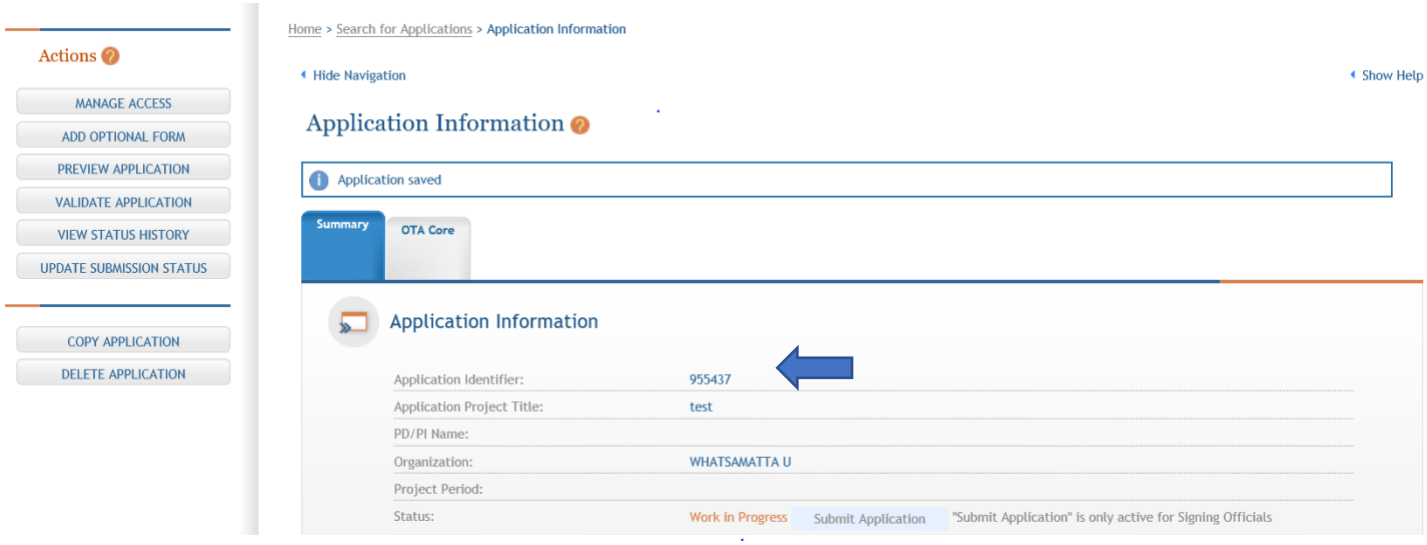
The Contact PD/PI fields may be pre-filled from Commons Username using the button or entered manually. These fields will be available to edit in the application once it has been initiated.

Once required fields have been satisfied, press the “Initiate Application” button to create an application record.



Navigating the Application

Each application in ASSIST receives a unique Application Identifier at creation. This value is displayed on the Application Information Summary page and can be used as a search key on the Search for Applications screen if returning to work on the application at a later time.



Actions are available on the left-hand side of the screen. The Summary screen and any constituent forms in the application are loaded as tabs to the right of the action menu.

Navigate to the OTA Core form by clicking the grey tab for the form. The currently active tab will be highlighted in blue.

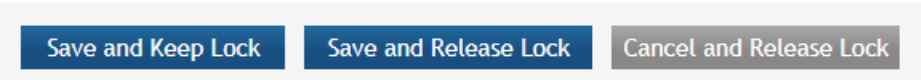
Completing the Application Form

All fields marked by an asterisk (*) are required.

All attachments should be in the format of a PDF file.

Click on “Edit” button to begin data entry. In the “edit” mode, data entry by other users is blocked until the lock-holder either releases the lock or it expires.

At the bottom of the form there are several saving options:

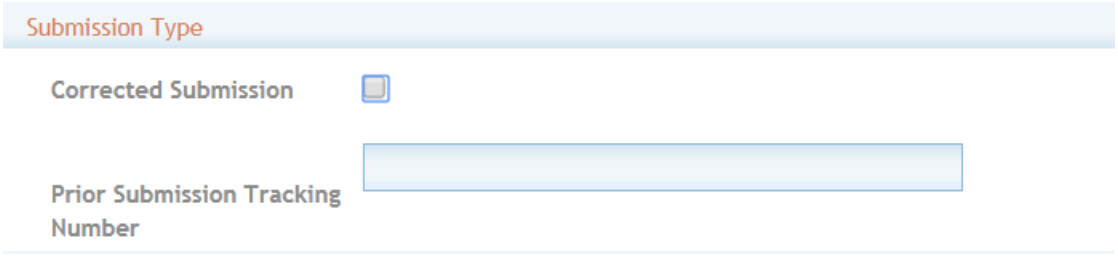


Save and Keep Lock: saves data and restricts data entry access to current user

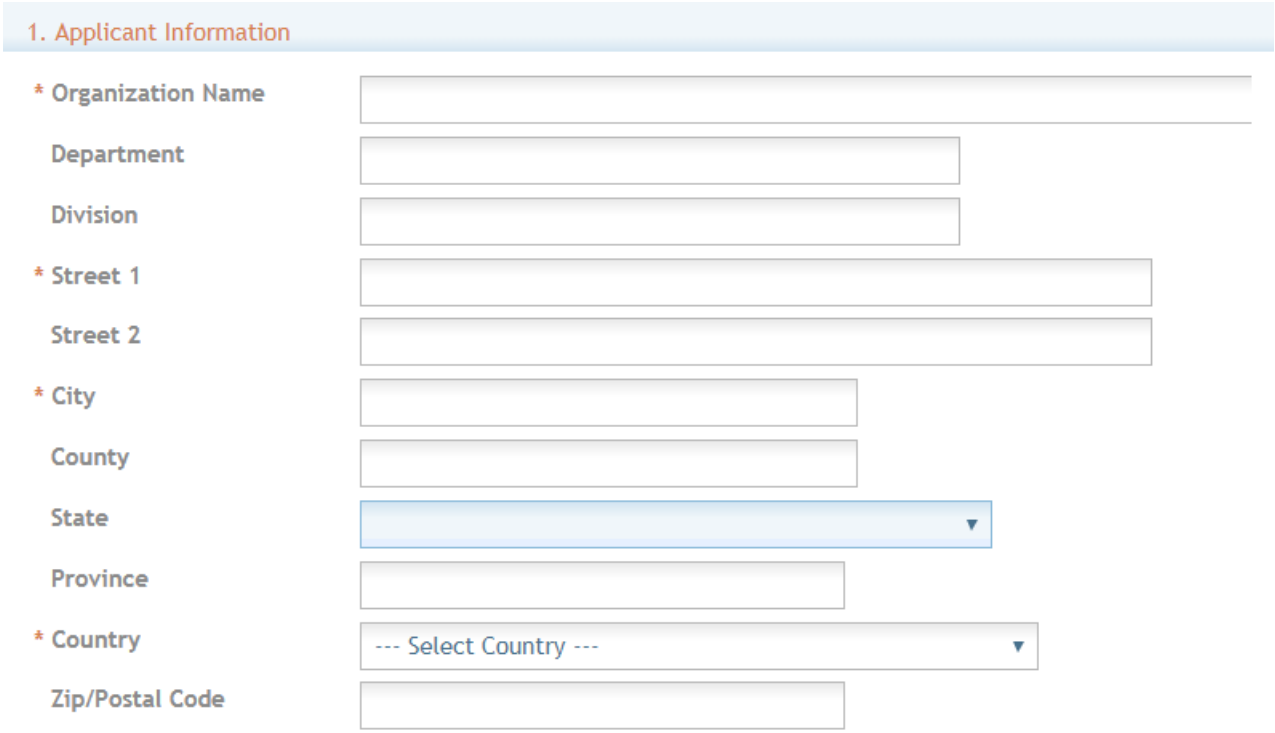
Save and Release Lock: saves data and releases form to other users

Cancel and Release Lock: does not save data and releases form to other users

Submission Type: If the application is being resubmitted after correcting errors/warning, check "Corrected Submission" and enter prior submission tracking number as it appears in the footer of the prior submission.



1. **Applicant Information**

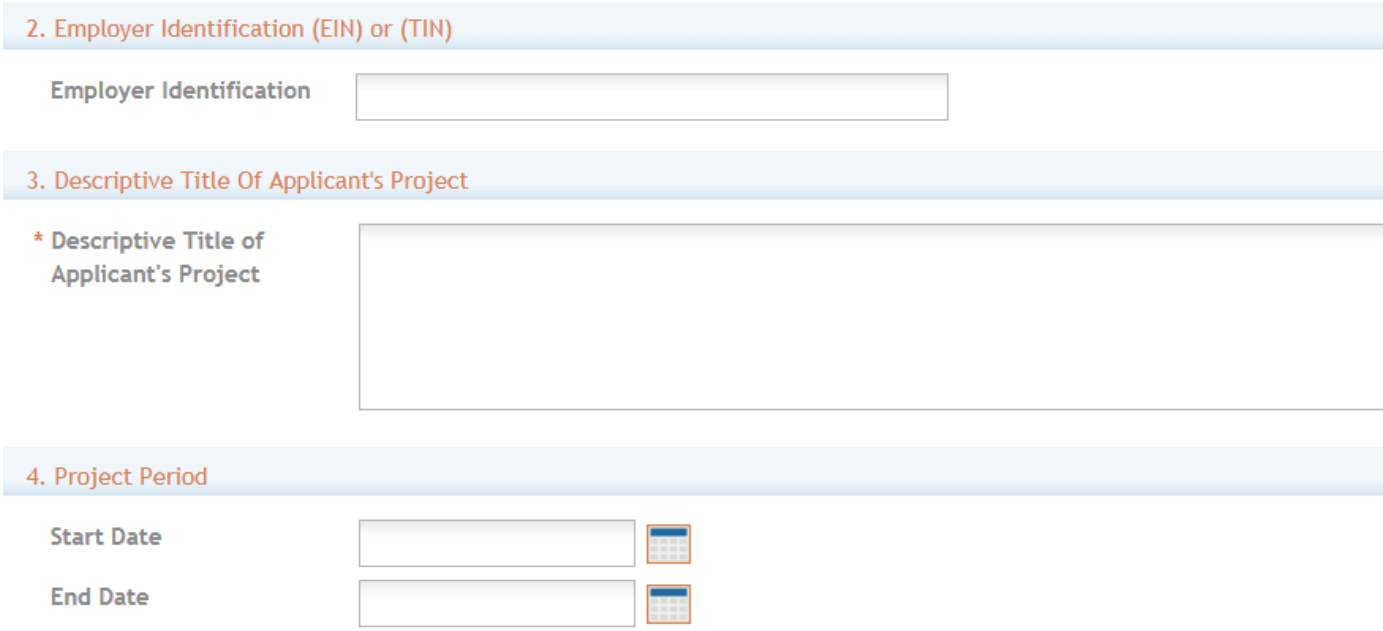


Organization Name: This field is required. Enter the name of the organization for the SO

Contact and Address fields: Enter the field data for the Applicant Organization as each label indicates. The Organization Name, Street 1, City, and Country are required fields. The State and ZIP/Postal Code fields will become required upon Country selection of United States. Note that ZIP/Postal Code must be entered in ZIP+4 (nine-digit postal code) format. Province is enabled for all non-US countries and required for Canada.

2. **Employer Identification (EIN) or (TIN):** Enter either the organization's Taxpayer Identification Number (TIN) or Employer Identification Number (EIN) as assigned by the

Internal Revenue Service. If your organization is not in the United States, enter 44-4444444. Your EIN may be 12 digits, and if this is the case, enter all 12 digits.



3. **Descriptive Title of Applicant's Project**: This field is required. The descriptive title is limited to 200 characters, including spaces and punctuation.

4. **Project Period**: Enter the proposed start date of the project.

The Start Date is an estimate. The project period should not exceed what is allowed in the ROA.

The End Date is an estimate and must occur in the future of the Start Date.

5. Project Director/Principal Investigator (PD/PI) Contact Information

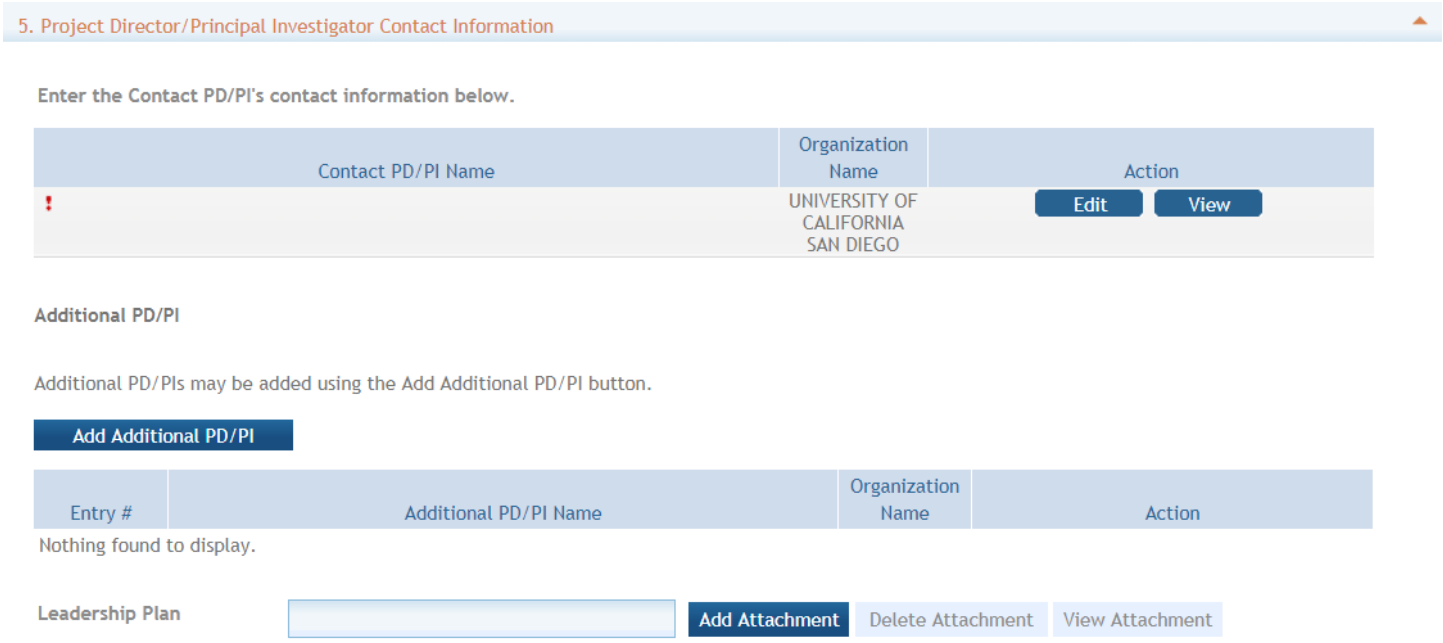
The PD/PI is the individual responsible for the overall scientific and technical direction of the project.
In the eRA Commons profile, the person listed here must be affiliated with the applicant organization entered in "1. Applicant Information".

If submitting an application with multiple PD/PIs, the main or primary PD/PI should be entered in the first or top section as the Contact PD/PI. The "Add Additional PD/PI" button may be used to add other PD/PIs. Following data entry, the user may "Edit" or "View" the PD/PI entries; Additional PD/PI entries can also be individually removed from the application.

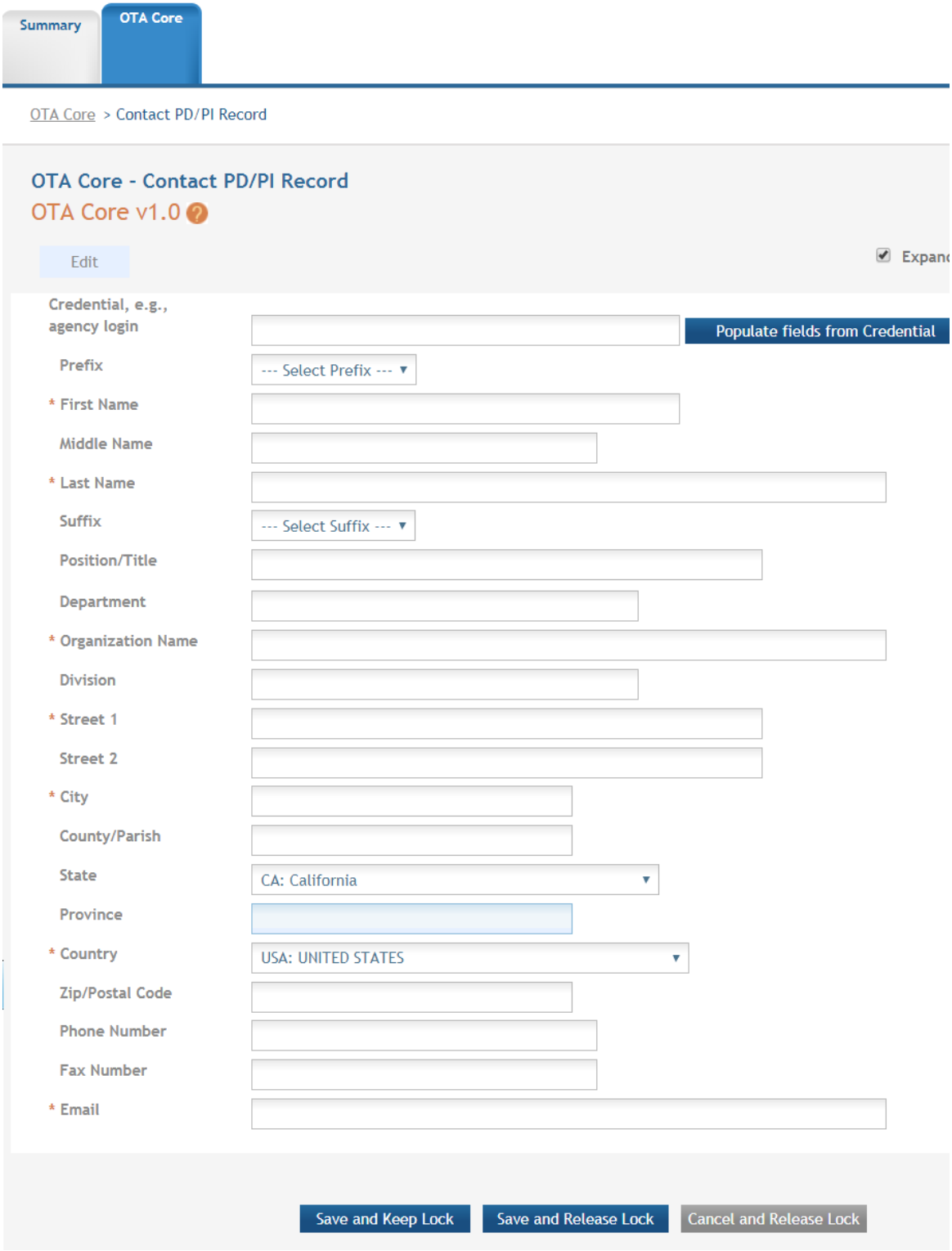
If the FOA requires a leadership plan for a multi-PD/PI application, provide the rationale for choosing a multiple PD/PI approach. The governance and organizational structure of the leadership team and the research project should be described, including communication plans,

processes for making decisions on scientific direction, and procedures for resolving conflicts. The roles and administrative, technical, and scientific responsibilities for the project or program should be delineated for the PD/PIs and other collaborators.

If budget allocation is planned, the distribution of resources to specific components of the project or the individual PD/PIs should be delineated in the Multiple PD/PI Leadership Plan. In the event of an award, the requested allocations may be reflected in a footnote on the Notice of Award.



The attachment should be attached as a PDF file.



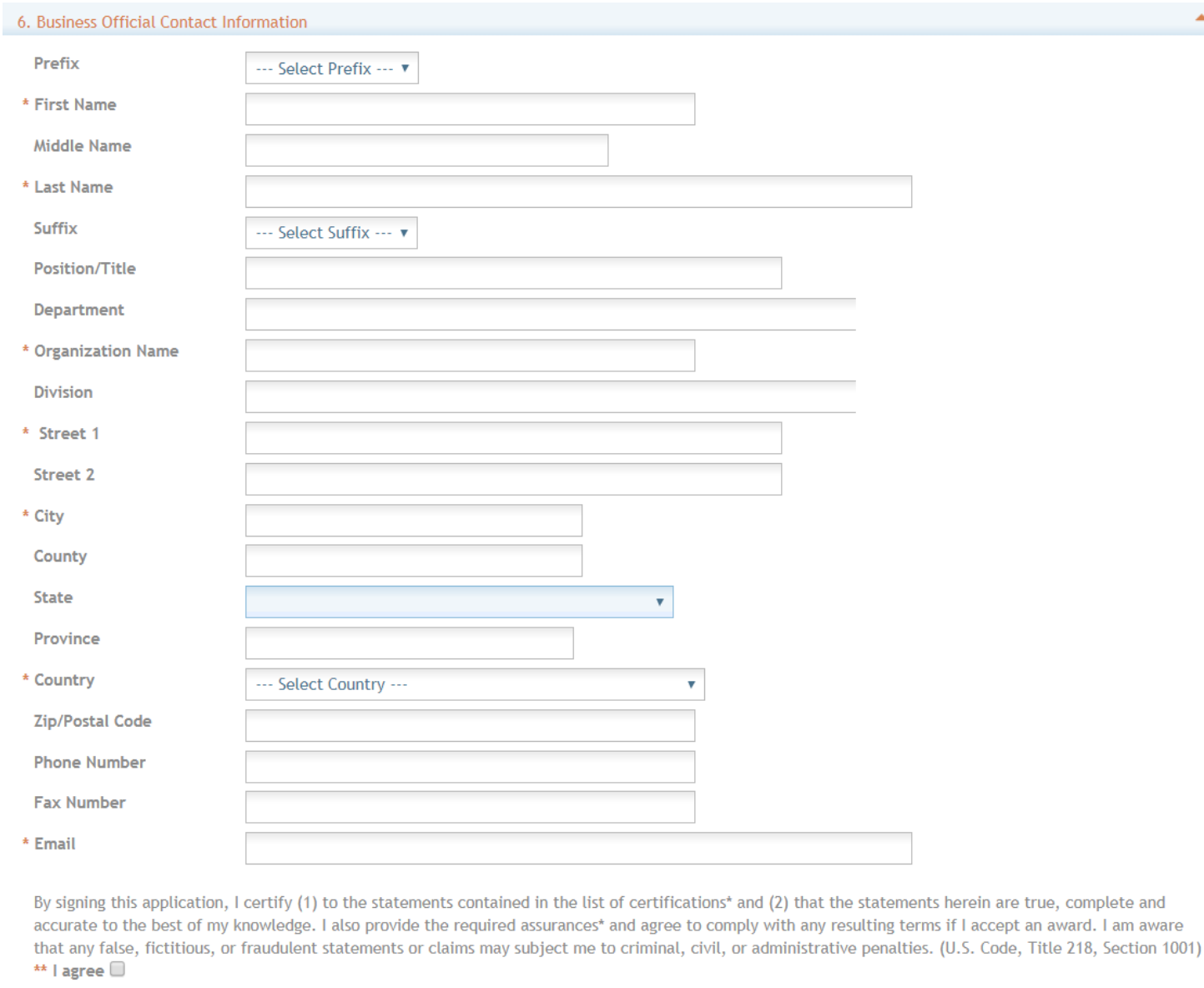
Credential, e.g. agency login: enter the eRA Commons user name for the PD/PI. This field is not required to save the form, however the Commons user Identification is for the Contact PD/PI at the time of submission.

Complete the personal information for the PD/PI.

Contact and Address fields: Enter the field data for the PD/PI as each label indicates. The Organization Name, Street 1, City, and Country are required fields. The State and ZIP/Postal Code fields will become required upon Country selection of United States. Note that ZIP/Postal Code must be entered in ZIP+4 (nine-digit postal code) format. Province is enabled for all non-US countries and required for Canada.

Following data entry, Save changes. To return to the main form use 'Save and Release Lock' or click the 'OTA Core' breadcrumb under the blue 'OTA Core' form tab.

6. Business Official Contact Information



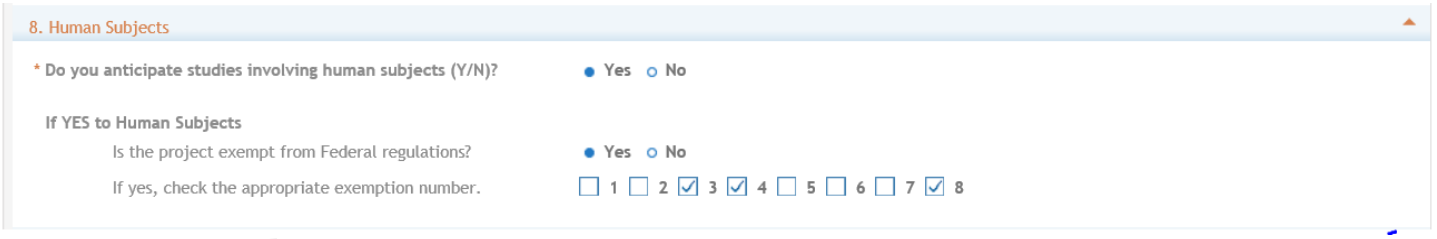
Complete the information for the Business Official Contact.

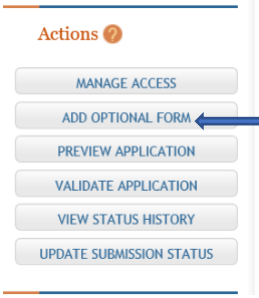
Contact and Address fields: Enter the field data for the Business Official as each label indicates. The Organization Name, Street 1, City, Country, and Email are required fields. The State and ZIP/Postal Code fields will become required upon Country selection of United States. Note that ZIP/Postal Code must be entered in ZIP+4 (nine-digit postal code) format. Province is enabled for all non-US countries and required for Canada.

Assurances: The applicant organization is responsible for verifying its eligibility and the accuracy, validity, and conformity with the most current institutional guidelines of all the administrative, fiscal, and scientific information in the application, including the Facilities and Administrative rate. Deliberate withholding, falsification, or misrepresentation of information could result in administrative actions, such as withdrawal of an application, suspension and/or termination of an award, debarment of individuals, as well as possible criminal and/or civil penalties. The signer further certifies that the applicant organization will be accountable both for the appropriate use of any funds awarded and for the performance of the grant-supported project or activities resulting from this application. The grantee institution may be liable for the reimbursement of funds associated with any inappropriate or fraudulent conduct of the project activity.
Check "I agree" to provide the required certifications and assurances.

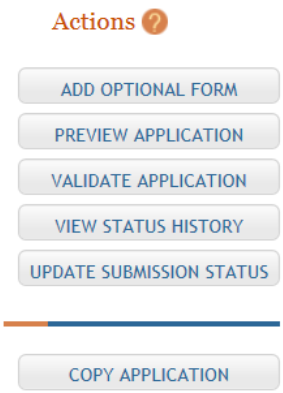
7. **Estimated Project Funding:** Enter the total federal funds, including Direct Costs and F&A (Indirect Costs) requested for the entire project period.

8. **Human Subjects**: Answer yes or no to the question regarding involvement of human subjects. If yes, indicate whether the studies are exempt from Federal Regulations. If required by the ROA, select the Human Subject Clinical Trial Information Form from the "Optional Forms" in the left navigation pane. Follow instruction in the SF424 instruction guide for this specific form.

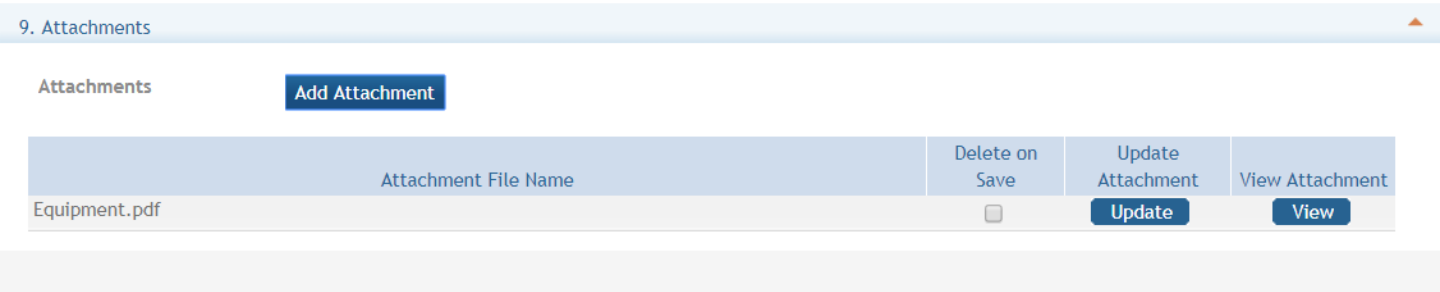




9. **Cover Letter**: The cover letter is for internal use only. It should be included any special considerations or explanatory details regarding the submission of the application. The letter should include the application title.



10. **Attachments:** attach PDF file in accordance with the FOA and/or specific instructions using the "add attachment" button. The attachment should be in the format of a PDF file. If multiple attachments are added, each should have a unique file name. To delete an uploaded attachment, check 'Delete on Save' and the file will be removed upon next save. Individual attachments can be replaced or updated by clicking the 'Update' button of the corresponding row.

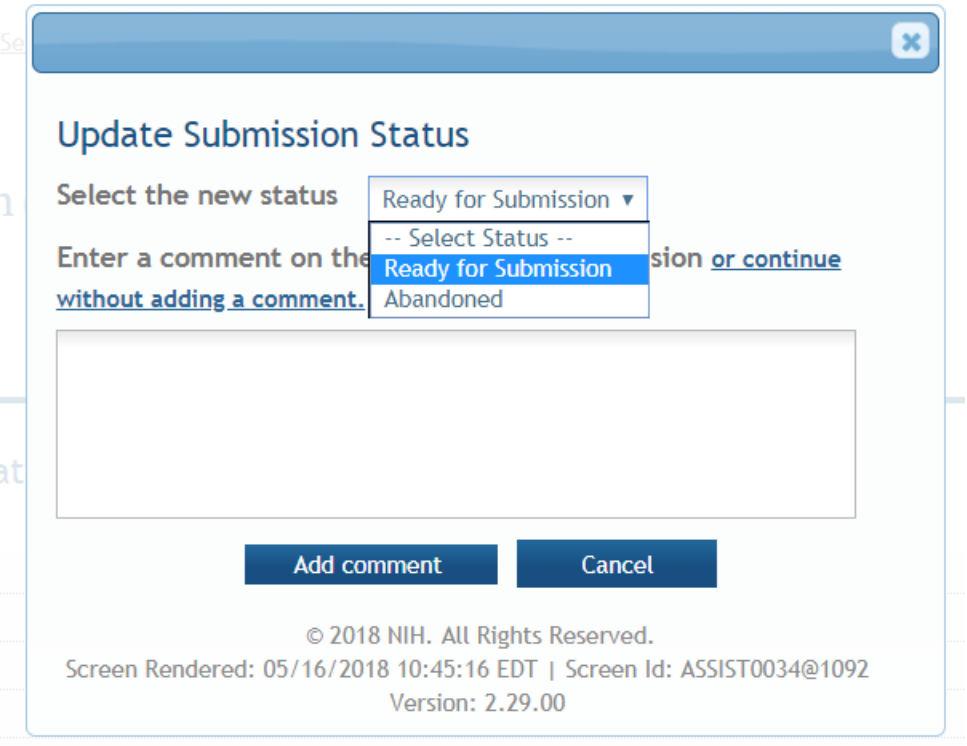


Submitting your Application

Validating your application: Select "Validate Application" from the left-hand panel to check your application for business errors and warning. Errors must be corrected prior to submission. Warnings will not stop prevent your application from being submitted.

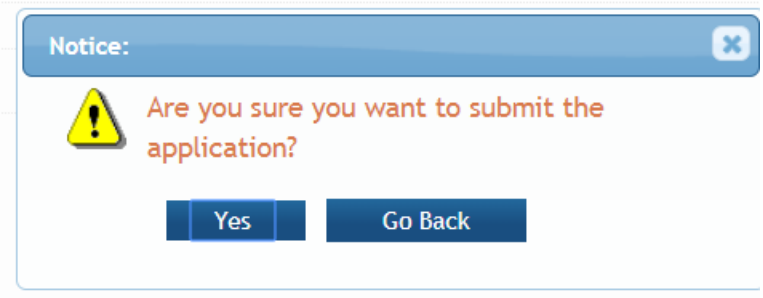
Preview Application: presents the PDF version of the application. This is a view of the assembled image is nearly identical to that the reviewers/evaluators will assess.

Update Submission Status: Prior to submission, the status must be updated to "Ready for Submission" which will also perform additional validations against business errors prior to changing the status. Click "Add Comment" or "continue without adding a comment" to continue.



Submission

Once the application has been placed in "Ready for Submission" status, the Signing Official will be able to click "Submit Application". After a confirmation that the SO does wish to submit the application, it will be sent to NIH for processing. The status can be viewed and updated by clicking the "View Submission Status Details" hyperlink next to the application status.



Follow the status of your submission in eRA Commons