The Study to Explore Early Development (SEED)
Phase 3

OMB Control # 0920-1171
Expiration date: 3/31/2020

Supporting Statement: Part A

August 8, 2019

Contact:
Karen Pazol, PhD, MPH
Health Scientist
Centers for Disease Control and Prevention
Email: kpazol@cdc.gov
Phone: 770-488-6305
Fax: 770.498.1541

Table of Contents

List of Attachments

10 Invitation Packets
10a Invitation Letter – ASD
10b Invitation Letter – DD
10c Invitation Letter – POP
10d Study Brochure – ASD
10e Study Brochure – DD
10f Study Brochure – POP
10g Response Card
11 Invitations Calls
11a Invitation Call Script, Previous ASD diagnosis
11b Invitation Call Script, DD
11c Invitation Call Script, POP
12 Enrollment Packets
12a Enrollment Packet Cover Letter ASD and POP
12b Enrollment Packet Cover Letter DD
12c Study Flow for Participant SEED 3
12d Study Bill of Rights
13 Follow-up Call Items Checklist
14 Follow-up Call 2 Items Self/Parent Administered Forms
15 Consent Forms
15a Overall Consent Document – ASD
15b Overall Consent Document – DD
15c Overall Consent Document – POP
16 Privacy Impact Assessment
16a PIA Approval Form
16b PIA Full Response to Item 37
17 SEED Data Confidentiality and Security Policy
18 Certificate of Confidentiality
18a Signed Certificate of Confidentiality, CDC
18b Signed Certificate of Confidentiality, CO
18c Signed Certificate of Confidentiality, MD
18d Signed Certificate of Confidentiality, MO
18e Signed Certificate of Confidentiality, NC
18f Signed Certificate of Confidentiality, WI
19 IRB Approval Letter
20 SEED Semi-annual Newsletter Sample
21 SEED Eligibility Criteria
22 ICD DSM Special Education code lists

- **Goal of the study:** SEED is a multi-site case-control study designed to investigate risk factors for Autism Spectrum Disorders (ASDs) and the health and behavioral characteristics of children with ASDs. SEED is specifically focused on preconception, prenatal, perinatal and early postnatal risk factors with primary emphasis on genetic, infectious, immunological, hormonal, and obstetric risk factors. Additional SEED focus areas include characterization of distinct ASD behavioral phenotypes and assessment of ASD health impacts, particularly gastrointestinal disorders and symptoms.

- **Intended use of the resulting data:** The data from SEED 3 will be combined with data from the first two SEED phases to enable investigators to conduct in depth analyses of ASD risk factors, including assessment of potential etiologic subgroups and gene-environment interactions. SEED findings will inform the public about potential causes of ASD. It will have implications for primary prevention, the health of children with ASD, and secondary prevention of associated sequelae.

- **Methods to be used to collect:** Data will be collected once from participant families using multiple methods including: 1) maternal telephone interviews with questions about maternal reproductive history and pregnancy with the index child, 2) parent-completed questionnaires about parental and child health and child development, 3) in-person child developmental evaluations, 4) maternal and child anthropometry measurements, and 5) biosampling from biological parents and child.

- **The subpopulation to be studied:** SEED 3 will enroll children who were born between 2014 and 2017 and are aged 2-5 at enrollment and their parents. Children will be identified at 6 participating sites (5 sites funded through a competitive funding process and 1 site in metropolitan Atlanta managed by CDC). Three groups of children will be included: children with ASDs (ASD case group – defined based on SEED developmental assessment), children with other developmental (non-ASD) conditions (DD comparison group), and children from the general population (POP comparison group). Potential ASD and DD children will be identified from multiple health and education sources at each site. POP children will be sampled from site birth certificates.

- **How data will be analyzed:** Children with ASDs will be compared to the two control groups (DD and POP). Odds ratios with 95% confidence intervals will be calculated for associations between ASD and various risk factors: overall; for etiologic ASD subgroups defined empirically based on analyses of the detailed behavioral and other phenotypic data; within strata defined by key demographic characteristics, such as child sex, race-ethnicity and maternal age; and after adjustment for demographic and perinatal characteristics. Interactions between genetic and non-genetic risk factors will be assessed through stratification and modelling. Child health characteristics will be assessed similarly through comparison of children with ASD to children in the DD and POP comparison groups.

Supporting Statement Part A.

A. Justification

A.1. Circumstances Making the Collection of Information Necessary

This Information Collection Request (ICR) is submitted under the classification "Extension." The length of data collection requested for Office of Management and Budget (OMB) approval is 2 years. "The Study to Explore Early Development (SEED): Child Development and Autism" was developed under the National Center on Birth Defects and Developmental Disabilities (NCBDDD) at CDC. NCBDDD at CDC is making this request as authorized by Section 301(a)[42 U.S.C. Section 241(a)] and 317(c) of the Public Health Service Act [42 U.S.C. 247b-4] (**Attachment 1a);** the Children's Health Care Act of 2000, Pub. Law No. 106-310 (**Attachment 1b**); the Combating Autism Act of 2006, Pub. Law No. 109-416 (**Attachment 1c**); and the Autism Collaboration, Accountability, Research, Education, and Support (CARES) Act of 2014, Pub. Law No. 113–157 (**Attachment 1d**).

The Children's Health Care Act of 2000 mandated CDC to establish autism surveillance and research programs to address the number, incidence, correlates, and causes of autism. Under the provisions of this act, NCBDDD implemented the "The Study to Explore Early Development". OMB first approved data collection for this study in October 2007 (OMB 0920-0741). The second phase of the effort (SEED 2) was inappropriately granted an OMB-PRA clinical research exemption. Most recently, information collection was approved for the third phase of this study (SEED 3) in March 2017 (OMB 0920-1171). We seek a 2-year extension of OMB-PRA approval so that data collection may continue beyond the current expiration date of 03/31/2020.

The overall purpose of SEED is to investigate risk factors for autism spectrum disorders ASD and symptom subgroups of ASD, using a case-control study design that includes ascertainment of case and comparison groups that represent diverse population subgroups. The prevalence of ASD, measured through the Autism and Developmental Disabilities Monitoring Network most recently in 2014, is estimated at 16.8 per 1,000 8-year-old children (1 in 59).[1] Apart from the identification of some rare genetic conditions that are commonly associated with autism,[2] causal mechanisms for the disorder largely remain unknown. While numerous genetic factors have been implicated in the etiology of ASDs and sibling and twin studies suggest high heritability,[3,4] the specific genetic mechanisms appear complex and research gaps remain. Moreover, select prenatal environmental factors and adverse perinatal outcomes have also been

---

1 Baio J. et al. 2018. Prevalence of Autism Spectrum Disorder Among Children Aged 8 Years — Autism and Developmental Disabilities Monitoring Network, 11 Sites, United States, 2014. MMWR Surveillance Summaries: 67(6);1–23. https://www.cdc.gov/mmwr/volumes/67/ss/ss6706a1.htm.

2 Miles, J. H. (2011). Autism spectrum disorders—A genetics review. Genetics in Medicine, 13(4), 278–294. http://doi.org/10.1097/GIM.0b013e3181ff67ba

3 Risch et al. 2014. Familial recurrence of autism spectrum disorder: evaluating genetic and environmental contributions. Am J Psychiatry. 2014 Nov 1;171(11):1206-13. https://www.ncbi.nlm.nih.gov/pubmed/24969362.

4 Hallmayer, J. (2011). Genetic Heritability and Shared Environmental Factors Among Twin Pairs With Autism. Archives of General Psychiatry, 68(11), 1095. http://doi.org/10.1001/archgenpsychiatry.2011.76.

associated with ASDs.[5,6,7] The composite evidence supports the likelihood of gene-environment interactions.[2,4]

In the face of these considerable research gaps, large population-based epidemiologic studies of ASD etiology are lacking. The data collection for SEED 3 addresses this critical need. While the composite sample from SEED 1 and 2 allows investigators to address many important research questions, and has resulted in a number of peer reviewed publications stemming from the analysis of this data (**Attachment 2**), it is insufficient for many analyses of interest, including examination of the many important risk factors with a prevalence of 1-5%, that are modestly associated with ASD (Odds ratio, OR<2.5) and associations with ASD subgroups. SEED 3 will increase the composite sample to ≥ 2,000 children in all groups – this will include an ASD case group and two comparison groups (children with other developmental disabilities [DD] and children from the general population [POP]). SEED 3 will thus, expanding our ability to analyze rare exposures and/or modest (yet scientifically important) associations, assess associations within key ASD subgroups, and explore potential effect modifications between various ASD risk factors, including the potential for gene-environment interactions. To reach a final sample exceeding 2,000 children in each of the tree study group, a target goal of 702 participants has been set for each group. Data collection began upon initial approval for SEED 3 in March 2017 (OMB 0920-1171). As of April 4, 2019, 229 participants in the ASD group, 435 participants in the POP group, and 386 participants in the DD group. The number of participants who have completed all steps of SEED 3 (N=1050) is therefore currently at just over half of our final target goal (N=2,106), and will require data collection past the current approval of March 31, 2020.

A.2. Purpose and Use of Information Collection

In planning and designing SEED, study investigators conducted an extensive review of the literature[8] and based on their findings, designated 6 primary research domains: 1) investigation of the ASD phenotype; 2) assessment of genetic risk factors and genetic differences in children with and without ASD; 3) assessment of prenatal infection and immunologic risk factors; 4) assessment of reproductive and hormonal risk factors (subsequently split into two domains – hormonal and obstetric risk factors); 5) assessment of child health with a focus on gastrointestinal symptoms; and 6) assessment of sociodemographic features associated with ASD. Additionally, several areas were designated as secondary research domains: assessment of substance use during pregnancy; assessment of maternal and paternal occupational exposures before and during pregnancy; assessment of maternal environmental exposures before and during pregnancy; assessment of hospitalizations and injuries of the child; and assessment of sleep disorders in the child. Secondary research domains were considered important areas with notable research gaps; yet they were prioritized below the primary research domains. The data collection protocol was designed to ascertain data on both primary and secondary research domains; however, more extensive data are collected to answer research questions pertaining to

---

[5] Schieve L, et al. 2015. Comparison of perinatal risk factors associated with autism spectrum disorder (ASD), intellectual disability (ID), and co-occurring ASD and ID. J Autism Dev Disord. 2015 Aug;45(8):2361-72. https://www.ncbi.nlm.nih.gov/pubmed/25739693

[6] Schieve L, et al. 2011. Have secular changes in perinatal risk factors contributed to the recent autism prevalence increase? Development and application of a mathematical assessment model. Ann Epidemiol. 21(12):930-45. https://www.ncbi.nlm.nih.gov/pubmed/22000328.

[7] Talbott E, et al. Air toxics and the risk of autism spectrum disorder: the results of a population based case-control study in southwestern Pennsylvania. Environ Health;14:80. https://www.ncbi.nlm.nih.gov/pubmed/26444407.

[8] Newschaffer C, et al. 2007. The epidemiology of autism spectrum disorders. *Annual Review of Public Health*, *28*(1), 235–258. http://doi.org/10.1146/annurev.publhealth.28.021406.144007.

the primary research domains. The original primary and secondary scientific interests of SEED were retained in SEED Phase 2 and SEED Phase 3.

The data from SEED 3 will be combined with data from the first two SEED phases to enable investigators to conduct in depth analyses of ASD risk factors including assessment of potential etiologic subgroups and gene-environment interactions. Altogether over 1,400 children with ASD are included in the combined SEED 1 and SEED 2 dataset, with approximately equal numbers in the two control groups. While this existing sample will allow us to address many important research questions within our primary and secondary research domains, the sample size and corresponding statistical power will not be adequate for many analyses of infrequent exposures or exposure subtypes (such as looking at type of maternal infection rather than broadly assessing maternal infection as a risk factor), ASD subtypes, and genetic associations.

The findings from analyses of SEED 1, 2, and 3 data will be published in peer-reviewed journals and presented at national scientific and public health meetings and at local community meetings at each site. A list of manuscripts that have been published to date using SEED 1 and SEED 2 data is included in **Attachment 2**. CDC also prepares summaries of key findings from these studies written in plain language so as to be accessible to the general public and makes them available on our website.[9] CDC also prepare webinars and reports detailing SEED findings for partner organizations and stakeholders. In addition, for participants who consent, identifying information is collected to maintain contact with the participants throughout the course of the study and for future contact for a follow-up study (The Study to Explore Early Development - Teen Follow-up Study, OMB 0920-1219).

As in SEED 1 and SEED 2, data for SEED 3 are collected from participant families using multiple methods, including: 1) maternal telephone interview with questions about maternal reproductive history and pregnancy with the index child **(Attachment 4)** parent-completed questionnaires about parental and child health and child development **(Attachment 6a-h)**, 3) in-person child developmental evaluation **(Attachment 7a-g)**, and 5), and anthropometric and biosampling from biological parents and child **(Attachment 8a-d)**. However, for the last approved OMB submission, the SEED 3 data collection protocol was reduced in comparison to SEED 1 and 2 in order to:

- Eliminate some of the most time-intensive data collection instruments; and
- Reduce processing and storage costs for the SEED Biorepository.

In reducing the data collection protocol, CDC carefully considered which instruments were no longer necessary because sufficient sample size had already been achieved in earlier phases of SEED to answer the specific relevant research questions pertaining to the instrument. Additionally, the SEED 3 data collection protocol was streamlined for the DD comparison group. The rationale for this reduction is that the data already collected in SEED 1 and 2 will be sufficient to address many important research questions needing this particular comparison group.

Specific data collection instruments that were retained in SEED 3 and justifications for their need are included below:

- **Social Communication Questionnaire (SCQ)**, administered to mothers of all children upon entry into the study (**Attachment 3**). This instrument is key in determining the data collection workflow for SEED participants upon enrollment in the study. Additionally data collected on the SCQ inform many analyses subsumed under the ASD Phenotype

[9] https://www.cdc.gov/ncbddd/autism/seed-research.html.

research domain; these data analyses are important for informing the scientific and clinical community on the range of characteristics and behaviors exhibited among children on the autism spectrum and in defining potential etiologic ASD subtypes to be used in other SEED analyses.

- **Maternal Interview (MI) and accompanying Pregnancy Reference Form (PRF)**, administered to mothers of all enrolled children – The MI (**Attachment 4**) is a telephone-assisted interview that includes questions on maternal reproductive history, maternal health and behaviors during the index pregnancy, maternal occupational history, and family demographics. The PRF (**Attachment 5a-b**) is a short questionnaire administered to the mother via phone in advance of the MI; it is designed to estimate dates pertaining to the index pregnancy and the breastfeeding periods. The MI and PRF instruments provide risk factor data for several primary and secondary research domains: Prenatal Infection and Immunologic Factors; Prenatal Hormonal Risk Factors; Obstetric Risk Factors; Sociodemographic Factors; Pregnancy Substance Use; and Maternal Occupational Exposures.

Self/Parent-Administered Forms (Attachments 6):

- **Maternal Medical History Form,** collected from mothers of all enrolled children **(Attachment 6a).** This instrument provides risk factor data for several primary research domains: Genetics (family history data on parental developmental and psychiatric conditions); Preconception and Prenatal Infection and Immunologic Factors; Preconception and Prenatal Hormonal Risk Factors; and Obstetric Risk Factors.
- **Paternal Medical and Occupational History Form,** collected from fathers (or mothers serving as respondent) of all enrolled children **(Attachment 6b).** This instrument provides risk factor data for several primary and secondary research domains: Genetics (family history data on parental developmental and psychiatric conditions); Preconception Immunologic Factors; and Paternal Occupational Exposures.
- **Child Health History Form,** collected from mothers of all enrolled children **(Attachment 6c).** This instrument provides risk factor data for several primary and secondary research domains: Genetics (sibling history collected in addition to index child for various health conditions); Preconception, Prenatal, and Postnatal Immunologic Factors; Sociodemographic Factors (specifically child health insurance which is not captured on the MI); Child Gastrointestinal Symptoms; Child Hospitalizations and Injuries; Child Sleep Disorders.
- **Maternal and Child Residential History Form,** collected from mothers of all enrolled children, **(Attachment 6d).** This instrument provides data that can be geocoded and linked to other data sources such as environmental monitoring data and Census data to examine risk factors related to Maternal Environmental Exposures and Maternal and Child Sociodemographic Features.
- **Child Behavior Checklist (CBCL),** collected from mothers of children in ASD and POP groups **(Attachment 6e).** This instrument informs many analyses subsumed under the ASD Phenotype research domain. Additionally, this instrument provides important information on the POP comparison group – specifically, information about the degree to which this group is typically developing versus experiencing symptoms consistent with developmental impairments.

- **Child Social Responsiveness Scale,** collected from mothers of children in ASD and POP groups **(Attachment 6f-g).** This instrument informs many analyses subsumed under the ASD Phenotype research domain. Additionally, this instrument provides important information on the POP comparison group – specifically, information about the degree to which this group is typically developing versus experiencing symptoms consistent with developmental impairments.
- **Child Services and Treatment Questionnaire,** collected from mothers of children in ASD group **(Attachment 6h).** This instrument informs many analyses subsumed under the ASD Phenotype research domain. Additionally, this instrument informs separate analyses of service needs and access to services among children with ASDs.

Clinical in-person developmental evaluation (ASD and POP workflows only; Attachment 7):

- **Mullen Scales of Early Learning,** collected from mothers of children in ASD and POP groups **(Attachment 7a).** This instrument informs many analyses subsumed under the ASD Phenotype research domain. Additionally, this instrument provides important information on the POP comparison group – specifically, information about the degree to which this group is typically developing versus experiencing symptoms consistent with developmental impairments.
- **Vineland Adaptive Behavioral Scales (VABS),** collected from mothers of children in ASD and POP groups **(Attachment 7b-c).** This instrument informs many analyses subsumed under the ASD Phenotype research domain. Additionally, this instrument provides important information on the POP comparison group – specifically, information about the degree to which this group is typically developing versus experiencing symptoms consistent with developmental impairments.
- **Autism Diagnostic Observation Schedule (ADOS),** collected from mothers of children in ASD group **(Attachment 7d-f).** This instrument is used along with the ADI-R to determine the final study group classification for children in the ASD work flow. Additionally data collected on the ADOS inform many analyses subsumed under the ASD Phenotype research domain.
- **Autism Diagnostic Interview-Revised (ADI-R),** collected from mothers of children in ASD group **(Attachment 7g).** This instrument is used along with the ADOS to determine the final study group classification for children in the ASD work flow. Additionally data collected on the ADI-R inform many analyses included within the ASD Phenotype research domain.
- **Anthropomorphic and Saliva Specimens,** collected from mothers and children in ASD and POP groups **(Attachment 8a-d).** This collection provides risk factor data for the Genetics research domain and also informs gene-environment interaction studies.

In addition to the data collected from the SEED study instruments, each site will obtain limited birth certificate data from all enrolled participants. These data on maternal and child demographics and pregnancy factors will be used in various analyses (both as adjustment factors and in some analyses as the primary risk factor of interest, e.g., preterm delivery). Sites also are asked to obtain to the extent possible, birth certificate data on all invited participants to enable comparison of responders and non-responders on key demographic and pregnancy factors. Some additional publicly available data might also be used in SEED analyses. For example, data collected on the Maternal and Child Residence History Form (**Attachment 6d**) will be geocoded

such that they can be linked to other data files such as various environmental monitoring databases such as those maintained by the EPA to study hazardous air pollutants as a potential risk factor for ASD.

Overall the various types of information collected in SEED collection are necessary to fulfill the study objectives. Collecting medical and health information will allow investigators to identify risk factors that might be associated with ASD. Without this information SEED will not be able to answer the scientific questions proposed as part of the purpose for this study and consequently, CDC will not be able to develop recommendations about primary prevention of ASDs or secondary prevention of associated sequelae.

A.3. Use of Improved Information Technology and Burden Reduction

NCBDDD will fund a Data Coordinating Center (DCC) and a Central Biosample Repository (Central Lab) for SEED. The DCC, at Michigan State University, has developed an electronic data collection system to centrally store (100%) of the data. The Central Biosample Repository is where all biosamples from the study will be shipped, processed, and stored. The DCC and Central Biosample Repository will work on an ongoing basis with the SEED investigators to implement the study.

SEED will apply information technology broadly to collect data efficiently, to assure both the quality of the collected data and the privacy and security of the collected data, and to minimize the burden to the study participants. The DCC will be responsible for the information technology aspects of the study. The DCC has previously created and hosted a custom web-based information system, called the CADDRE Information System (CIS), which was used for the first two phases of the study. The same system will be used in SEED 3. CIS was carefully designed to directly support all of SEED data collection workflows, data quality assurance processes, and provide secure database and Internet transaction services. Relevant services provided by CIS include:

- Generation of customized task lists specific to the role of each authenticated user.
- Role-based security that restricts user access privileges to the minimum required.
- Automated tracking of participant progress.
- Generation of bar code labels to identify all study documents and biologic samples.
- Computer-assisted-telephone-interviews (CATI) for MI and PRF instruments **(Attachments 4-5a&b)**.
- Double data entry for data collected on paper forms.
- Support for data entry and coding of copyrighted clinical assessment instruments by interfacing with other approved electronic systems.
- Ongoing data quality assurance checks
- Automated tracking and quality assurance reports.

- Comprehensive audit logging functions

- User support services.

While the CIS will be used by the study personnel only, the use of a facilitate computer-assisted-telephone-interviews (CATIs) for the Maternal Interview (MI) and Pregnancy Reference Form **(Attachments 4, 5a&b)** will reduced participant burden; the required logical branching is automatically provided by the CIS guidance to the interviewer during the interview. This implementation improves data quality and reduces errors to preclude the burden of follow-up calls to participants. During the MI and PRF interview and during all other calls with the participants, study staff also employ the CIS system to aid the interviewer in tracking completeness of response and thus reducing time for the participant.

Additionally, the CIS proactively tracks all aspects of participant's needs, requests, scheduled activities, and study protocol requirements. All contacts with each participant are tracked to ensure the efficient execution of the study. Staff are alerted automatically at login to all pending actions/tasks. Special care is given to preparation for the clinical visits. Automatic alerts for the clinical visit are provided to the staff about a participant's special needs, prior special requests, allergies, sibling child care, incentives, and any pending paper forms that still need to be completed.

Electronic data collection systems are used as much as possible in SEED both to minimize participant burden and reduce data entry errors. Nonetheless, for several data collection instruments participants are provided with paper forms and are given the option of completing the forms on their own or of having a staff member complete the forms for them over the phone. These forms include medical history checklist type forms **(Attachments 6a-c, h)** similar to those individuals are asked to complete at their doctor's offices, and standardized checklist child development forms that cannot be reformatted into a CATI because of copyright constraints (e.g., Social Responsiveness Scale, **Attachment 6f-g**). In SEED 1 and 2, we found that many participants preferred to complete these forms on their own rather than with a staff member. To ensure high data quality, SEED study staff members carefully review all forms upon receipt and follow-up with participants about inconsistent or ambiguous information; additionally, all paper forms are entered into CIS using a duplicate data entry function to minimize data entry errors.

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

No data collection activities currently supported by the Department of Health and Human Services (DHHS), other government institutions, or other private agencies, are comparable to the SEED proposed data collection. The Collaborative Programs of Excellence in Autism (CPEA) network – co-funded by the National Institute for Child Health and Development (NICHD), the National Institute of Deafness and Other Communication Disorders, and the National Center for Complementary and Alternative Medicine – were investigating the cause of autism at 25 sites in the United States, Canada, Great Britain, France and Germany. The National Institutes of Health (NIH) and the Interagency Autism Coordinating Committee (IACC), established the Studies to Advance Autism Research and Treatment (STAART) Network to conduct basic and clinical research in autism at eight centers in the United States. In 2007, NIH initiated the Autism Centers of Excellence (ACE) program to support studies covering a broad range of autism research areas, including early brain development and functioning, social interactions in infants, rare genetic variants and mutations, associations between autism-related genes and physical

traits, possible environmental risk factors and biomarkers, and a potential new medication treatment. Although some CPEA, STAART and ACE grantees have in the past, or are currently, investigating research domains similar to those in SEED, the CPEA, STAART or ACE sites do not all adhere to a common protocol. Use of a common protocol will allow SEED sites to pool data, resulting in a sample of >2,000 children in each of the three study groups – ASD, DD, and POP – at the end of SEED 3. Not only does the large sample size increase study power and statistical precision overall, but it also makes possible stratified analyses of phenotypic subtypes within the ASD case group as well as stratification on other factors across all subject groups.

Another recent autism epidemiologic project is the California Childhood Autism Risks from Genetics and the Environment (CHARGE) study. The CHARGE Study was funded by the National Institute of Environmental Health Sciences, the United States Environmental Protection Agency, and the University of California Davis Medical Investigation of Neurodevelopmental Disorders (MIND) Institute. CHARGE is investigating factors in the environment that are associated with autism in some children and families. Although the CHARGE study utilized data collection methods similar to SEED, there are multiple differences between CHARGE and SEED. The CHARGE sample is only 25% that of SEED. Moreover, CHARGE is collecting data only in the state of California, and therefore is less generalizable to a national population; CHARGE relied on a single source (Department of Developmental Services) for case ascertainment while SEED uses multiple source case ascertainment to achieve both a large and demographically diverse sample of study participants; CHARGE case and developmental delay comparison groups are more narrowly defined than those for SEED (thus the SEED ASD case group includes a more complete representation of children across the autism spectrum, which allows for more complete phenotype analyses and more accurate ASD subtypes for etiologic analyses); the research goals and corresponding data collection batteries differ somewhat between the SEED and CHARGE studies (while CHARGE collects more data on some environmental exposures than SEED, SEED collects much more detailed information on child health including gastrointestinal function and sleep features; and child behavioral phenotype).

A literature review conducted for SEED protocol development identified other case-control population-based studies on the pre- and perinatal etiological risk factors for autism; although none have utilized comparable data collection procedures.[10,11,12,13] For instance, previous investigations have used relatively small sample sizes, did not verify autism case status, and did not employ as detailed exposure data collection methods as SEED. This comprehensive literature review helped detect gaps in our current understanding of ASD, which, in turn, led to identification of high priority research domains.

The SEED sample size and unique data being collected in SEED allow for study of research domains not covered or not covered as fully by CHARGE or other autism studies. Additionally, while there is some overlap in data collection between SEED and other studies, this will permit replication of many smaller analyses published by other studies. In fact, many aspects of SEED data collection were explicitly set up to enable this kind of replication, which will allow for comparison of results. Autism is a complex neurological disorder that is difficult to diagnose because there is no clear biologic marker; early studies suggest a multi-factorial

---

[10] Burd L, et al. 1999. Prenatal and perinatal risk factors for autism. J Perinat Med;27(6):441-50. https://www.ncbi.nlm.nih.gov/pubmed/?term=Burd+1999+autism.

[11] Hultman C, et al. 2002 Perinatal risk factors for infantile autism. Epidemiology. 2002 Jul;13(4):417-23. https://www.ncbi.nlm.nih.gov/pubmed/12094096

[12] Juul Dam N, et al. 2001. Prenatal, perinatal, and neonatal factors in autism, pervasive developmental disorder-not otherwise specified, and the general population. Pediatrics;107(4):E63. https://www.ncbi.nlm.nih.gov/pubmed/11335784

[13] Glasson e, et al. Perinatal factors and the development of autism: a population study. Arch Gen Psychiatry. 2004 Jun;61(6):618-27. https://www.ncbi.nlm.nih.gov/pubmed/15184241.

etiology with the likelihood of both gene-gene and gene-environment interactions. Thus, duplication of findings in multiple independent studies is a key component of assessing the causality of identified risk factors.

Finally, the Associate Director for Science of the National Center on Birth Defects and Developmental Disabilities at CDC, Dr. Stuart Shapira, is a member of the Interagency Autism Coordinating Committee (IACC). The IACC was established in accordance with the Combating Autism Act of 2006 (**Attachment 1c**) and coordinates all efforts within DHHS concerning ASD. Through its inclusion of both Federal and public members, the IACC helps to ensure that a wide range of ideas and perspectives are represented and discussed in a public forum providing ongoing opportunities to identify overlap with other information collections.

A.5. Impact on Small Businesses or Other Small Entities

This data collection will not involve small businesses.

A.6. Consequences of Collecting the Information Less Frequently

The information collected from each SEED participant will only be collected once and has not been collected previously. The SEED case-control study is the first and largest multi-site, population based study on ASD planned and implemented to date, and the findings from this study will be essential to advancing the understanding of the causes of autism and ASDs. This data collection is in response to a mandate for research into the causes of ASD in the Children's Health Act of 2000; the Combatting Autism Act of 2006; and the Autism Collaboration, Accountability, Research, Education, and Support (CARES) Act of 2014 and has been directed by the IACC. The consequence of not collecting the information would be to severely limit information on autism causes from studies of US children. If these data were not collected researchers' ability to provide timely and important information related to the risk factors and causes of ASD and the characteristics and health of children with ASD would be greatly impacted. There are no legal obstacles to reduce the burden.

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

This request fully complies with the guidelines of 5 CFR 1320.5.

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

A 60-day notice was published in the *Federal Register* on May 24, 2019 Vol. 84, No. 101, pp 24144–24146 to make the public aware of this proposed information collection (**Attachment 9a**). No public comments were received.

We have consulted a number of persons outside CDC to ensure that this data collection is not duplicative and that the study design, data elements, and instruments are appropriate. The principal investigators (PIs) at each of the SEED 1 sites played an integral role in the design and the development of SEED. They conducted an extensive review of the literature, identified the research domains, selected the study design and data collection instruments, and developed the study protocol. The same sites were funded in SEED 2 and the PIs from these sites worked together throughout study implementation to discuss and resolve any issues related to implementing the study protocol, maintaining data security, and analyzing the data. Please see **Attachment 9b** for a list of SEED 2 PIs.

In December 2003, prior to submission to CDC IRB, the SEED group established a five person external peer review panel. This panel consisted of experts in clinical research, epidemiology, genetics, immunology, and advocacy, who were chosen on the basis of their expertise, balance, independence, and lack of conflicts of interest. Each of the panel members reviewed the SEED protocol and appendices with regard to several factors, including: the relevance of the proposed research domains and associated hypotheses, the effectiveness and feasibility of the scientific study plan, the appropriateness of the study design, study population, eligibility criteria, and case determination, adequacy of the sample size and study power, and, appropriateness of the data collection instruments and methods. The SEED protocol was revised based on the panel's feedback.

After protocol approval, CDC assembled periodic peer review panels to assess the portfolio of research conducted in the Developmental Disabilities Branch, including the SEED project. The most recent of these was in January 2015. Please see **Attachment 9c** for the list of the external review panel members. Although the purpose of that review was not specifically to suggest revisions to the SEED protocol (since the protocol necessarily needs to be consistent across phases to enable data harmonization), the CDC presented the panel members with a comprehensive overview of the SEED program and progress to date and asked panel members to consider SEED in the context of some specific issues related to SEED data collection, such as maximizing the use of the biospecimens already collected) and to consider any research gaps not covered by the current SEED protocol. Based on their review, CDC made the decision to continue collection of data with this current round (SEED 3) to ensure we have an appropriate sample size to answer the critical research questions that SEED was designed to address.

In addition to the more formal peer review panels that have been assembled, we have periodically sought consultation from various individuals at academic institutions with particular areas of expertise (such as sampling, statistical analyses, genetics, etc.) needed to address various challenges we faced throughout the course of SEED implementation.

Finally, through his role as a member of the IACC, Dr. Stuart Shapira, Associate Director for Science for the National Center on Birth Defects and Developmental Disabilities at CDC helps to ensure that a wide range of ideas and perspectives are represented and discussed in a public forum. Dr. Shapira provides updates on SEED progress at all annual IACC meetings.

A.9. Explanation of Any Payment or Gift to Respondents

All SEED families have young children and two-thirds of SEED families include children with autism or other developmental disabilities. These parents cope with challenges above and beyond what parents of typically developing children face. In general, tokens of appreciation have been found to be important for encouraging participation in federal research especially for more

reluctant responders,[14,15,16] including populations that SEED is specific designed to include (i.e., minorities and mothers with lower education, literacy or income).[17] Moreover, given that SEED includes an intensive data collection protocol with multiple data collection components occurring at varying points in time over a period of several months, it is important to include strategies to retain participants who are successfully enrolled. Other studies of mothers with young children have shown the importance of strategies to retain mothers of young children in studies given that participating in even simple activities that can be challenging in the face of financial constraints, childcare duties, and fatigue.[18,19] These factors may be particularly salient for families enrolled in SEED: these families all have at least 1 pre-school aged child; two-thirds have at least 1 pre-schooled aged child with a developmental disability; and many have multiple young children and multiple children with developmental disabilities. Thus, even for phone interviews and completion of self-administered forms, finding sufficient periods of uninterrupted time can be challenging for mothers. This problem is often exacerbated for many SEED families who are of lower SES – groups SEED was specifically designed to include.

In considering these tokens of appreciation, it is also very important to bear in mind that while they are linked to certain data collection components, those particular components do not account for the total time we ask a participant to spend on SEED study activities. For example, we have several follow-up calls with subjects to discuss study steps, schedule visits, etc. We send several packets of materials to participants that we ask them to review (for example a packet of materials to help them prepare their child for the lengthy in person visit that takes place at the end of the data collection protocol). We also offer participants help in completing self-administered forms via phone and offer to go over study materials we send them. This adds time to the protocol for some participants, but we find it is an important option that many families take. Given the diversity of our study population, we need to be mindful that some participants have low literacy level and thus need more support from SEED staff. These participants might also have the largest time constraints.

Thus, we propose providing tokens of appreciation to participants for completing each study step in SEED 3 to ensure a more representative study sample. These tokens of appreciation are lower than those provided to participants in SEED 1-2. Additionally, tokens of appreciation for the DD group has been further reduced commensurate with the reduction of the data collection protocol for this group.

Description and specific justification of tokens of appreciation in SEED 3 are as follows:

---

14 Berry SH, Pevar J, Zander-Cotugno M (2008). Use of Incentives in Surveys Supported by Federal Grants. Rand Corporation, March 2008: http://www.copafs.org/seminars/use_of_incentives_in_surveys.aspx.

15 Singer E, Ye C. (2013) The Use and Effects of Incentives in Surveys. Annals of the American Association of Political and Social Science, 645:112-141: http://journals.sagepub.com/doi/pdf/10.1177/0002716212458082.

16 Singer E, Kulka RA. (2002). Paying Respondents for Survey Participation. In *Studies of Welfare Populations: Data collection and research issues*.105-28. Washington DC: National Academy Press: https://aspe.hhs.gov/system/files/pdf/174381/04.pdf.

17 Kim SY, Tucker M, Danielson M, Johnson CH, Snesrud P, Shulman H. (2008). How can PRAMS survey response rates be improved among American Indian mothers? Data from 10 States. Matern Child Health J, 12(Supp 1):119-125: https://www.ncbi.nlm.nih.gov/pubmed/18350261.

18 Van Ryswyk EM, Middleton PF, Hague WM, Crowther CA (2015). Women's views on postpartum testing for type 2 diabetes after gestational diabetes: Six month follow-up to the DIAMIND randomized controlled trial. Prim Care Diabetes. 2015 Aug 27. pii: S1751-9918(15)00100-X. doi: 10.1016/j.pcd.2015.07.003: https://www.ncbi.nlm.nih.gov/pubmed/26320407.

19 Nicklas JM, Zera CA, Seely EB, et al. (2011). Identifying postpartum intervention approaches to prevent type 2 diabetes in women with a history of gestational diabetes. BMC Pregnancy and Childbirth 11:23: https://www.ncbi.nlm.nih.gov/pubmed/25837258.

1. Amount of $30-40 at multiple times during the study period for data collection via phone interview or completion of self-administered forms.

2. Amount of $75 POP group and $200 ASD group for in person assessment. This is a particularly intensive and burdensome visit in terms of both participant time and acceptability data collection components.

Time: Actual visit time is nearly 2 hours for the POP group and 5.5 hours for the ASD group, In addition, to visit time, these visits most often take place in a clinic which can be many miles from the participant's home. At some sites, some participants must travel an hour or more one way to reach the clinic location.

As discussed above, for most SEED families, time to participate in research studies can be very limited. Thus, incentives allow participants to purchase meals, childcare, or even take time off work without pay. The incentive for this data collection component must also cover travel expenses.

Acceptability of Data Collection Components: This visit includes simultaneous data collection activities for both mothers and their children. Mothers undergo intensive developmental interviews while children undergo a developmental assessment. While all SEED study staff are professionally trained with past experience conducting assessments with children, some children nonetheless have difficulty working with a stranger. Some mothers will bring along another trusted caregiver to help the child get comfortable with the visit. In other cases, the mother must help the child get started and thus start her own data collection components later, which adds to the total time burden.

During these visits, we also ask mothers to allow us to collect biosamples, including saliva, from themselves and their children. This can be a particularly difficult component for the child. Thus, while families can refuse any data collection component and we do not put undue pressure on them to participate in biosample collections, we do include this as a possible data collection component we ask them to consider.

Given the issues with both time burden and acceptability it is important to include tokens of appreciation to motivate families to complete the SEED study data collection protocol and to offset the time demands of the study.

The following are the specific structures according to data collection workflow for all enrolled participants.

Incentives for ASD workflow mother-child pairs (NOTE: only enrolled mother-child pairs receive incentives)

GROUP: ASD

	Data Collection Step*	Time to Complete	Amount
1	Invitation Packet/Response Card (**Attachment 10a,d,g**)	10 min	$30

2	Screening and Invitation Phone Call (Includes eligibility screen, description of the study, consent [**Attachment 11a**], and administration of the SCQ [**Attachment 3**])	30 min	
3.	Enrollment Packet (includes incentive for the eligibility screening **(Attachment 12a, c, d**)	20 min	
4.	Follow Up Phone Call (includes administration of Pregnancy Reference Form **(Attachment 5 a, b)** and discussion of maternal interview) **(Attachment 13)**	15 min	
5.	Maternal Interview Call (includes administration of interview [**Attachment 4]** and discussion of next steps in study)	60 min	$30
6.	Mailing of Self/Parent Administered Forms Packet (**Attachment 6a-e, 6f or 6g, 6h-i, 6k-l, and 6o-p**) includes forms to complete on parental and child health and child development, and materials about clinic visit preparation)	105 min	$40
7.	Follow Up Phone Call 2 (includes answering questions, help with self-administered forms as needed, and preparation for clinic visit) **(Attachment 14)**	20 min	
8.	Clinic/Home Visit (includes in depth developmental assessments, anthropometry, collection of biosamples, signing consents, and completing any remaining forms, including overall consent **(Attachments 7a-g, Attachments 8a-d, Attachment 15a)**	375 min	$200
	TOTAL	**10 hours, 35 minutes**	**$300**

Note: Steps 1 through 7 involve contact with and data collection from the biological mother. During the clinic/home visit, data are collected from both the mother and the index child.

Tokens of Appreciation for POP workflow mother-child pairs (NOTE: only received by enrolled mother-child pairs)

GROUP: POP			
	Data Collection Step*	**Time to Complete**	**Amount**
1	Invitation Packet/Response Card (**Attachment 10c,f,g**)	10 min	

2	Screening and Invitation Phone Call (Includes eligibility screen, description of the study, consent [**Attachment 11c**], and administration of the SCQ [**Attachment 3**])	30 min	
3.	Enrollment Packet (includes incentive for the eligibility screening [**Attachment 12, c, d**])	20 min	$30
4.	Follow Up Phone Call (includes administration of Pregnancy Reference Form (**Attachment 5a, b**) and discussion of maternal interview (**Attachment 13**)	15 min	
5.	Maternal Interview Call (includes administration of interview **[Attachment 4]** and discussion of next steps in study)	60 min	$30
6.	Mailing of Self/Parent Administered Forms Packet (**Attachment 6a-e, 6f or 6g, 6h-i, 6k, and 6n-p**) includes forms to complete on parental and child health and child development, and materials about clinic visit preparation)	105 min	$40
7.	Follow Up Phone Call 2 (includes answering questions, help with self-administered forms as needed, and preparation for clinic visit) (**Attachment 14**)	20 min	
8.	Clinic/Home Visit (includes developmental assessments, anthropometry, collection of biosamples, signing consents and completing any remaining forms) **(Attachments 7a-c, Attachments 8a-d, Attachments 15c)**	155 min	$75
	TOTAL	**6 hours, 55 minutes**	**$175**

Note: Steps 1 through 7 involve contact with and data collection from the biological mother. During the clinic/home visit, data are collected from both the mother and the index child.

GROUP: DD			
	Data Collection Step*	**Time to Complete**	**Amount**
1	Invitation Packet/Response Card (**Attachment 10b,e,g**)	10 min	

2	Screening and Invitation Phone Call (Includes eligibility screen, description of the study, consent [**Attachment 11b**], and administration of the SCQ [**Attachment 3**])	30 min	$30
3.	Enrollment Packet (includes incentive for the eligibility screening [**Attachment 12b-d**)	20 min	
4.	Follow Up Phone Call (includes administration of Pregnancy Reference Form [**Attachment 5a, b**] and discussion of maternal interview [**Attachment 14**])	15 min	
5.	Maternal Interview Call (includes administration of interview (**Attachment 4**) and discussion of next step in study)	60 min	$30
6.	Mailing of Self/Parent Administered Forms Packet (includes 4 forms to complete **[Attachments 6a-d, 6j, 6m, and 6o-p]**) and overall consent (**Attachment 15b**)	75 min	$40
	TOTAL	**2 hours, 30 minutes**	**$100**

Tokens of Appreciation for DD workflow mother-child pairs (NOTE: only received by enrolled mother-child pairs)

A.10. Protection of the Privacy and Confidentiality of Information Provided by Respondents

CDC's Information Systems Security Officer has determined that the Privacy Act does apply (**Attachment 16a&b**). The SORN is 09-20-0136, "Epidemiologic Studies and Surveillance of Disease Problems."

Data are collected and entered by each of the six study sites and are maintained by a Data Coordinating Center (DCC). The DCC has previously created and hosted a custom web-based information system, called the CADDRE Information System (CIS), which was used for the first two phases of the study. The same system will be used in SEED 3. Each site collects and maintains PII for recruitment of potential participants, identifying duplicate records, and for future contact for those who consent to future studies. A 12-digit identification number is used to encode the participant identity on data collection forms, specimens, and various other study materials. All transactions across the Internet of study data to/from the DCC occur over an encrypted connection, and then stored in the database in only an encrypted form (all 18 HIPAA 'Safe Harbor' identifiers as applicable). No documentation concerning the encryption/decryption secrets is saved on the database server.

The DCC also contracts with the Internet System for Assessing Autistic Children (ISAAC), a web-based application for administering and managing health research projects/studies for some of the data entry tools. ISAAC is a third-party internet portal scoring service for copyrighted assessment instruments. ISAAC access requires a username and password; all communication with the ISAAC servers uses 128-bit SSL encryption. Once SEED data are transferred from ISAAC to DCC, the data in ISAAC are expunged.

Access to CIS and ISAAC follows a least privilege model to prevent unnecessary viewing

of PII. The CADDRE System Security Plan (**Attachment 17**) describes the user privileges and who should have access to what PII maintained in the system. Secure logins are used to prevent unauthorized access from the application. CIS enforces a limited number of invalid access attempts by a user before lockout and utilizes FIPS-compliant encryption. The server room remains locked at all times through the use of RFID key cards and personal security passcodes assigned to individual authorized IT staff with proper security privileges. Physical measures, policies, and procedures are in place at each SEED site to protect information, buildings, and equipment from unauthorized intrusions, environmental hazards, and natural hazards.

When not in use by authorized project staff, all hard copy material and physical media containing confidential data will be stored in locked containers, locked file cabinets, or locked rooms. Access to locked storage areas will be limited to project staff who have completed confidentiality training and have been designated by the PI or Project Coordinator. This procedure will apply to all physical media containing confidential data, including data collections forms, printouts, diskettes, cds/dvds, flash drives, memory cards, laptop computers, and magnetic data tapes. Staff working with confidential materials during forms processing and data handling will have access only to the materials that they are currently processing. When confidential records are in use, they must be kept out of sight of persons not authorized to work with these records.

Except as needed for operational purposes, photocopies of confidential records are not to be made. If photocopies are necessary, care should be taken that all copies and originals are recovered from the copy machines and work areas. Whenever practicable, copy machines should be the type that do not retain hard drive copies of documents. In cases where copy machines retain electronic images, SEED Data or communications images should be deleted immediately. All confidential paper records will be destroyed as soon as operational requirements permit by burning or shredding the documents.

No data (or copies of data) are to be retained by a contractor after completion of the period of performance of the contract, or as specified in the contract for reasonable handling of data on back-up tapes/drives.

The DCC exports all of the data needed for analyses (i.e. nearly all data collected other than participant tracking data) into analytic data files. As part of this export process, the DCC de-identifies all date data using a date shifting algorithm and removes any open string text field data elements that are inherently identifiable or deemed very likely to be identifiable (such as name of child's hospital of delivery). The DCC uploads the analytic data files that have been nearly de-identified to a remote data access server (RDA) for access by the site investigators. The RDA meets CDC security requirements and is located in a HIPAA-compliant data-center with full redundant power and security measures. The DCC also delivers a complete set of the analytic data files to the CDC on an encrypted hard drive. The CDC stores these data in a secure location on SQL servers with limited access to the databases and identifiers are encrypted.

The only approved mechanism for accessing SEED analytic data sets is through the RDA maintained by the DCC. Scientists, colleagues, and collaborators who are given access to clinical, interview and biologic data from SEED through the RDA must sign a confidentiality and data use oath that describes how the data should be used and stored. All analyses are conducted on the RDA and aggregate results are downloaded but individual-level data are not. DCC actively monitors the RDA and ensures that data are not downloaded. The Principal Investigator of each SEED site has full and direct responsibility for tracking the use of SEED data at their site and assuring that each person who has access to the data has read and signed the

confidentiality and data use oath. Each site maintains files of the signed confidentiality and data use oaths. Signed statements will also be kept on file at the CDC.

For a number of reasons the analytic data files available on the RDA will not (and cannot) be considered completely de-identified:

- Some genetic data derived from analyses of biosamples will be compiled into analytic files and will be shared with SEED investigators through a second RDA server maintained by the DCC. Because genetic data are unique to individuals, they are not considered de-identified. Nonetheless, CDC wants to ensure that these important data are available to the SEED investigators to maximize their use in research analyses. The RDA approach offers a secure platform for these analyses. The second server will include all security provisions as the original RDA. Additionally, access to the genetic data files will be more restrictive. Only investigators with a clearly identified need to use the genetic data in their analyses will be provided access.
- The address data collected on the Maternal and Child Residence History form will not be included directly on the RDA. However, each site will be asked to allow DCC to derive various geocode variables from the data collected on this form and/or to themselves derive various geocode variables. If sites derive geocoded variables themselves, these data will be sent to DCC following an encryption protocol and the DCC will compile the information. The DCC will provide access to the geocode data on a designated restricted folder on one of the DCC RDA servers to individual SEED investigators who require access for data linkage purposes. The investigator will have to demonstrate a specific need for these sensitive data and that proposed analysis must be approved by the CDC and the CADDRE/SEED Data Sharing Committee (comprised of site principal investigators). DCC will provide such investigators time-limited access to the geocoded data such that they can link the data with other data files such as environmental monitoring data or Census data. Once the linkage occurs, the geocodes will be stripped from the linked data file. DCC will move this de-identified file to the main section of the RDA whereby all SEED investigators needing such data will have access and can combine these data with other SEED data.
- Beyond genetic and geocode data, the composite SEED data files will contain up to 20,000 variables per participant for participant families. Although the major issues related to identifiability will be addressed, with this volume of data it is not possible to ensure that combination of data about a research participant will not result indirectly, by reference to other information to identification of study participants.

Biologics samples will be stored in two ways, or destroyed at the end of the study, based on a choice by the study participants. The first way of storing the samples is to keep them linked to personal information (through a study ID). This will allow study investigators to contact participants again in the future. Participants who agree to have a sample stored with the study ID link intact are informed that they are only agreeing to potentially being contacted for future studies (which will require additional consent from participant). They will also be told they have the option to request this link be broken in the future, and are requested to do this by sending a written, signed letter to the study staff. Study participants will also have the option to store their samples without a link to personal identifiers. Under this approach the link between the participant's study ID and their biologic samples will be destroyed at the end of the study. Participants can also request to have their biologics samples destroyed at the end of the study.

Specific consent language has been included in the written consent forms to permit, if signed by the participant, the participant's genetic data to be placed in the National Database for Autism Research (NDAR) and the Database for Genotypes and Phenotypes (dbGaP) **(Attachment 15a&c)**. Both databases are run by the National Institutes of Health (NIH) that allows researchers studying autism to easily share and pool information with each other. To protect participant privacy, all identifying information such as name and address will be removed and replaced with an NDAR specific code number. The consent form language was developed in accordance with the evolving understanding of the identifiability of genetic data such that parents are fully informed of risks and benefits when providing consent for genetic data sharing.

Consent for the data collection will be obtained on four separate occasions for mother-child pairs included in the ASD and POP groups. First, during the initial telephone contact after the individual has received the invitation letter, verbal consent will be obtained to conduct an ASD screen (**Attachment 3**) and then to participate in the study. Second, verbal consent for the maternal interview (**Attachment 4**) will be obtained during the interview phone call. Third, enrollment packets will contain written informed consent forms (**Attachments 15a&c**) as well as a bill of rights for participants (**Attachment 12d**). Staff will obtain written, oral, or implied consent for the Medical History and Child Development Forms (**Attachments 6a-e, 6f or 6g, and 6h**), depending on whether the forms are completed face-to-face, over the telephone, or self-administered. Fourth, during the clinic visit, study staff will collect the informed consent for the overall study (**Attachment 15a&c**), including developmental assessments (**Attachments 7a-g**), and collection of anthropometric measurements and saliva swabs (**Attachments 8a-d**).

Given the streamlined protocol for mother-child pairs included in the DD workflow, consent for the data collection will be obtained on three separate occasions. First, during the initial telephone contact after the individual has received the invitation letter, verbal consent will be obtained to conduct an ASD screen (**Attachment 3**) and then to participate in the study. Second, verbal consent for the maternal interview (**Attachment 4**) will be obtained during the interview phone call. Third, staff will obtain written, oral, or implied consent (**Attachment 15b**) for the Medical History and Child Development Forms (**Attachments 6a-d**).

Due to the sensitive nature of certain data collection components, as in SEED 1 and SEED 2, a 301(d) Certificate of Confidentiality for protection of the individual participants at all six sites has been approved for SEED 3 (**Attachment 18a-f**).

A.11. Institutional Review Board Approval and Justification for Sensitive Questions

IRB approval was granted on 12/7/2015, and the most recent renewal extends through 12/7/2019. The current IRB approval letter is included as **Attachment 19**. The consent forms for the parents are included as **Attachment 15a-c**. Due to the age of the children involved in this study (2-5 years), parental consent alone is sufficient and the explicit assent of the child is not required. During the consent process, participants are fully informed about the potential uses of the information and the fact that their participation is completely voluntary. Participants are also assured that their decision about participating in the study will not affect their child's medical care. In addition, participants are given a chance to receive a semi-annual *Participant Newsletter* **(Attachment 20)** which keeps them informed about the study's progress and when the study results will be shared in general medical and public health journals.

SEED participants will be interviewed on multiple occasions and will be asked some questions that are potentially sensitive, including items on alcohol use during pregnancy, adverse pregnancy outcomes, family medical history, family income questions, and other lifestyle questions. The interviews have some risk of psychological discomfort, but women will be told at

the beginning of each interview that they may choose to skip any question at any time during the interview. In addition, we will accommodate the participants' wishes with regards to the timing of the interviews and will hire interviewers who are sensitive to the well-being of participants who are emotionally vulnerable. Extensive training will be held with all interviewers to address these issues.

Questions of particular sensitivity can be found in the following:

1. Maternal Interview **(Attachment 4)**

2. Self/Parent Administered Forms **(Attachment 6a-h)**

3. Child developmental evaluation **(Attachment 7a-g)**

We have included these items despite their potential sensitivity because research suggests that they are 1) potential risk factors for ASDs and the associations need further clarification; 2) important health conditions potentially related to ASDs in need of further study; 3) behaviors and symptoms important in further characterizing the spectrum of autism.

Specifically, some of these questions explore risk factors that may be:

- Direct hazards to the developing fetus (e.g., recreational drugs use during pregnancy, infectious diseases of the genitourinary system, medications taken during pregnancy)
- Pathways of exposure to potentially harmful agents to the developing fetus (e.g., infectious disease transmission associated with sexual intercourse)
- Related to poor reproductive outcomes (e.g., abnormal menstrual patterns or indicators of abnormal hormonal patterns such as menstrual history and fertility treatments).

Throughout the data collection process, subjects repeatedly are reminded that they may choose to skip any question that causes them undue discomfort and that their answers are not divulged to anyone outside the research group. Prior to beginning the *Maternal Interview* **(Attachment 4)** interviewers notify participants 'You may find some of the questions sensitive in nature but you can choose not to answer any question you wish' and, again that 'You may feel uncomfortable answering sensitive questions or discussing your pregnancies. Again, you can choose not to answer any question that makes you feel uncomfortable.' Participants in ASD and POP workflows sign a *written informed consent* **(Attachment 15a&c)** at the initial clinic visit. It informs participants that: 'You can refuse any task and still participate in the study.'

A.12. Estimates of Annualized Burden Hours and Costs

Although children are initially identified as potentially being eligible for a given group – ASD, DD, or POP – the final study group classification is determined from standardized research developmental assessments. Upon enrollment, all children are screened for possible autism characteristics through their mother's completion of the Social Communication Questionnaire (SCQ). Children with SCQ scores above a predetermined threshold (>11) are designated as potential ASD cases regardless of how they were initially identified. Additionally, all children who had a previous ASD diagnosis or autism special education classification are designated as potential ASD cases regardless of their SCQ scores. Potential ASD or DD or POP participants are designated into specific protocols, and these protocols are called workflows.

Children in the potential ASD workflow will have a more comprehensive assessment than children in the other groups. The clinic visit includes a developmental assessment, anthropometric measurements, and biologic sample collection. The child and mother parts of the clinic visit can happen simultaneously, and fathers are mailed the saliva kits. In addition to a general developmental assessment, children are administered the Autism Diagnostic Observation Schedule (ADOS) and their mothers are administered the Autism Diagnostic Interview revised (ADI-R). Final ASD case classification are based on the ADOS and ADI-R scores.

Those in the potential POP workflow will have a general developmental assessment resulting in a shorter clinic visit, fewer paper forms to fill out, and less burden than the potential ASD workflow group. Participants in the potential DD workflow will not have a clinic visit and will have even fewer paper forms to fill out than the ASD or POP workflow groups.

To estimate annualized burden hours, we have made the following assumptions based on data that have been collected already:

Mother-child pairs sampled from birth records for potential POP workflow group

- Of *potentially* eligible participants sent invitation mailings, study staff will have contact with 50%.
- Of those with contact, 24% will be ineligible.
- Of those with contact who are eligible, 60% will enroll.
- Of those eligible, consented, and enrolled, 90% of mothers will complete the first follow-up phone call and pregnancy reference form and the maternal interview.
- Of those eligible, consented, and enrolled, 80% of mothers will complete the self-administered forms and second follow-up phone call to review the forms.
- Of those eligible, consented, and enrolled, 70% will complete the clinic visit.
- Of those who complete the clinic visit, 50% of fathers will provide saliva specimens.

Mother-child pairs identified from health and school sources for potential ASD or DD workflow groups

- Of *potentially* eligible participants sent invitation mailings, study staff will have contact with 50%.
- Of those with contact, 22% will be ineligible.
- Of those with contact who are eligible, 70% will enroll.
- Of those who do not have a previous ASD diagnosis and are initially identified as potential DD workflow, 40% will screen positive on the ASD screen and will be evaluated in the ASD workflow instead of the DD workflow.
- Of those eligible, consented, and enrolled, 90% of mothers will complete the first follow-up phone call and pregnancy reference form and the maternal interview.
- Of those eligible, consented, and enrolled, 80% of mothers will complete the self-administered forms and second follow-up phone call to review the forms.
- Of those eligible, consented, and enrolled, 70% will complete the clinic visit (ASD workflow only)
- Of those who complete the clinic visit, 50% of fathers will provide saliva specimens (ASD workflow only).

Given the complexity of a study with participants in three different workflow groups, the table below presents the estimates by pre-enrollment, enrollment, and post-enrollment of each

workflow group (ASD, POP, and DD) to give the most accurate estimate of participant burden. The estimated annual burden is 7,118 hours.

A.12.A Estimated Annualized Burden Hours

Type of Respondents	Form Name	No. of Respondents	No. Responses per Respondent	Average Burden per Response (in hours)	Total Burden Hours
ASD Workflow					
PRE-ENROLLMENT					
Mother *All potential participants sent mailing*	**Invitation Packet/Response Card** (Attachments 10a, 10d, 10g)	1,718	1	10/60	286
ENROLLMENT					
Mother *Potentially eligible with contact by study staff*	**Invitation Call Script** (Attachment 11a) **and SCQ** (Attachment 3)	859	1	30/60	430
Mother *Eligible, consented, and enrolled; assigned to the ASD workflow based on enrollment intake*	**Enrollment Packet** (Attachments 12a,c,d) – Reference materials only; could be read on own or reviewed during subsequent calls.	469	1	20/60	156
POST-ENROLLMENT					
Mother *Completed this study step*	**Follow-up Phone Call Script and Checklist** (Attachment 13) **and Pregnancy Reference Form** (Attachments 5a and 5b)	422	1	15/60	106
Mother *Completed this study step*	**Maternal Interview Call** (Attachment 4)	422	1	60/60	422
Mother *Completed this study step*	**Self-Administered Forms** (Attachment 6a-e, 6f or 6g, 6h-i, 6k-l, and 6o-p)	375	1	105/60	656

Mother *Completed this study step*	**Follow-up Call 2,** answer questions, provide help with self-administered forms as needed, and prepare for clinic visit (Attachment 14)	375	1	20/60	125
Mother *Completed this study step*	**Clinic / Home Visit – Developmental Assessment, saliva collection and overall consent** (Attachments; 7b;7c; 7g; 8a-d, 15a)	328	1	225/60	1,230
Father *Completed this study step*	**Clinic / Home Visit – Saliva Collection (optional - on own)** (Attachments 8b-d)	164	1	15/60	41
Child *Completed this study step*	**Clinic / Home Visit – Developmental Assessment and saliva collection** (Attachments 7a; 7d or 7e or 7f; 8a-d)	328	1	135/60	738
ASD Workflow SUBTOTAL					**4,190**
POP Workflow					
PRE-ENROLLMENT					
Mother *All potential participants sent mailing*	**Invitation Packet/Response Card** (Attachments 10c, 10f, and 10g)	1,466	1	10/60	244
ENROLLMENT					
Mother *Potentially eligible with contact by study staff*	**Invitation Call Script** (Attachment 11c) **and SCQ** (Attachment 3)	733	1	30/60	367
Mother *Eligible, consented,*	**Enrollment Packet** (Attachments 12a,c,d) Reference materials only;	334	1	20/60	111

and enrolled; assigned to the POP workflow based on enrollment intake	could be read on own or reviewed during subsequent calls.				
POST-ENROLLMENT					
Mother *Completed this study step*	**Follow-up Phone Call Script and Checklist** (Attachment 13) **and Pregnancy Reference Form** (Attachments 5a and 5b)	301	1	15/60	75
Mother *Completed this study step*	**Maternal Interview Call** (Attachment 4)	301	1	1	301
Mother *Completed this study step*	**Self-Administered Forms** (Attachment 6a-e, 6f or 6g, 6h-i, 6k, 6n-p)	267	1	105/60	467
Mother *Completed this study step*	**Follow-up Call 2** answer questions, provide help with self-administered forms as needed, and prepare for clinic visit (Attachment 14)	267	1	20/60	89
Mother *Completed this study step*	**Clinic / Home Visit – Developmental Assessment, saliva collection and overall consent** (Attachments 8a-d,15c)	234	1	50/60	195
Father *Completed this study step*	**Clinic / Home Visit – Saliva Collection (optional - on own)** (Attachments 8b-d)	117	1	15/60	29
Child *Completed this study step*	**Clinic / Home Visit – Developmental Assessment, saliva collection and overall consent** (Attachments 7a; 8a-d)	234	1	90/60	351

POP Workflow SUBTOTAL					**2,229**
DD Workflow					
PRE-ENROLLMENT					
Mother *All potential participants sent mailing*	**Invitation Packet/Response Card** (Attachments 10b, 10e, and 10g)	641	1	10/60	107
ENROLLMENT					
Mother *Potentially eligible with contact by study staff*	**Invitation Call Script** (Attachment 11b) **and SCQ** (Attachment 3)	321	1	30/60	161
Mother *Eligible, consented, and enrolled; assigned to the DD workflow based on enrollment intake*	**Enrollment Packet** (Attachment 12b-d) Reference materials only; could be read on own or reviewed during subsequent calls.	175	1	20/60	58
POST-ENROLLMENT					
Mother *Completed this study step*	**Follow-up Phone Call Script and Checklist** (Attachment 13) **and Pregnancy Reference Form** (Attachments 5a and 5b)	158	1	15/60	40
Mother *Completed this study step*	**Maternal Interview Call** (Attachments 4)	158	1	1	158
Mother *Completed this study step*	**Self-Administered Forms** (Attachments 6a-d, 6j, 6m, and 6o-p)	140	1	55/60	128
Mother *Completed this study step*	**Follow-up Call 2** answer questions, provide help with self-administered forms as needed (Attachment 15b)	140	1	20/60	47
DD Workflow SUBTOTAL					**699**

GRAND TOTAL	**7,118**

A.12.B. Estimated Annualized Burden Costs

Annualized burden costs are summarized in the table below. The hourly wage estimates are based on the Bureau of Labor Statistics May 2018 National Occupational Employment and Wage Estimates (available at http://www.bls.gov/oes/current/oes_nat.htm). The mean hourly wage rate for all occupations ($24.98) was used.

Type of Respondents	Form Name	Total Burden Hours	Average Hourly Wage Rate ($)	Total Respondent Costs ($)
ASD Workflow				
Mother N=1,718	**Invitation Packet/Response Card** (Attachments 10a, 10d, 10g)	286	$24.98	$7,144.28
Mother N=859	**Invitation Call Script and SCQ** (Attachment 3)	430	$24.98	$10,741.40
Mother N=469	**Enrollment Packet,** (Attachments 12a,c,d)	156	$24.98	$3,896.88
Mother N=422	**Follow-up Phone Call Script and Checklist** (Attachment 13) **and Pregnancy Reference Form** (Attachments 5a and 5b)	106	$24.98	$3,996.80
Mother N=422	**Maternal Interview Call** (Attachment 4)	422	$24.98	$10,541.56
Mother N=375	**Self-Administered Forms** (Attachment 6a-e, 6f or 6g, 6h-i, 6k-l, and 6o-p)	656	$24.98	$16,386.88
Mother N=375	**Follow-up Call 2** (Attachment 14)	125	$24.98	$3,122.50
Mother N=328	**Clinic / Home Visit – Developmental**	1,230	$24.98	$30,725.40

	Assessment (Attachments; 7b;7c; 7g; 8a-d, 15a)			
Father N=164	**Clinic / Home Visit – Saliva Collection (optional - on own)** (Attachments 8b-d)	41	$24.98	$1,024.18
Child N=328	**Clinic / Home Visit – Developmental Assessment** (Attachments 7a; 7d or 7e or 7f; 8a-d)	738	$0	$0
ASD SUBTOTAL				**$87,579.88**
POP Workflow				
Mother N=1466	**Invitation Packet/Response Card** (Attachments 10c, 10f, and 10g)	244	$24.98	$6,095.12
Mother N=733	**Invitation Call Script and SCQ** (Attachment 3)	367	$24.98	$9,167.66
Mother N=334	**Enrollment Packet** (Attachments 12a,c,d)	111	$24.98	$2,772.78
Mother N=301	**Follow-up Phone Call Script and Checklist** (Attachment 13) **and Pregnancy Reference Form** (Attachments 5a and 5b)	75	$24.98	$1,873.50
Mother N=301	**Maternal Interview Call** (Attachment 4)	301	$24.98	$7,518.98
Mother N=267	**Self-Administered Forms** (Attachment 6a-e, 6f or 6g, 6h-i, 6k, 6n-p)	467	$24.98	$11,665.66
Mother N=267	**Follow-up Call 2** (Attachment 14)	89	$24.98	$2,223.22
Mother N=234	**Clinic / Home Visit – Developmental**	195	$24.98	$4,871.10

	Assessment (Attachments 8a-d,15c)			
Father N=117	**Clinic / Home Visit – Saliva Collection (optional - on own)** (Attachments 8b-d)	29	$24.98	$724.42
Child N=234	**Clinic / Home Visit – Developmental Assessment** (Attachments 7a-c; 8a-d)	351	$0	$0
POP SUBTOTAL				**$46,912.44**
DD Workflow				
Mother N=641	**Invitation Packet/Response Card** (Attachments 10b, 10e, and 10g)	107	$24.98	$2,672.86
Mother N=321	**Invitation Call Script** (Attachment 11b) **and SCQ** (Attachment 3)	161	$24.98	$4,021.78
Mother N=175	**Enrollment Packet (Attachment 12b-d)**	58	$24.98	$1,448.84
Mother N=158	**Follow-up Phone Call Script and Checklist** (Attachment 13) **and Pregnancy Reference Form** (Attachments 5a and 5b)	40	$24.98	$999.20
Mother N=158	**Maternal Interview Call** (Attachment 4)	158	$24.98	$3,946.84
Mother N=140	**Self-Administered Forms** (Attachments 6a-d, 6j, 6m, and 6o-p)	128	$24.98	$3,197.44

Mother N=140	**Follow-up Call 2** (Attachment 15b)	47	$24.98	$1,174.06
DD SUBTOTAL				**$17,461.02**
GRAND TOTAL				**$151,953.34**

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

There are no costs to respondents associated with either capital and startup efforts or operation and maintenance of services for this project.

A.14. Annualized Cost to the Government

The average annualized cost to the Government to collect this information is:

Federal Government Personnel costs	CDC Site Principal Investigator	$90,000
	CDC Project Officer	$60,000
	CDC Site Co-Principal Investigator	$80,000
	CDC Site Co-Principal Investigator	$80,000
	CDC Site Co-Principal Investigator	$80,000
	CDC Health Scientist	$75,000
	CDC Public Health Analyst	$65,000
	CDC Collaborator	$50,000
	CDC Collaborator	$50,000
	Subtotal for Government Personnel Costs	$630,000
Contractor and Grantee Costs	GA SEED (CDC)	$770,377
	Awardee #1	$770,377
	Awardee #2	$770,377
	Awardee #3	$770,377
	Awardee #4	$770,377
	Awardee #5	$770,377
	Awardee #6 Biorepository	$250,000
	Awardee #7 Data Coordinating Center (DCC)	1,200,000
	Subtotal for Contractor and Grantee Costs	**$6,072,262**
Total Government Costs		**$6,702,262**

A.15. Explanation for Program Changes or Adjustments

There are no changes to this data collection.

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

Data collection commenced 1 month after OMB approval is obtained and is expected to 4 years. Data cleaning and analytic preparation and QC of all SEED analytic data files (including harmonization of SEED 3 data with data from SEED 1 and 2) will take up to 1 year following data collection. Data analysis will begin as soon as the analytic files are finalized.

In most risk factor and child health analyses, children with a final classification of ASD will be compared to children in the two control groups (DD and POP). Odds ratios with 95% confidence intervals will be calculated for associations between ASD and various risk factors: overall; for etiologic ASD subgroups defined empirically based on analyses of the detailed behavioral and other phenotypic data; within strata defined by key demographic characteristics, such as child sex, race-ethnicity and maternal age; and after adjustment for demographic and perinatal characteristics. Interactions between genetic and non-genetic risk factors will be assessed through stratification and modelling. Child health characteristics will be similarly assessed through comparison of children with ASD to children in the DD and POP comparison groups. Separate in depth analyses of the ASD group will also be performed that examine the proportions of children with various behavioral and developmental symptoms and thus characterize ASD subgroups and the extent to which various ASD symptoms and other health characteristics cluster together.

Project Time Schedule	
Activity	**Time Schedule***
Letters of invitation sent to potential participants	Immediately after initial OMB approval
Data collection begins	1 month after initial OMB approval
Complete data collection	4 years after initial OMB approval
Finalize data cleaning and entry	4.5 years after initial OMB approval
Prepare analytic data files and harmonize SEED 1 and 2 data files	5 years after initial OMB approval
Begin to analyze data	5 years after initial OMB approval
Prepare first manuscript	5.5 years after initial OMB approval
Publication of first manuscript	6 years after initial OMB approval

A.17 Reasons Display of OMB Expiration Date Is Inappropriate .

The display of the OMB expiration date is appropriate, no exception is sought.

A.18. Exceptions to Certification for Paperwork Reduction Act Submissions

There are no exceptions to the certification.