Supporting Statement A

Paltoo, Dina (NIH/OD) [E]
NIH\OD
[Pick the date]

Supporting Statement A for

National Institutes of Health Information Collection Forms to
Support Genomic Data Sharing for Research Purposes (OD)
OMB Number 0925-0670 expiration Date 03/31/2016

Date: 2/23/2016

Name: Dina N. Paltoo, PhD, MPH

Address: 6705 Rockledge Dr., Suite 750, Bethesda, MD 20892

Telephone: 301-496-9838

Fax: 301-496-9839

Email: paltood@od.nih.gov

Check off which applies:

X New Revision

- Reinstatement with Change
- Reinstatement without Change
  Extension
- Emergency
- Existing

Table of Contents

A. Justification

Sharing of research data is an integral element of 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 and 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 NIH Genomic Data Sharing Policy (GDS Policy).[1] The GDS Policy, an extension of the 2008 *NIH Policy for Sharing of Data Obtained in NIH Supported or Conducted Genome-Wide Association Studies* (the GWAS Policy), applies to all 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 GDS Policy addresses data submission, data maintenance, and data 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 data repository. Human genomic data submissions and access, whether via an unrestricted-or controlled-access mechanism, are managed through a central data 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 data repositories such as dbGaP are made available for secondary research only after investigators have submitted a request and obtained approval from a NIH Data Access Committee (DAC) for a particular project. Approved investigators are granted access to millions of dollars of genomic research data free of charge.

We are requesting a revision with a three year approval of this information collection in order to continue supporting NIH's goals for sharing of genomic data.

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 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).

[1] NIH Genomic Data Sharing Policy. See http://gds.nih.gov/PDF/NIH_GDS_Policy.pdf.

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 in a 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.

Investigators interested in using controlled-access human data for secondary research must apply through dbGaP and be granted permission to use the data from the relevant NIH DACs. As part of the application process, investigators 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.

This application is for a revision to an existing Office of Management and Budget (OMB) approved collection of information, 0925-0670 expiration Date 03/31/2016. Changes are documented in Attachment 1 and 2, but mainly focus on implementation of mechanisms to ease completion of necessary information and provide greater consistency among submissions and access requests.

A.2 Purpose and Use of the Information Collection

We are requesting a revision of this information collection in order 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 implementation of mechanisms to ease completion of necessary information and provide greater consistency among submissions and access requests. For example the submission and access systems now contains categories or options for certain responses instead of free text, in an attempt to ease completion of materials and provide greater consistency among submission. These changes are documented in Attachment 1 and 2.

We estimate that we will have more respondents, but the number of responses per respondent and the average burden per response will remain the same. The cost to the Government has decreased from $9.25 million to $5.99 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. Attachments 1 and 2 illustrate the information that will be collected for dbGaP study registration and data submission.

Data Access

The 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 NIH Genomic Data User Code of Conduct;[2] and certification by the investigators and institution a Signing Official at the investigator's institution that data will not be sold, distributed, re-identified, or used by unauthorized users and collaborators not named in the data access request (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 renewal and closeout requests 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 investigator and Signing Official must be registered in the NIH electronic Research Administration (eRA) Commons.[3] Most investigators seeking access to dbGaP data have already registered with eRA Commons when applying for NIH funding. For investigators 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 renewal and close-out request 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

In lieu of using time-consuming paperwork traditionally completed by investigators and then submitted to NIH staff, NCBI has created an online system in dbGaP 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 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 name, institution, 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

[2] NIH Genomic Data User Code of Conduct. See http://gds.nih.gov/pdf/Genomic_Data_User_Code_of_Conduct.pdf
[3] eRA Commons. See https://commons.era.nih.gov/commons/.

data fields on the renewal and 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.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 12/1/2015 vol. 80 page 75120 and allowed 60 days for public comment. A public comment was received.

NIH Genomic Data Sharing (GDS) Policy governance has discussed the comment regarding the consideration of use of a single Institutional Certification form with a footnote stating the expectations for materials obtained prior to or after January 25, 2015. Based on the discussion and in working with institutions at different levels of experience in submitting the necessary information and data, and in an effort to decrease burden on and confusion of these institutions, NIH has decided not to use one form. NIH has provided more information on its GDS website regarding the Institutional Certification and clarified when each type should be used by both the intramural and extramural research community. Please see: https://gds.nih.gov/Institutional_Certifications.html.

Regarding the inclusion/clarification of additional language in the Institutional Certification clarifying the IRB's limited review capacity, we acknowledge that IRBs may have limited review capacity to certify for all tribal and other state laws, and confirm that the insertion of the phrase "as appropriate" is intended to accommodate such limitations.

NIH has developed "Points to Consider for Institutions and Institutional Review Boards in Submission and Secondary Use of Human Genomic Data under the National Institutes of Health Genomic Data Sharing Policy," which was recently made available (June 2016) on the NIH GDS Policy website at: https://gds.nih.gov/pdf/GDS_Points_to_Consider_for_Institutions_and_IRBs.pdf.

Information regarding IRB review of proposals for data submission and sharing included in funding applications is found on page 2. Additionally, on pages 3-4, the document also discusses risks associated with genomic data submission and broad sharing of these data.

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 GWAS Policy[4], during a 60-day public consultation period which involved requests for public comment in the *Federal Register*,[5] request for information in the *NIH Guide for Grants and Contracts*[6] and through a public webinar that involved an open question and answer session.[7] NIH received comments from 107 commenters[8] 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 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.

Ongoing opportunities exist for data submitters and requesters to provide feedback to NIH about how the study registration, and data submission access processes could be improved, such as through annual reporting. This input has helped NIH recognize the need for automating the process and shaping the design of the online system for study registration, data submission, and data access. More specifically in response to comments from data submitters and users NIH has developed standard categories for consent groups, an auto-populated Data Use Certification, fillable PDF version of the Institutional Certification, and 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.

---

[4] http://grants.nih.gov/grants/guide/notice-files/NOT-OD-07-088.html

[5] 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

[6] 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..

[7] NIH GDS Policy Public Consultation. See https://webmeeting.nih.gov/p7sqo6avp6j/.

[8] Compiled Public Comments on the Draft NIH GDS Policy. See http://gds.nih.gov/pdf/GDS_Policy_Public_Comments.pdf.

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). Approved users of the data are both NIH-funded and non-NIH-funded investigators. Making the names of submitters and users 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 expanding the relevant data access enterprise, e.g. NIH-designated data repositories and DACs, is under development.

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; and (4) institutional Signing Officials who certify that they will abide by the 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 2015, and the number of registrations anticipated for 2016, based on NIH funding allocated for genomic data research grants that will fall under the GDS Policy. Based on this calculation, NIH estimates 300 respondents per year (150 investigators and 150 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, 0.50 hours to certify the information. Multiplying the frequency of response (1) by the total number of investigators (150) 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 150 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 75 hours. The total annual burden for both groups is 225 hours.

Data Access

The investigator completes two of the three forms for a given project (i.e., initial and renewal data request or a close-out report). In order to request data from dbGaP, an investigator must complete a DAR form. This step also requires the Signing Official to review the DAR, and both the investigator and Signing Official must certify that they agree to the terms and conditions for use of the data and will adhere to the GDS Policy. Upon approval of their request, investigators 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 renewal form to extend the access period for another year or complete a project close out. Information from the original DAR is automatically incorporated into the investigator's renewal and close-out forms, and the investigator 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.

Investigators who do not wish to renew their access request and complete a project close-out form additionally must provide NIH with information on final project outcomes and confirmation that the project has been discontinued and the data were deleted from the institution's computers.

The annual time burden estimates are calculated separately for the initial DAR and renewal/project close-out forms, however these would be the same individuals. To estimate the number of data requesters (respondents) and the number of requests made per year, NIH totaled the number of requesters and DARs that were submitted over a two-year period and divided by two, for an average of 633 respondents and 933 DARs per year.

To calculate the number of annual requests per respondent, NIH divided the number of DARs per year by the number of respondents, and estimated two DARs per year, per respondent. Since each DAR is either renewed or closed out at the end of the one-year access period, the renewal/close-out process also involves 633 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 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 0.75 hours to complete the DAR and 0.50 hours for the Signing Official to review and certify the DAR. To derive the annual hour burden for investigators, NIH multiplied the number of DARs submitted per year (2) by the number of investigators (633) by the time it takes an investigator to complete a DAR (0.75 hour), for a total of 950 hours. The same formula was used to derive the total annual burden for Signing Officials (i.e., the frequency was multiplied by the number of Signing Officials by the time it takes a Signing Official to complete the certification process). The total annual hour burden for Signing Officials is 633 hours, with a total annual burden of 1,583 hours for all respondents.

The burden associated with renewal and project close-out involves the time and effort necessary for requesters to (1) prepare information on their research use of NIH datasets, (2) complete the renewal or close-out form, and (3) have the Signing Official review and approve the form. Importantly, when investigators stop using the data, NIH expects them to delete the datasets from the institution's computer systems and mobile devices, and Signing Officials are expected to verify that the data were deleted. Investigators, 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 renewal/close-out forms, significantly reducing data entry time.

Table 12-1 shows the estimated burden hours for renewal/project close-out forms.

Based on simulations, NIH estimated that it will take an average of 0.25 hours for the investigator to provide the required information and 0.30 hours for the 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 investigators, NIH multiplied the frequency of response (2) by the number of investigators (633) by the time it takes an investigator to complete a renewal or close-out form (0.25 hour), for a total of 317 hours. The same formula was used to derive the total annual burden for Signing Officials (i.e., the frequency was multiplied

by the number of Signing Officials by the time it takes a Signing Official to complete the certification process). The total annual hour burden for Signing Officials is 380 hours, with a total annual burden of 697 hours for all respondents.

A.12-1 Estimated Annualized Burden Hours

Form Name	Type of Respondent	Number of Respondents	Number of Responses per Respondent	Average Burden Per Response (in hours)	Total Annual Burden Hour
Study Registration and Data Submission	Investigator Submitting Data	150	1	1	150
	Institutional Official to Certify Submission	150	1	30/60	75
Requesting Access to Data	Investigator Requesting Data	633	2	45/60	950
	Signing Official to Certify Request	633	2	30/60	633
Renewal/project Close-out Process	Investigator Requesting Data	633 (same individuals as listed above)	2	15/60	317
	Signing Official to Certify Request	633 (same individuals as listed above)	2	18/60	380
Grand Total		1,566	5,064		2,505

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 2014, National Occupational Employment and Wage Estimates.[8] The mean hourly wage of

$38.11 for Life Scientists was used for investigators, and the mean hourly wage of $48.99 for Education Administrators, Postsecondary, was used for Institutional/Signing Officials. The annual cost burden estimates are calculated separately for the initial DAR and renewal/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 2014, National Occupational Employment and Wage Estimates.[9] The mean hourly wage of $38.11 for Life Scientists was used for investigators, and the mean hourly wage of $48.99 for Education Administrators, Postsecondary, was used for Institutional Officials. The estimated annual cost of $5,716.50 for investigators to register studies and submit data was calculated by multiplying the number of investigators (150) by frequency of response (1) by the average time per response (1 hour) by the hourly wage rate for investigators ($38.11). The estimated annual cost of $3,674.25 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 $9,390.75.

Data Access

Table 12-2 provides the estimated annual cost for respondents to complete an initial DAR form. The estimated annual cost of $36,204.50 for investigators to complete an initial DAR was calculated by multiplying the number of investigators (633) by frequency of response (2) by the average time per response (0.75 hour) by the hourly wage rate for investigators ($38.11). The estimated annual cost of $31,010.67for Signing Officials to certify the DAR was similarly calculated (i.e., multiplying the number of 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 $67,215.17.

Table 12-2 provides the estimated annualized cost for respondents to complete a renewal or close-out report. The estimated annual cost of $12,080.87 for investigators to complete a renewal or close-out report was calculated by multiplying the number of investigators (633) by frequency of response (2) by the average time per response (0.25 hour) by the hourly wage rate for investigators ($38.11). The estimated annual cost of $18,616.20 for Signing Officials to certify the report was similarly calculated (i.e., multiplying the number of 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 $30,697.07.

[9]

A.12-2 Annualized Cost to the Respondents

	Type of Respondent	Total Burden Hours	Hourly Wage Rate[8]	Respondent Cost
Study Registration and Data Submission				
	Investigator Submitting Data	150	$38.11	$5,716.50
	Institutional Official to Certify Submission	75	$48.99	$3,674.25
Requesting Access to Data				
	Investigator Requesting Data	950	$38.11	$36,204.50
	Signing Official to Certify Request	633	$48.99	$31,010.67
Renewal/project Close-out Process				
	Investigator Requesting Data	317	$38.11	$12,080.87
	Signing Official to Certify Request	380	$48.99	$18,616.20

[9] U.S. Bureau of Labor and Statistics. May 2011. National Occupational Employment and Wage Estimates. See http://www.bls.gov/oes/current/oes_nat.htm#19-0000 and http://www.bls.gov/oes/current/oes_nat.htm#11-0000.

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 $5.99 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 $3.49 million, based on 26 programmers (contractors and federal) at an average annual rate of

$134,000 (salary and benefits). The estimated cost of computer hardware and software is $2.5 million.

Staff	Grade/Step	Salary	% of Effort	Fringe (if applicable)	Total Cost to Gov't
Project Team Lead [1] [Section Leader – Senior Scientist]	AD-00	$210,000	0.270		$56,700
Project Team Lead [1] [Section Leader – Staff Scientist]	AD-00	$185,000	.1000		$18,500
NIH Content Specialist [6] [Staff Scientist]	AD-00	$130,000	5.2		$676,000
NIH Software Developer [7] [Staff Scientist]	AD-00	$136,000	6.15		$836,400
Contractor Content Specialist [5] [2MSC & 3 ComputerCraft]		$170,000	5		$850,000
Contractor Software Developer [6] [5 MSC & 1 ComputerCraft]		$194,000	5.4		$1,047,600
Travel					
Other Cost					
Computer Hardware and Software					$2,500,000

A.15 Explanation for Program Changes or Adjustments

This application is for a revision to a previous approved collection of information. Since the last approved PRA, minor changes have been made that focus on implementation of mechanisms to ease completion of necessary information and provide greater consistency among submissions and access requests. For example the submission and access systems now contains categories or options for certain responses instead of free text, in an attempt to ease completion of materials and provide greater consistency among submission. These changes are documented in Attachment 1 and 2.

We estimate that we will have more respondents, but the number of responses per respondent and the average burden per response will remain the same. The cost to the Government has decreased from $9.25 million to $5.99 million.

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.