Supporting Statement A for

NIH Information Collection Forms to Support Genomic Data Sharing for Research Purposes

(OD)

OMB# 0925-0670 and expiration date, 07/31/2019

Date: July 1, 2019

Check off which applies:

- New
- X Revision
- Reinstatement with Change
- Reinstatement without Change
- Extension
- Emergency
- Existing

Name: Lyric A. Jorgenson, Ph.D.
Acting Director, Division of Scientific Data Sharing Policy
Office of Science Policy, NIH
Address: 6705 Rockledge Dr., Suite 750, Bethesda, MD 20892
Telephone: 301-496-6837
Fax: 301-496-9839
Email: jorgensonla@mail.nih.gov

Table of Contents

Attachments

Attachment 1 – Study Registration and Institutional Certification template

Attachment 2 – Data Access Request, Renewal, and Close-out forms

Attachment 3 – Supplementary Document: dbGaP Registration Reminder Notice

Attachment 4 – Privacy Impact Assessment

A. ABSTRACT

Sharing research data is integral to the mission of the National Institutes of Health (NIH), as it advances our understanding of factors that influence health and disease, while also providing opportunities to accelerate research through the power of combining large, information-rich datasets. To promote robust sharing of human and non-human genomic data from a wide range of large-scale genomic research, and to provide appropriate protections for research involving human data, the NIH issued the Genomic Data Sharing (GDS) Policy. The NIH GDS Policy, an extension of the 2008 *NIH Policy for Sharing of Data Obtained in NIH-Supported or Conducted Genome-Wide Association Studies*, applies to NIH-funded research that generates large-scale human or non-human genomic data as well as the use of these data for subsequent research.

The NIH GDS Policy addresses data submission, maintenance, and access. Under the Policy, NIH-funded investigators generating large-scale genomic data are expected to submit these data and associated phenotypic data to an NIH-designated repository. Human genomic data submission and access, whether via an unrestricted or controlled-access mechanism, are managed through a central repository, the database of Genotypes and Phenotypes (dbGaP), administered by the National Center for Biotechnology Information (NCBI), part of the National Library of Medicine at NIH.

Controlled-access data in NIH-designated repositories (e.g., dbGaP) are made available for secondary research after requesters have submitted a request and obtained approval from an NIH Data Access Committee (DAC) for a particular project. Approved requesters are granted access to millions of dollars of genomic research data without charge.

A.1 Circumstances Making the Collection of Information Necessary

The collection of information activities set forth herein would be conducted under the authorities granted in Section 301 of the Public Health Service Act, 42 USC 241 and in accordance with the policies and procedures set forth in the NIH GDS Policy.

Under the NIH GDS Policy, all investigators who receive NIH funding to conduct research studies generating large-scale human genomic data maintained in controlled-access are expected to register the studies in dbGaP, no matter which NIH-designated data repository will maintain the data. As part of the registration process, investigators must provide basic study information such as the type of data that will be submitted (e.g., individual-level genotype data, aggregate-level association data), a description of the study, and an institutional assurance of the data submission (i.e., an Institutional Certification) which delineates any limitations on the use of the data for secondary research (e.g., data can be used only for research of particular diseases).

Although NIH encourages genomic data sharing through this Policy, circumstances beyond the control of investigators may preclude submission of human genomic data to NIH-designated data repositories (e.g., country or state laws that prohibit data deposition in a U.S. database). In such cases, an exception to submit genomic data to an NIH-designated data repository may be granted. For transparency purposes, when exceptions are granted, studies will still be registered

in dbGaP, the reason for the exception will be included in the registration record, and a reference will be provided to an alternative data-sharing plan or resource, if available.

Requesters interested in using controlled-access human data for secondary research must apply through dbGaP and be granted permission to use the data from a relevant NIH DAC. As part of the application process, requesters and their institution must provide information such as a description of the proposed research use of the data that conforms to the data use limitations, agree to the Genomic Data User Code of Conduct, and agree to the terms of data access through a Data Use Certification agreement.

A.2 Purpose and Use of the Information Collection

This application is for an extension to an existing Office of Management and Budget (OMB) approved collection of information, 0925-0670, to continue supporting NIH's goals for sharing of genomic data. The online forms reduce burden for investigators who submit data to dbGaP and for requesters who wish to access the data for secondary research. Since the last approved PRA, minor changes have been made that focus on reducing redundancy among data submissions and access requests and increasing awareness of the ethical responsibilities associated with the responsible use of genomic information. For example, NIH updated the access model to Genomic Summary Results (*Update to NIH Management of Genomic Summary Results (GSR)* (NOT-OD-19-023)) to increase the accessibility of GSR from most NIH-funded studies in a manner that promotes public benefit from the federal investment in genomics research while considering potential risk to research participants. Therefore, institutions are now expected to indicate in their Institutional Certification if GSR from a submitted study should be provided only though controlled-access or unrestricted access (see Attachment 1, slide 7). Additionally, the NIH observed the input of redundant information by requesters in the *Research Progress* and *Research Plan for the Next Year* sections of the *Project Renewal*. To reduce redundancy, and thus, burden, NIH removed the *Research Plan for the Next Year* section from the Project Renewal forms. These changes are documented in Attachment 2.

We estimate that we will have more respondents, and thus proportional more responses per responded. However, despite the proportional increase in responses as a result of more respondents, the average burden per response will remain the same. Since 2015, the cost to the Government has decreased from $5.99 million to $5.00 million.

Study Registration and Data Submission

The required elements for study registration in dbGaP include basic study information and an Institutional Certification approved by the Institutional Official. Basic study information includes the name of the investigators, funding information, a description of the study and data, and a statement that secondary researchers can use to acknowledge the original data collection. Through the Institutional Certification, the institution assures that study data submission is consistent with all applicable laws, regulations, and institutional policies; outlines data use limitations, if any; certifies that the identities of research participants will not be disclosed to dbGaP; and documents that an Institutional Review Board (IRB) and/or Privacy Board reviewed and approved the plan to submit the data to dbGaP. Data files are submitted prior to registration.

Attachment 1 illustrates the information that will be collected for dbGaP study registration and data submission.

Data Access

The initial data access request (DAR) application to request controlled-access dbGaP data includes information about the requesting investigator's credentials; the proposed use of the data; an agreement to comply with NIH policies and the Genomic Data User Code of Conduct;[1] and certification by the requester and an Institutional Signing Official at the requester's institution that data will not be sold, distributed, re-identified, or used by unauthorized users and collaborators not named in the DAR. Requests to renew data access and reports to close out data use are done annually and are similar to the initial DAR. Additional information obtained in the Project Renewal and Project Close-out forms includes information about publications, presentations, or intellectual property based on the research conducted with the accessed data as well as any data security issues or other data management incidents.

As a prerequisite to applying for controlled-access data, the requester and an Institutional Signing Official must be registered in the NIH electronic Research Administration (eRA) Commons.[2] Most requesters seeking access to dbGaP data have already registered with eRA Commons when applying for NIH funding. For requesters not previously registered in eRA Commons, this step is not expected to add significantly to the burden. Several of the mandatory fields in the DAR form will be automatically filled based on eRA registration information, and most fields in the Project Renewal and Project Close-out forms will be automatically filled from the initial DAR form. Attachment 2 illustrates the investigator and institutional information that will be collected for data access.

A.3 Use of Information Technology and Burden Reduction

NCBI developed an online dbGaP system to register studies, submit data, and request access to data. The online system allows investigators to submit the required information directly to dbGaP, thereby minimizing burden not only for investigators and institutions, but also for NIH staff. The online system uses time-saving features, such as the use of pull-down and scrolling menus to fill data fields, "find as you type" (or "type ahead") functionality, and text fields that allow investigators and requesters to cut and paste information from other sources. Where possible, data fields are automatically filled with information from other data sources, such as eRA Commons, which provides the investigator's or requester's name, institution, Institutional Official, and Institutional Signing Official. Also, information from one online form can be used to pre-fill data fields on subsequent forms. For example, information from an Institutional Certification will be used to pre-fill many data fields in the study registration system, and likewise, information in a DAR will be used to pre-fill many of the data fields on the Project Renewal and Project Close-out forms. In addition, the data submission system has been designed to allow investigators to submit their data in any format based on the genomics platform used. A

[1] NIH Genomic Data User Code of Conduct. See http://osp.od.nih.gov/wp-content/uploads/Genomic_Data_User_Code_of_Conduct.pdf
[2] eRA Commons. See https://commons.era.nih.gov/commons/.

Privacy Impact Assessment has been completed for the online dbGaP system to register studies, submit data, and request access to data (Attachment 3).

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

Currently, no comparable genomic data repository exists. dbGaP is unique in that it provides a single site for all NIH-funded human genomic research studies, as well as to many non-NIH-funded human genomic research studies.

A.5 Impact on Small Businesses or Other Small Entities

The information collection does not have a differential effect on small businesses.

A.6 Consequences of Collecting the Information Less Frequently

Delaying submission of the information to dbGaP would impede secondary research studies of data in dbGaP and lessen the value of NIH's investment.

Following the initial request and approval to use controlled-access data, requesters must provide annual updates on their research progress and renew access to the dataset(s) for another year or close-out access to the dataset(s). The consequence of not submitting the required information annually is a reduction in NIH oversight of dbGaP data, and reduced knowledge of how the data are being used, possibly leading to an increased rate of adverse data management incidents.

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

This collection fully complies with 5 CFR 1320.5.

A.8.1 Comments in Response to the Federal Register Notice

In accordance with 5 CFR 1320.8(d), the proposed information collection was published in the *Federal Register* (FR) on May 1, 2019, (84 FR 18555) and allowed 60 days for public comment. No public comments were received.

A.8.2 Efforts to Consult Outside Agency

NIH engaged with stakeholders on the development of the NIH GDS Policy, an extension of the *NIH Policy for Sharing of Data Obtained in NIH Supported or Conducted Genome-Wide Association Studies* (*GWAS Policy*),[3] during a 60-day public consultation period that involved requests for public comment in the *Federal Register*,[4] request for information in the *NIH Guide for Grants and Contracts*[5] and through a public webinar that involved an open question and answer session.[6] NIH received comments from 107 respondents that included professional societies, scientists, ethicists, IRB administrators, privacy advocates, patient advocacy groups, tribal representatives, and members of the general public. The comments, which reflected a

[3] http://grants.nih.gov/grants/guide/notice-files/NOT-OD-07-088.html
[4] Request for Information (RFI): Draft NIH Genomic Data Sharing Policy. *Federal Register*. 78(183): 57860-57865. See http://www.gpo.gov/fdsys/pkg/FR-2013-09-20/pdf/2013-22941.pdf
[5] Request for Information (RFI): Input on the Draft NIH Genomic Data Sharing Policy. NIH Guide. NOT-OD-13-119. September, 27, 2013. See http://grants.nih.gov/grants/guide/notice-files/NOT-OD-13-119.html.
[6] NIH GDS Policy Public Consultation. See https://webmeeting.nih.gov/p7sqo6avp6j/.

variety of interests and perspectives, were considered in the development of the final Policy and in the governance and oversight procedures established to carry out the policy.[7]

Ongoing opportunities exist for data submitters and requesters to provide feedback to NIH about how the study registration, data submission, and access processes could be improved, such as through annual reporting and Requests for Information (RFI). On February 21, 2017, NIH issued an RFI on *Processes for dbGaP Data Submission, Access and Management* (NOT-OD-17-044). During 45-day comment period, NIH received comments from 43 respondents that included scientists, institutional officials, bioethicists, and members of the public (Attachment 3). This input has helped NIH recognize the need for the continued expansion of automated processes and streamlining the design of the online system for study registration, data submission, and data access. More specifically in response to comments from data submitters and requesters, NIH streamlined access for investigators who must retrieve the data they generated from an NIH-designated data repository for their primary research use, and allowed the use of cloud computing to store and analyze dbGaP controlled-access data.

A.9 Explanation of Any Payment of Gift to Respondents

No gifts or payments are to be offered in regard to this information collection.

A.10 Assurance of Confidentiality Provided to Respondents

The names and institutional affiliations of the respondents (both data submitters and data requesters) will be posted publicly on the dbGaP website, and thus there is no assurance of confidentiality afforded to the respondents. However, it is important to emphasize that no personal information is requested from investigators submitting or accessing data beyond their name and institutional affiliation. Submitters of data to dbGaP are largely NIH-funded investigators whose names and institutional affiliations are already a matter of public record (see http://projectreporter.nih.gov/reporter.cfm). Requesters of the data are both NIH-funded and non-NIH-funded investigators. Making the names of submitters and requesters of the data publicly available is an important ethical underpinning of the NIH GDS Policy as it allows NIH to be transparent in informing research participants, the scientific community, and the general public on how genomic data are being shared, with whom, and for what research purpose in addition to helping to foster future research collaborations.

A System of Records Notice (SORN) 09-25-0036, is in place for eRA Commons, which dbGaP uses for credentialing purposes of submitters and data requestors (https://oma.od.nih.gov/forms/Privacy%20Documents/PAfiles/0036.htm). Federal Register Notice published on September 26, 2002 (Vol. 67, No. 187, page 60751 -https://www.federalregister.gov/articles/2002/09/26/02-23965/privacy-act-of-1974-annual-publication-of-systems-of-records#h-167). Another SORN 09-90-1401 expanding the relevant data access enterprise, (e.g. NIH-designated data repositories and DACs) is in place (https://www.federalregister.gov/documents/2018/03/14/2018-05176/privacy-act-of-1974-system-of-records). Federal Register Notice published on March 14, 2018 (Vol. 83, No.50, page 11214 - https://www.govinfo.gov/content/pkg/FR-2018-03-14/pdf/2018-05176.pdf).

[7] Final NIH GDS Policy. See https://grants.nih.gov/grants/guide/notice-files/not-od-14-124.html

A.11 Justification for Sensitive Questions

No questions of a sensitive nature are included in this data collection.

A.12.1 Estimated Annualized Burden Hours

The burden associated with this information collection is calculated in two parts: the burden associated with registering genomic studies and submitting data to dbGaP; and the burden associated with applying for controlled-access data in dbGaP. Respondents are (1) investigators submitting data to dbGaP; (2) Institutional Officials who assure the data submission; (3) investigators who request and receive data (requesters); and (4) Institutional Signing Officials who certify that they will abide by the NIH GDS Policy.

Study Registration and Data Submission

The burden associated with registering genomic studies and submitting data to dbGaP involves the time and effort necessary for the investigator to complete the online form, and upload the data to dbGaP. The burden for the Institutional Official includes the time and effort to review and approve the registration and data submission, including completion of the Institutional Certification. Because NIH makes it possible to submit the information through tailored, user-friendly online forms, the time and effort is considerably reduced. Moreover, some of the information investigators will need to complete the forms will be automatically provided from eRA Commons and other information can be retrieved from the investigator's NIH grant award.

To estimate the annual number of registrants, NIH averaged the number of study registrations and data submissions received by NCBI in 2018, and the number of registrations anticipated for 2019, based on NIH funding allocated for genomic data research grants that will fall under the NIH GDS Policy. Based on this calculation, NIH estimates 600 respondents per year (300 investigators and 300 Institutional Officials). NIH does not expect the number of registrations to increase significantly in the next three years. Based on prior experience, most investigators do not make changes to their initial registration information, but may add additional datasets. Since dbGaP accepts various data formats, additional datasets submitted after the initial registration for the same project do not require any reformatting, and thus does not add additional burden to the registration process. Therefore, the average frequency of response per investigator is estimated to be once.

Table 12-1 provides the estimated burden hours for registration and data submission of NIH-funded human genomic studies. Based on a simulation of the registration and submission processes using the online forms, NIH estimates that it will take an investigator, on average, 1 hour to enter the study information and upload the study data, and that it will take an Institutional Official, on average, 30/60 hours to certify the information. Multiplying the frequency of response (1) by the total number of investigators (300) by the time it takes an investigator to register a study and submit data (1 hour), yields a total annual hour burden for submitting investigators of 300 hours. The same formula was used to derive the total annual burden for Institutional Officials (i.e., the frequency was multiplied by the number of Institutional Officials by the time it takes an Institutional Official to carry out the certification process). The total

annual hour burden for Institutional Officials is 150 hours. The total annual burden for both groups is 450 hours.

Data Access

On an annual basis, the requester completes only one of the three forms for a given project (i.e., initial, Project Renewal, or Project Close-out). In order to request data from dbGaP, a requester must complete a data access request (DAR) form of the project request. This step also requires the Institutional Signing Official to review the DAR, and both the requester and Institutional Signing Official must certify that they agree to the terms and conditions for use of the data and will adhere to the NIH GDS Policy. Upon approval of their request, requesters are granted access to the data for one year. At the end of the year, investigators who wish to continue to use the data must complete a Project Renewal to extend the access period for another year. Information from the original DAR is automatically incorporated into the requester's Project Renewal form, and the requester provides additional information such as a description of research progress, publications, presentations, and intellectual property that are based on the secondary use of dbGaP data.

Requesters who do not wish to renew their access request are expected to complete a Project Close-out. Information from the original DAR or Project Renewal is automatically incorporated into the Project Close-out form, and the requester provides additional information such as a description of publications, presentations, and intellectual property that are based on the secondary use of dbGaP data. Completion of the Project Close-out provides NIH with information on final project outcomes, and also provides the oversight mechanism by which the Institutional Signing Official confirms that the project has been discontinued and that all copies, versions, and data derivatives of the dataset(s), on both local servers and hardware, have been deleted.

The annual time burden estimates are calculated separately for the initial DAR and Project Renewal/Project Close-out forms. To estimate the number of respondents and the number of requests made per year, NIH totaled the number of respondents and DARs that were submitted over a two-year period and divided by two, for an average of 3000 respondents (1500 requesters and 1500 Institutional Signing Officials) and 2300 DARs per year.

To calculate the number of annual requests per respondent, NIH divided the number of DARs per year by the number of requesters, and estimated two DARs per year. Since each DAR is either renewed or closed out at the end of the one-year access period, the Project Renewal/Project Close-out process also involves 3000 respondents submitting two reports per year.

The burden associated with an initial DAR includes the time and effort necessary for investigators to (1) identify the studies and datasets in dbGaP of interest, (2) prepare information for their proposed research use statement, and (3) complete the DAR form. The burden also involves the Institutional Signing Official's review and certification of the DAR.

Table 12-1 provides the estimated burden hours for completing a DAR. Based on simulations, NIH estimates that it will take a requester an average of 45/60 hours to complete the DAR and 30/60 hours for the Institutional Signing Official to review and certify the DAR. To derive the annual hour burden for requesters, NIH multiplied the number of DARs submitted per year (2) by the number of requesters (1500) by the time it takes a requester to complete a DAR (45/60 hour), for a total of 2250 hours. The same formula was used to derive the total annual burden for Institutional Signing Officials (i.e., the frequency was multiplied by the number of Institutional Signing Officials by the time it takes an Institutional Signing Official to complete the certification process). The total annual hour burden for Institutional Signing Officials is 1500 hours, with a total annual burden of 3750 hours for all respondents.

The burden associated with renewal and close-out involves the time and effort necessary for requesters to (1) prepare information on their research use of NIH datasets, (2) complete the Project Renewal or Project Close-out form, and (3) have the Institutional Signing Official review and approve the form. Importantly, when requesters stop using the data, Institutional Signing Officials verify the data has been deleted from the institution's computer systems, cloud storage, if applicable, and mobile devices. Requesters, who are registered in eRA Commons or those who have made previous data requests or submitted data to dbGaP, will have information from those systems automatically transferred to the Project Renewal or Project Close-out forms, significantly reducing data entry time.

Table 12-1 shows the estimated burden hours for Project Renewal/Project Close-out forms.

Based on simulations, NIH estimated that it will take an average of 15/60 hours for the requester to provide the required information and 18/60 hours for the Institutional Signing Official to review and approve the form and to confirm that data were deleted from the institution's computer system for projects that are closed-out. To derive the annual hour burden for requesters, NIH multiplied the frequency of response (2) by the number of requesters (1500) by the time it takes an investigator to complete a renewal or close-out form (15/60 hour), for a total of 750 hours. The same formula was used to derive the total annual burden for Institutional Signing Officials (i.e., the frequency was multiplied by the number of Institutional Signing Officials by the time it takes an Institutional Signing Official to complete the certification process). The total annual hour burden for Institutional Signing Officials is 900 hours, with a total annual burden of 1650 hours for all respondents.

A.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
Study Registration and Data Submission					

dbGaP Registration and Submission	Investigator Submitting Data	300	1	1	300
Institutional Certification	Institutional Official to Certify Submission	300	1	30/60	150
Requesting Access to Data					
Data Access Request	Requester Submitting Request	1,500	2	45/60	2,250
Data Access Request	Institutional Signing Official to Certify Request	1,500	2	30/60	1,500
Project Renewal or Project Close-out					
Project Renewal or Project Close-out form	Requester Submitting Request	1,500	2	15/60	750
Project Renewal or Project Close-out form	Institutional Signing Official to Certify Request	1,500	2	18/60	900
Total			12,600		5,850

A.12-2 Annualized Cost to Respondents

To estimate the annual cost to respondents, NIH used statistics from the U.S. Bureau of Labor and Statistics 2017, National Occupational Employment and Wage Estimates.[8] The mean hourly

wage of $40.80 for Life Scientists was used for investigators and requesters, and the mean hourly wage of $51.77 for Education Administrators, Postsecondary, was used for Institutional Officials/Institutional Signing Officials. The annual cost burden estimates are calculated separately for the initial DAR and Project Renewal or Project Close-out forms.

Study Registration and Data Submission

Table 12-2 provides the estimated annualized cost to register a study and submit data in dbGaP. To estimate the annual cost to respondents, NIH used figures from the U.S. Bureau of Labor and Statistics 2017, National Occupational Employment and Wage Estimates.[8] The mean hourly wage of $40.80 for Life Scientists was used for investigators, and the mean hourly wage of $51.77 for Education Administrators, Postsecondary, was used for Institutional Officials. The estimated annual cost of $12,240.00 for investigators to register studies and submit data was calculated by multiplying the number of investigators (300) by frequency of response (1) by the average time per response (1 hour) by the hourly wage rate for investigators ($40.80). The estimated annual cost of $7,765.50 for Institutional Officials to certify study registration and data submission was similarly calculated (i.e., multiplying the number of Institutional Officials by the frequency of response by the average time per response by the hourly wage rate). The total annual cost of study registration and data submission is $20,005.50.

Data Access

Table 12-2 provides the estimated annual cost for respondents to complete an initial DAR form. The estimated annual cost of $91,800.00 for requesters to complete an initial DAR was calculated by multiplying the number of requesters (1500) by frequency of response (2) by the average time per response (45/60 hour) by the hourly wage rate for requestors ($40.80). The estimated annual cost of $77,655.00 for Institutional Signing Officials to certify the DAR was similarly calculated (i.e., multiplying the number of Institutional Signing Officials by the frequency of response by the average time per response by the hourly wage rate). The total annual cost for respondents to complete an initial DAR is $169,455.00.

Table 12-2 provides the estimated annualized cost for respondents to complete a Project Renewal or Project Close-out report. The estimated annual cost of $30,600.00 for requesters to complete a Project Renewal or Project Close-out report was calculated by multiplying the number of requesters (1500) by frequency of response (2) by the average time per response (15/60 hour) by the hourly wage rate for requesters ($40.80). The estimated annual cost of $46,593.00 for Institutional Signing Officials to certify the report was similarly calculated (i.e., multiplying the number of Institutional Signing Officials by the frequency of response by the average time per response by the hourly wage rate). The total annual cost of study registration and data submission is $77,193.00.

A.12-2 Annualized Cost to the Respondents

[8] U.S. Bureau of Labor and Statistics. May 2011. National Occupational Employment and Wage Estimates. The investigators and requesters wag rate was obtained from the http://www.bls.gov/oes/current/oes_nat.htm#19-0000 and the Institutional Official or Institutional Signing Official wage rate was obtained from the https://www.bls.gov/oes/current/oes119033.htm.

Type of Respondents	Total Annual Burden Hours	Hourly Respondent Wage Rate[8]	Respondent Cost
Study Registration and Data Submission			
Investigator Submitting Data	300	$40.80	$12,240.00
Institutional Official to Certify Submission	150	$51.77	$7,765.50
Requesting Access to Data			
Requester Submitting Request	2250	$40.80	$91,800.00
Institutional Signing Official to Certify Request	1500	$51.77	$77,655.00
Project Renewal/Project Close-out Process			
Requester Submitting Request	750	$40.80	$30,600.00
Institutional Signing Official to Certify Request	900	$51.77	$46,593.00

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

Other than the respondents time, there are no additional costs associated with this data collection.

A.14 Annualized Cost to the Federal Government

The estimated annualized cost to the Federal Government to support this information collection is $5 million, which is comprised of program personnel costs and computer

hardware/software costs associated with the project's implementation and operation. The estimated personnel cost is $2.50 million, based on 20 programmers (contractors and federal) at an average annual rate of $124,845 (salary and benefits). The estimated cost of computer hardware and software is $2.5 million.

Cost Descriptions	Grade/Step	Salary	% of Effort FTE equiv	Fringe (if applicable)	Total Cost to Gov't
NIH Project Tech Lead [2] [Staff Scientist]	AD-00	$158,647	0.75		$135,876
NIH Content Team Lead [1] [Staff Scientist]	AD-00	$164,200	0.52		$113,663
NIH Content Specialist [5] [Staff Scientist]	AD-00	$123,833	4.45		$725,467
NIH Software Developer [2] [Staff Scientist]	AD-00	$128,819	0.4		$68,928
NIH Data Specialist [3] [Staff Scientist]	AD-00	$118,556	2.4		$369,133
Contractor Content Specialist [2]		$148,914	2		$297,828
Contractor Software Developer [6]		$191,348	3.55		$715,254
Contractor Project Manager [1]		$176,877	0.4		$70,751
Travel					
Other Cost					
Computer Hardware and Software					$2,500,000
Total					$4,996,900

A.15 Explanation for Program Changes or Adjustments

This application is for a revision to an existing Office of Management and Budget (OMB) approved collection of information, 0925-0670. Changes in Attachments 1 and 2 are described below and noted in red boxes in the attachments. Some changes focus on

reducing redundancy among data submissions and access requests and increasing awareness of the ethical responsibilities associated with the responsible use of genomic information.

In the dbGaP registration and submission forms (Attachment 1): additional questions on publication embargo deleted because this is no longer applicable (slide 4), added ability for investigator to add additional investigators that are notified when a submitted dataset is used (slide 4), added ability to review previously entered submission information for accuracy (slide 4), added ability to auto-create the DUC from existing information in the submission package (slide 5), added ability to auto-create the DUC from existing information in the submission package (slide 6), added a column for "Collaborating Site" that automatically adds additional rows to form when information is entered for all rows and columns (slide 7), and created new ability for investigators and Institutional Officials to digitally sign form (slide 8).

In the dbGaP data access, renewal, and close-out forms (Attachment 2): information on the login page have been reorganized and consolidated to make the application information more apparent and prominent (slide 3), a pop-up reminder has been added to the "create decryption password" field to alert requester additional information is needed (slide 7), the "Research Plans with Approved Dataset(s)" field has been removed as it is redundant with information provided in the "Research Progress" and "Research Plans with Approved Datasets" sections (slide 22).

We estimate that we will have more respondents, and thus proportional more responses per responded. However, despite the proportional increase in responses as a result of more respondents, the average burden per response will remain the same.

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

Information submitted to dbGaP to register studies or access genomic data will be made available to the public via the dbGaP website once the registration is complete and the study has been released. The only submitted information that will not be publicly displayed on the website is contact information of the submitting investigators and requesters (i.e., the investigator's email addresses, and phone and fax numbers).

NIH will post on their public website, on a bi-annual basis, an analysis of data usage that would include summary statistics such as the number of registered studies, the number of submitted/approved DARs, trends in usage, and number of publications.

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

No exemption is requested.

A.18 Exceptions to Certification for Paperwork Reduction Act Submissions

No exceptions are requested.