Supporting Statement A for

National Institute of Mental Health
Data Repositories
Data Submission Request
and
Data Access Request & Use Certification
OMB # 0925-0667

December 17, 2014

Keisha Shropshire, MPH
National Institute of Mental Health
6001 Executive Boulevard, Room 6202, MSC 9667
Bethesda, Maryland 20852-
Email: kshropsh@mail.nih.gov
Phone: 301-451-0292
Fax: 301-443-1362

TABLE OF CONTENTS

A.1 Circumstances Making the Collection of Information Necessary

The National Institutes of Mental Health (NIMH) Data Repositories are a group of Federal data repositories based on an informatics platform for human-subjects research domains related to mental health, initially established as the National Database for Autism Research to support autism-related research. As of June 2014, the system has expanded to include the following repositories:

- 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

In 2013, NIMH received approval from OMB for use of the NIMH Data Access Request and Use Certification (DUC) Form to meet the unique data access needs of all existing NIMH data repositories, which at the time consisted of NDAR, PedsMRI, and the NIMH Clinical Research Datasets (NCRD)—OMB# 0925-0667 (Expiration: 09/30/2016). Now in 2014, two new databases have been added and integrated into the NDAR infrastructure, NDCT and RDoCdb. At this time, NIMH is seeking OMB approval to revise, the all-purpose NIMH Data Repositories Data Access & Use Certification form, as well as add a new all-purpose NIMH Data Repositories Data Submission form. As the data repositories have matured, and with the introduction of the new databases—namely NDCT and RDoCdb—the information being collected for data submission has become more complex, rendering an OMB-approved submission form a new necessity.

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, which will use the NDAR model. NDCT will provide 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). NDCT will encompass data from the NCRD.

NCRD: The data from NIMH-supported clinical trials are an important scientific resource (http://www.nimh.nih.gov/funding/clinical-trials-for-researchers/datasets/nimh-procedures-for-requesting-data-sets.shtml). It is the view of NIMH that their full value can only be realized if they are made available, under appropriate terms and conditions, in a timely manner to the wider scientific community, provided that the privacy of study participants are protected to extent provided by law.

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 data submitted by other interested investigators, regardless of funding source.

The potential for public benefit to be achieved through sharing autism, pediatric MRI, and clinical trials research data is significant. However, genotype and phenotype information generated about individuals, such as data related to the presence or risk of developing brain disorders and information regarding paternity or ancestry, may be sensitive. Therefore, protecting the privacy of the research participants and the confidentiality of their data are critically important. Risks to individuals, groups, or communities should be balanced carefully with potential benefits of the knowledge to be gained through NIMH data repositories/sets.

The information requested from the investigator seeking to **submit** data, as part of the Data Submission Request for these repositories, 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 repository/set data, as part of the Data Access Request for these repositories and datasets, 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 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, (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." The NIH System of Record Notice was previously published in the Federal register on September 26, 2002, Volume 67, No 187, page 60742.

A.2 Purpose and Use of the Information Collection

The primary uses of this information are to document, track, monitor, and evaluate the use of the repository datasets, as well as to notify interested submitters and recipients of updates, corrections, or other changes to the database. The type of information requested in the new NIMH Data Repositories Submission Agreement and in the revised NIMH Data Repositories Access Request/Use Certification satisfies the terms and conditions of the data sharing policies for these data. The new forms contain a section wherein investigators can provide a description of the research project they either have performed or are proposing to perform with the data, as the case may be. The terms and conditions associated with the NIMH Data Repositories remind investigators to provide an annual summary of research accomplishments and publications from using the repository/set in an updated biographical sketch or CV, which can be completed on the NIMH Data Repositories' websites. As investigators typically update their sketches and CVs on a regular basis, this is unlikely to be an undue burden when requested. This valuable information will help NIH understand and evaluate the use of repositories/datasets 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, scan, and upload or email the Data Submission Agreement (SA) to the repository web portal. The document must include the Federal-wide Assurance (FWA) number of the investigator's affiliated institutions, and be co-signed by an NIH-recognized Business/Institutional Official. Thus, the process for submitting data to the repositories is designed to be both electronic (information may be

typed into the form and the form is uploaded via a web portal/emailed) and mechanical (signatures are requested on the form, which is then scanned and uploaded/emailed):

The NIMH Submission Agreement requests the following information:

- The title and a brief summary/abstract of the Research Project to which repository data are being submitted. A single paragraph is sufficient.
- Contact information for the investigator seeking to submit (the Submitter).
- A statement indicating whether data from biospecimens have any use restrictions and what those restrictions are.
- A statement indicating whether the Submitter has applied for or obtained a Certificate of Confidentiality for the data.
- Electronic copies of the study protocol, questionnaires, study manuals, variables measured, and other supporting documentation, as appropriate.
- Co-signatures from the Submitting Investigator and the Investigator's Institutional Official certifying that they will abide by the SA and the NIH principles, policies and procedures for the use of the repository. Investigators also acknowledge that they have shared the Data Submission Agreement document and the NIH policies and procedures with any research staff who will participate in the use of the repository. 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 FWA number.

To gain **access** to data, an investigator must obtain data access privileges. To obtain these privileges, an investigator must complete, sign, scan, and upload or email the Data Access Request/Data Use Certification to the repository web portal. The document must include the Federal-wide Assurance (FWA) number of the investigator's affiliated institution, and be co-signed by an NIH-recognized Business/Institutional Official. Thus, the process for obtaining access to data within the repository is designed to be both electronic (information may be typed into the form and the form is uploaded via a web portal/emailed) and mechanical (signatures are requested on the form, which is then scanned and uploaded/emailed):

The NIMH Data Access Request/Use Certification requests the following information:

- The title and a brief summary/abstract of the Research Project for which repository data are sought. A single paragraph is sufficient.
- 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.
- 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 Data Access Agreement document and the NIH policies and procedures with any research staff who will participate in the

use of the repository. 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 FWA number.

Once completed, the request package is then sent for adjudication to the relevant Data Access Committee or authority (for NDAR, Pediatric MRI, NDCT/NCRD), and RDoCdb) established to oversee submission to the repository and access to the shared data. When the investigator's request is approved, the investigator is notified by e-mail.

A Privacy Impact Assessment (PIA) for NIMH Data Access Request/Use Certification was approved on September 28, 2012.

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

To protect and assure the privacy of all research participants whose data have been submitted to the repositories, investigators who seek access to these data are expected to adhere to the specifications of the principles outlined in the repository Data Sharing Policy (for NDAR, see http://ndar.nih.gov/ndarpublicweb/Documents/NDAR_Policy.pdf, section entitled, "Data Access"; for PedsMRI, see http://pediatricmri.nih.gov/nihpd/info/irb.html; for NDCT, see http://grants.nih.gov/grants/guide/notice-files/NOT-MH-14-015.html). Furthermore, each research project is unique, and collecting information about these projects, through the new NIMH Data Repositories forms, will enable NIH to document, track, monitor, and evaluate the use of the NDAR, PedsMRI, NDCT, and RDoCdb data, as well as to notify interested recipients of updates, corrections, or other changes to the database.

Due to the sensitive nature of the data contained in the repositories/datasets, and in accordance with existing NIH policies, such as that for NDAR, PedsMRI, and 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 study.

A.6 Consequences of Collecting the Information Less Frequently

The information requested in the NIMH Data Repositories 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 on a needed basis. We anticipate no more than once a year per researcher/investigator request.

Additionally, the NIMH Data Repositories forms state that data submitters and recipients may be asked to provide an annual summary of research accomplishments from using the repositories in an updated biographical sketch or CV (as noted above, data submission/access approvals are granted for one year and may be renewed thereupon).

As investigators typically update their sketches and CVs on a regular basis, this is unlikely to be an undue burden when requested.

As stated before, protecting the privacy of the research participants and the confidentiality of their data is critically important. Essential aspects of that protection are careful screening who may submit or obtain access to the database, and ongoing monitoring of the use of those data.

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

Not Applicable.

A.8 Comments in Response to the Federal Register Notice and Efforts to Consult Outside Agency

A Federal Register Notice was published on October 7, 2014, Vol. 79 FR 60479 (https://www.federalregister.gov/articles/2014/10/07/2014-23959/proposed-collection-60-day-comment-request-nimh-data-repositories-data-submission-request-nimh-data).. No public comments were received. No outside agency consult is necessary or required. The NIMH Data Repositories will enact the Data Access Committee (DAC) pertaining to the specific repositories included (i.e., NDAR, PedsMRI, NCDT, etc.).

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. he 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 Data Submission Request for these repositories, 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 repository data, as part of the NIMH Data Access Request and Use Certification, may be made public in part or in whole for tracking and reporting purposes. The NIMH Data Repositories 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 the NIH will not hold direct identifiers to individuals within the NIH data repository, 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 NIMH Data Access Request and Use Certification, co-signed by the investigator and the designated Institutional Official(s) (see http://ndar.nih.gov/ndarpublicweb/Documents/NDAR_Policy.pdf, http://pediatricmri.nih.gov/nihpd/info/irb.html, and http://www.nimh.nih.gov/trials/datasets/nimh-policy-for-distribution-of-data.shtml). The NDAR Data Access Committee reviews and approves all submission and access requests. (see http://ndar.nih.gov/ndarpublicweb/policies.html#AccessCommittee). PedsMRI, NDCT, and RDoCdb have their own separate, independent Data Access Committees or authorities.

A.11 Justification for Sensitive Questions

NDAR, PedsMRI, NDCT, and RDoCdb do 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, do no need to provide a justification for this type of information. The NIMH Data Repositories will not distribute sensitive data.

Upon submission of data, the repository 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 Estimates of Hour Burden Including Annualized Hourly Costs

There are two scenarios for completing the form. Sometimes the Principal Investigator completes the whole document, and other times he/she has a Research Assistant complete it (after which the Investigator reviews and signs it).

A. Estimates Annual Burden Hours				
Form	**Estimated Number of Respondents**	**Estimated Frequency of Response**	**Average time per response (in hours)**	**Estimated Total Annual Burden Hour Requested**
Data Submission Agreement	40	1	95/60	63
Data Access Request	100	1	95/60	158
Total	140			221
B. Estimates of Total Annual Cost Burden				
Form	**Estimate Total Annual Burden Hours**	**Wage rate**	**Total Costs**	
Data Submission Agreement	63	$ 91.00	$ 5,733	
Data Access Request	158	$ 91.00	$ 14,378	

Salary/Wage Source: Office of Personnel Management 2014 General Schedule Locality Salary Table for various GS-levels; Bureau of Labor Statistics/Occupational Employment "Life, Physical, and Social Scientist" http://www.bls.gov/oes/current/oes_13644.htm#19-0000, occupational code 19-2099.

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

Staff	Grade/Salary	Repository	NIMH	Repository

	(Percent FTE or Effort)	Operations Time	Repository Hourly Rate	Operations Cost*
Repository Operations Staff Tier 1	$36,711 (< 2% Effort)	2.25 hours	17.59	$39.58
Repository Operations Staff Tier 2	$149,993 (< 2% Effort)	1.25 hour	71.87	$89.84
Data Access Committee (DAC) Staff				
Division Director	GS-15 Step 4 (137,494) (< 1% FTE)	1.25	$ 65.88	82.35
Branch Chief	GS-14 Step 4 (116,887) (< 1% FTE)	1.25	$ 56.01	70.01
Program Officer	GS-14 Step 4 (109,804) (< 1% FTE)	1.25	$ 56.01	70.01
Contractor	$149,994 (< 1% FTE)	1.25	$79.06	98.83
Total per Data Repositories Access, Submission, and Use Certification Request Forms				$233.13
Annual # of Data Access Requests				100
Total Annualized Cost				$23,763.62

Salary/Wage Source: Office of Personnel Management 2014 General Schedule Locality Salary Table for various GS-levels; Bureau of Labor Statistics/Occupational Employment "Life, Physical, and Social Scientist" http://www.bls.gov/oes/current/oes_13644.htm#19-0000, occupational code 19-2099.

A.15 Explanation for Program Changes or Adjustments

This is a revised data collection. At this time, NIMH is seeking OMB approval to revise, the all-purpose NIMH Data Repositories Data Access & Use Certification form, as well as add a new all-purpose NIMH Data Repositories Data Submission form.

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

There is no specific plan to publish the data collected from this form. These data are for internal monitoring purposes.

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

Not applicable.

A.18 Exceptions to Certification for Paperwork Reduction Act Submissions

Not applicable.