Supporting Statement A for

The National Institute of Mental Health Data Archive (NDA), NIMH

OMB Control No. 0925-0667; REVISION

Date: September 13, 2017

Federal Government Employee Information:

Name: Gregory K. Farber, Ph.D.

Address: 6001 Executive Blvd., NSC 7162 Rockville, MD 20892-9640

Telephone: 301-435-0778

Email: farberg@mail.nih.gov

Table of contents

Attachments

Privacy Impact Assessment (PIA)

NDA Data Submission Agreement (DSA)

NDA Data Use Certification (DUC)

Tutorials that include screen shots and/or videos of the entire process related to the DSA and DUC are available at https://data-archive.nimh.nih.gov/training/training-modules.

A. Justification

The National Institute of Mental Health (NIMH) is requesting approval of the revised Data Submission Agreement (DSA) and Data Use Certification (DUC) information collection documents, which collect information about researchers submitting data and requesting access to shared data in the National Database for Autism Research (NDAR), and to change the repository name to the NIMH Data Archive (NDA). The NDA is an infrastructure for sharing human subjects research data and tools to further collaboration and scientific discovery. While the NDA is a single system, it is comprised of several research domains each with their own associated permission group. The information collected in the DSA and DUC is needed to provide appropriate permissions to either submit or access data, monitor the expiration of these permissions, track and report on permissions and requests, ensure that the terms of submission/access are followed, and provide appropriate attribution for data contributions, and communicate important information about the NDA.

A.1 Circumstances Making the Collection of Information Necessary

In February 2015, the National Institute of Mental Health (NIMH) received approval from OMB for use of the NIMH Data Repository DUC Form and the NIMH Data Repository DSA Form under OMB# 0925-0667 (expiration date 02/28/2018). NIMH is now requesting approval of updates to the DUC and DSA, to include additional terms/options for data submission and access to meet the needs of the expanding resource, and to change the repository name to the NIMH Data Archive (NDA). The NDA is a group of Federal data repositories based on an informatics platform for human-subjects research domains related to several scientific domains. Initially established as the National Database for Autism Research (NDAR) to support autism-related research, other research areas recognized the utility and benefit of the system prompting the expansion into other areas of mental health research prompting the NIMH to make data sharing an expectation for all clinical research it funds (see NOT-MH-14-015 and NOT-MH-15-012). Although the NDA is, in fact, a single system, the data for each logically delineated scientific domain is separated into permission groups. As of June 2017, the system has expanded to include the following research domains/permission groups, with additional domains planned/anticipated for inclusion during the next OMB approval period.

- National Database for Autism Research (NDAR)—data submission and access

- National Database for Clinical Trials Related to Mental Illness (NDCT)—data submission and access
- NIH Pediatric MRI Database (PedsMRI)—data access only
- Research Domain Criteria Database (RDoCdb)—data submission and access
- Adolescent Brain Cognitive Development Study (ABCD) – data access only
- Osteoarthritis Initiative (OAI) – data access only

NDAR: The NIH created NDAR (http://ndar.nih.gov), an informatics system and central data repository, housed at the NIH, to support and accelerate research in the prevention, cause, diagnosis, and treatment of autism spectrum disorder (ASD). NDAR collects a wide range of data types, including phenotypic, clinical, genomic, and neurological, as well as medical images, derived from individuals who participate in ASD research, regardless of the source of funding. NDAR provides the infrastructure to store, search across, retrieve, and analyze these varied types of data.

PedsMRI: The PedsMRI (www.pediatricmri.nih.gov) was created by four NIH institutes (NICHD, NIMH, NIDA and NINDS) using a contract mechanism to collect brain magnetic-resonance imaging data (anatomic MRI, supplemented by proton spectroscopy and diffusion tensor imaging) and correlated clinical/behavioral data from over 500 healthy, psychiatrically normal children, ages newborn to late adolescent/young adult. The goal of PedsMRI is to provide a normative reference that might be used to study healthy brain development as a basis for studying child and adolescent brain disorders and diseases, and for developing tools (e.g., image processing tools, atlases, etc.) that might be used to study brain development. PedsMRI has been moved to the NIH (housed at the NIH Data Center) and into the infrastructure provided by NDAR.

NDCT: NIMH has made data sharing an expectation for all future clinical trials funded by NIMH (see NOT-MH-14-015). Researchers are expected to submit both positive and negative data and results from NIMH-funded clinical trials to the NDCT (https://data-archive.nimh.nih.gov/ndct/), which will use the NDAR model. NDCT provides a system to support the submission, sharing and access of relevant data at all levels of biological and behavioral organization and for all data types. At present, data submitted to NDCT will be the result of grants funded through a series of NIMH funding opportunity announcements (FOAs).

RDoCdb: The Research Domain Criteria (RDoC) project aligns research in genetics, neuroscience, and behavioral science to develop a precision-medicine approach for classifying mental illnesses. In contrast to current symptom-based diagnostic systems for mental illnesses, precision medicine integrates many levels of information for each patient to define a precise diagnosis. Data submitted to the RDoC Database (RDoCdb) will include the results of grants funded through a series of NIMH FOAs in support of the RDoC project, as well as relevant mental health data submitted by other interested investigators, regardless of funding source. More information on the RDoC project and related FOAs can be found at http://www.nimh.nih.gov/research-priorities/rdoc/index.shtml.

ABCD: The ABCD Study is a long-term study of brain development and child health in the United States. Multiple NIH Institutes and Centers and additional federal partners are supporting this ambitious project. The ABCD Consortium consists of a Coordinating Center, a Data Analysis and Informatics Center, and 21 research sites across the country where investigators will perform regular, comprehensive biological and behavioral assessments on more than 10,000 children beginning when they are ages 9 or 10, continuing throughout adolescence into early adulthood. A more complete description of the study is available at https://abcdstudy.org.

OAI: The Osteoarthritis Initiative (OAI) is a multi-center, longitudinal, prospective observational study of knee osteoarthritis (OA). The overall aim of the OAI is to develop a public domain research resource to facilitate the scientific evaluation of biomarkers for osteoarthritis as potential surrogate endpoints for disease onset and progression. The OAI will establish and maintain a natural history database for osteoarthritis that will include clinical evaluation data, radiological (x-ray and magnetic resonance) images, and a biospecimen repository from 4796 men and women ages 45-79 enrolled between February 2004 and May 2006. Four 3.0 Tesla MRI scanners, one at each clinical center, are dedicated to imaging the knees of OAI participants annually over four years. The seven-year project will recruit participants who have, and those who are at high risk for developing, symptomatic knee osteoarthritis. Access to biospecimens will be by application to the National Institute of Arthritis, Musculoskeletal and Skin Diseases (NIAMS).

The potential for public benefit to be achieved through sharing autism, pediatric MRI, and clinical research data continues to be significant. The National Institutes of Health (NIH) and NIMH have expanded the NDA to store the collection of data from participants in research studies related to

mental health, regardless of the source of funding. The extensive information collected by these studies, and subsequently made available via NDAR, PedsMRI, NDCT, RDoCdb, ABCD, OAI, and other future permission groups provides a rare and valuable scientific resource.

A.2 Purpose and Use of the Information Collection

The primary uses of the information requested on the DSA and DUC forms are to a) help ensure that data submitted to the NDA are appropriate; and, b) help NDA staff to document, track, monitor, and evaluate the use of NDA, as well as to notify interested submitters and recipients of updates, corrections, or other changes to NDA. The type of information requested in the DUC and DSA satisfies the terms and conditions of the data sharing policies for NDA. When submitting data, the DSA reminds investigators that the NDA only accepts data from human subjects who have been appropriately consented for data sharing. The DSA also defines the requirement for data to be de-identified and describes to the researcher that certain expectations are in place in addition to simply submitting data. Further, researchers submitting data are provided an option to determine how others can access data through either an institutional sponsorship or individual sponsorship. Gathering information on the researcher and research project from which data will be submitted is important so that submitting researchers can be assisted throughout the multi-year process of data submission and to notify them of certain expectations for task completion based on data sharing terms often associated with grant awards. Similarly, the DUC defines the conditions under which users can access and use data from the NDA. Information on the proposed research question or scientific research development is needed by the Data Access Committee (DAC) to review and adjudicate data access requests. In addition, requestors must specify the permission group (NDAR, NDCT, RDoCdb, PedsMRI, ABCD, or OAI) from which they are requesting data access so that NDA staff may provide appropriate access if approved by the DAC. Practically speaking, the information requested on these forms helps to identify the users of the NDA system. The relationship with investigators submitting data to the NDA averages approximately 5 years while the relationship with users accessing data is about 1.5 years at a minimum. System enhancements and changes as well as new data will become available during these timeframes. Communication of these is paramount to the success of the NDA. Additionally, the valuable information requested on the DSA and DUC will help the NIH understand and evaluate the use of NDA in the research community.

A.3 Use of Information Technology and Burden Reduction

To **submit** data, an investigator must obtain data submission privileges. To obtain these privileges, an investigator must complete, sign and submit the DSA to NDA staff. The process for submitting the DSA is designed to be digital or paper based. Information may be typed or hand written into the form and the form can be uploaded via a web portal, emailed, or sent by US Mail or Courier. The DSA requests the following information:

- Contact information for the investigator seeking to submit (the Submitter).
- Information on the research project from which data will be submitted including the title, description, funding source, grant/contract number, funding amount, project dates and Clinical Trial ID (if applicable).
- Information on whether the project has or will collect tissue/biomedical samples, genomic data, or other data and, if so, the location into which these samples/data will be submitted.
- Selection of the type of sponsorship (Individual or Institutional) required for users to access data from the project
- The name and email address for an authorized institutional official (if Institutional Sponsorship data access type is selected).
- Co-signatures from the Submitting Investigator and the Investigator's Institutional Official certifying that they will abide by the DSA and the NIH principles, policies and procedures for the use of NDA. Investigators also acknowledge that they have shared the DSA document and the NIH policies and procedures with any research staff who will participate in the submission of data to the NDA. The Institutional Business Official(s) also acknowledges that they have shared this document and the relevant NIH policies and procedures with appropriate institutional organizations.

To gain **access** to data, an individual must obtain data access privileges. To obtain these privileges, an individual must complete, sign and submit the DUC to NDA Staff. The process for submitting the DUC to NDA is designed to be digital or paper based. Information may be typed or hand written into the form and the form can be uploaded via a web portal, emailed, or sent by US Mail or Private Courier. The NDA is in the process of implementing a fully online version of the data access request process with associated workflows allowing an individual to complete the DUC online and have it electronically sent to an Institutional Signing Official (if a permission group

requiring Institutional Sponsorship is requested) for electronic signature after which the request will be electronically sent to NDA staff. Electronic review and adjudication of requests for access will also be implemented for the Data Access Committee portion of the workflow.

The DUC requests the following information:

- The selection of the data requested by permission group. Users may request a single or multiple permission groups with one request thereby limiting the duplication of information needed.
- Contact information for the investigator seeking access (the Data Recipient), as well as for key/senior personnel in the Recipient's laboratory who will also require access as part of the Research Project.
- The title and a brief summary/abstract of the Research Project for which repository data are sought. A single paragraph is sufficient.
- Information on whether Co-signatures from the Recipient Investigator and the Investigator's Institutional Official certifying that they will abide by the DUC and the NIH principles, policies and procedures for the use of the repository/dataset. Investigators also acknowledge that they have shared the DUC and the NIH policies and procedures with any research staff who will participate in the use of NDA. The Institutional Business Official(s) also acknowledges that they have shared this document and the relevant NIH policies and procedures with appropriate institutional organizations.
- The institution's Federal Wide Assurance (FWA) number.

Once completed, the DUC is then sent for adjudication to the NDA Data Access Committee (DAC). The DAC is responsible for approving submission and access privileges to NDA. A Privacy Impact Assessment (PIA) for the DUC was approved on September 28, 2012.

A.4 Efforts to Identify Duplication and Use of Similar Information

Data access requestors using the DUC and users requesting permission to submit data via the DSA have the option of either completing a fillable pdf form by typing information, completing the form by handwriting the information or by utilizing the newly released and planned online process. This online process reduces the information that must be provided by the individual completing the form. For example, if the individual already has an eRA Commons ID then they may use this to begin to complete the DSA or DUC. Information on institutional business officials with signing

authority (as defined with an SO designation in the eRA Commons) will automatically be imported for selection by the data access requestor. This process can simplify the process and reduce the information requested from individuals. Similarly, when the institutional business official completes their section of either the DSA or the DUC document, their contact information can be imported. A digital signature is accepted for users and institutional business officials using either the fillable pdf form or the online workflow. For data access renewals, users can simply update information relevant to the new aspects of the project without updating basic information such as contact information thereby reducing the burden on the user. Due to the sensitive nature of the data contained in the NDA, and in accordance with existing NIH policies, such as genome-wide association studies (GWAS, see http://grants.nih.gov/grants/gwas/index.htm), data submission and access approvals are granted for one year and may be renewed thereupon.

A.5 Impact on Small Businesses or Other Small Entities

No small businesses will be involved in this project.

A.6 Consequences of Collecting the Information Less Frequently

The information requested in the NDA forms does not ask investigators to generate any new information, because the type of information being requested is fundamental to conducting any research study. The data are collected as needed with the primary purpose of ensuring that the users are aware of and follow the terms and conditions related to data submission and/or data access. The DSA needs to only be completed once per competitive grant. The DUC is required to be completed no more than once a year per lead researcher/investigator request. Additionally, the DUC states that data recipients may be asked to provide an annual summary of research accomplishments from using data accessed from the NDA. This is similar to updating a biographical sketch or CV, which is typically updated on a regular basis. Therefore, this is unlikely to be an undue burden when requested. Protecting the privacy of the research participants and the confidentiality of their data is critically important. Essential aspects of that protection are careful screening of who may submit or obtain access to NDA, and ongoing monitoring of the use of the NDA.

A.7 Special Circumstances Relating to the Guidelines of 5 CFR 1320.5

Guideless of 5 CFR 1320.5 are not applicable to this project.

A.8.1 Comments in Response to the Federal Register Notice

A Federal Register Notice for this revision request was published on July 11, 2017, Vol 82 FR 32005 (https://www.federalregister.gov/documents/2017/07/11/2017-14451/proposed-collection-60-day-comment-request-the-national-institute-of-mental-health-data-archive-nda). No public comments were received.

A.8.2 Efforts to Consult Outside Agency

Investigators and institutional business officials are consulted on a regular basis during the performance of other activities related to NDA data submission and/or data access. While proactive conversations are not elicited, investigators and institutional business officials have provided feedback on the type of information requested on the DSA and the DUC. Additionally, feedback is provided on the terms and conditions included in these NDA forms as well as instructions for completion of the forms. The revised DSA and DUC submitted for approval include the comments from such representatives in an effort to reduce burden and allow for the completion of the forms in a timelier manner.

A.9 Explanation of Any Payment of Gift to Respondents

No payment or gift will be provided to respondents.

A.10 Assurance of Confidentiality Provided to Respondents

The Federal Privacy Act ensures that no sensitive or personally identifiable information, located in federal systems of records (e.g., Recipient NIH records), is being shared. A system of records is any group of records under the control of a federal agency from which information is retrieved by the name of the individual or by some identifying number, symbol, or other identifying particular assigned to the individual. The NIH and any sites that are provided access to the datasets will have access to the data collected from the Recipient for the purposes described above. In addition, the Act allows the release of some information in the Recipient's records without his/her permission; for example, if it is required by members of Congress or other authorized individuals. The information requested is voluntary, but necessary for obtaining access to data.

The information requested from the investigator seeking to submit data, as part of the DSA, may be made public in part or in whole for tracking and reporting purposes. Each Data Access

Request provides a Privacy Act Notification pursuant to Public Law 93-579, Privacy Act of 1974, 5 U.S.C. Section 552a. Authority for the collection of the information requested from the submitting investigators comes from the authorities regarding the establishment of the National Institutes of Health, its general authority to conduct and fund research and to provide training assistance, and its general authority to maintain records in connection with these and its other functions (42 U.S.C. 203, 241, 289l-1 and 44 U.S.C. 3101), and Section 301 and 493 of the Public Health Service Act. These records will be maintained in accordance with the Privacy Act System of Record Notice 09-25-0200 (http://oma.od.nih.gov/public/ms/privacy/pafiles/0200.htm) covering "Clinical, Basic and Population-based Research Studies of the National Institutes of Health (NIH), HHS/NIH/OD."

The information requested from the investigator seeking access to NDA data, as part of the DUC, may be made public in part or in whole for tracking and reporting purposes. The DUC Form provides a Privacy Act Notification pursuant to Public Law 93-579, Privacy Act of 1974, 5 U.S.C. Section 552a. Authority for the collection of the information requested from the recipient investigators comes from the authorities regarding the establishment of the National Institutes of Health, its general authority to conduct and fund research and to provide training assistance, and its general authority to maintain records in connection with these and its other functions (42 U.S.C. 203, 241, 289l-1 and 44 U.S.C. 3101), and Section 301 and 493 of the Public Health Service Act. These records will be maintained in accordance with the Privacy Act System of Record Notice 09-25-0156, September 26, 2002, 67 FR 60742-60794 (http://oma.od.nih.gov/ms/privacy/pa-files/0156.htm) covering "Records of Participants in Programs and Respondents in Surveys Used to Evaluate Programs of the Public Health Service, HHS/PHS/NIH/OD."

Although the repository data will be coded (or de-identified) and NIMH will not hold direct identifiers to individuals within NDA, the agency recognizes the personal and potentially sensitive nature of the genotype-phenotype data. Investigators and institutions seeking access to data or images from the repository are expected to meet data security measures and to submit a DUC, co-signed by the investigator and the designated Institutional Official, as applicable.

A.11 Justification for Sensitive Questions

The NDA does not ask any questions of a sensitive nature, such as sexual behavior and attitudes, religious beliefs, and other matters that are commonly considered private; and therefore, there is

no need to provide a justification for this type of information. The NDA will not distribute sensitive data. Upon submission of data, NDA staff performs a quality control review to ensure that no personally identifiable information (PII) is contained in the dataset or supporting documentation. Only data that have undergone a quality control review are approved for sharing with the research community.

A.12.1 Estimates of Hour Burden Including Annualized Hourly Costs

Table 12-1: Estimated Annualized Burden Hours

Form Name	Type of Respondents	Number of Respondents	Number of Responses per Respondent	Average Burden Per Response (in hours)	Total Annual Burden Hours
NDA Data Submission Agreement (DSA)	Researchers submitting data	250	1	90/60	375
NDA Data Use Certification (DUC)	Researchers requesting access to data	750	1	90/60	1125
TOTAL		1000	1000		1500

A.12-2 Annual Cost to respondents

Table 12-2: Annualized Cost to Respondents

Type of Respondents	Total Annual Burden Hours	Hourly Respondent Wage Rate*	Respondent Cost
Researchers submitting data	375	$45.68	$17,130
Researchers requesting access to data	1125	$45.68	$51,390
TOTAL	1500		$68,520

**Bureau of Labor Statistics: May 2016 National Occupational Employment and Wage Estimates*
https://www.bls.gov/oes/current/oes191042.htm

A.13 Estimate of Other Total Annual Cost Burden to Respondents or Record Keepers

There are no additional costs other than the respondents' burden given in A12.

A.14 Annualized Cost to the Federal Government

The total annualized cost to the Federal Government is $27,108. No additional operational expenses such as equipment, overhead, printing, and support staff will be needed. The Division Director, who serves as the DAC Director, is involved in the policy aspects of the DAC and DAC membership selection. They are not involved in the review and adjudication of data access requests as a general rule. The Administrative Data Analyst will coordinate the daily activities related to the collection of information. The Senior Program Analyst will provide clarification and guidance for cases beyond the expertise of the Administrative Data Analyst.

Cost Descriptions	Grade/Step	Salary	% of Effort	Fringe (if applicable)	Total Cost to Gov't
Federal Oversight					
Data Access Committee (DAC) Members					
Division Director	GS-15, Step 4	$144,945	<1% FTE		$347
Program Officer	GS-14, Step 4	$123,223	6% FTE		$7,380
Contractor Cost					
Senior Program Analyst		$100,000		$20,000	$1,500
Administrative Data Analyst		$50,000		$35,000	$17,881
Travel					$0
Other Cost					$0
Total					$27,108

** Salary/Wage Source: Office of Personnel Management 2017 Salary Table for the Locality Pay Area of Washington-Baltimore-Arlington https://www.opm.gov/policy-data-oversight/pay-leave/salaries-wages/salary-tables/pdf/2017/DCB.pdf.*

A.15 Explanation for Program Changes or Adjustments

NIMH is requesting approval of revisions to the DUC and DSA, to include additional terms/options for data submission and access to meet the needs of the expanding resource, and to change the repository name to the NIMH Data Archive (NDA). As data sharing expectations have become more prevalently applied to awards of NIH funded grants and contracts, the NDA has continued to expand in an effort to accommodate data yielded from these projects. The increase in estimated

hourly burden and respondent cost noted in Section A12 above, as compared to the estimates in our 2015 submission, is directly related to the growth and expansion of the NDA within the scientific community and clarifications made to the DSA and DUC terms.

We expect the volume of DUC and DSA instruments received will also increase given the additional permission groups and types of data access sponsorship. To provide a segmentation of the research domains supported, we added permission groups on the DUC. In the DUC, data access requestors may choose one or more permission groups even though the location of the underlying data are co-located. It is expected that additional research domains will be supported over the next few years prompting the creation of additional permission groups. Another overarching change relates to the requirements for accessing data. In the past all data access requests have required sponsorship by an Institution with an active FWA, which includes the signature of an institutional signing official. New research domains have been incorporated into the NDA that do not carry this level of requirement. A newly added Individual sponsorship does not require an active FWA or institutional business official signature, but does still require the requestor to agree to the terms of data access. In both cases, new groups from which data are requested have been added to the DUC. The DSA has added the collection of additional grant related information that is targeted toward researchers not funded by the NIH who are requesting permission to submit data. Because the funding information is not accessible by NDA staff, the information is requested on the form.

The NIH and NIMH seek to encourage the use of these resources to achieve rapid scientific progress. In order to take full advantage of such resources and maximize their research value, it is important that data is broadly made available, on appropriate terms and conditions, to the largest possible number of qualified investigators in a timely manner.

A.16 Plans for Tabulation and Publication and Project Time Schedule

Information collected on the DSA and DUC may be published on the NDA-supported websites and may be used for internal monitoring purposes. Publication of requested information is done under the authority provided by the document signatories given the specific term included in both documents.

A.17 Reason(s) Display of OMB Expiration Date is Inappropriate

The OMB Control Number (0925-0667) and Expiration Date will be displayed appropriately.

A.18 Exceptions to Certification for Paperwork Reduction Act Submissions

None.