The Study to Explore Early Development (SEED)

SEED Phase 3

OMB Control # 0920-1171

Expiration date: 3/31/2020

Supporting Statement: Part B

August 1, 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

Contents

List of Attachments

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

B.1 Respondent Universe and Sampling Methods

SEED is a case control study. Three groups of children and their mothers are enrolled. Case children are those with ASD as determined by the study protocol, which relies on developmental assessments administered in a standardized manner by study staff. Two control groups are additionally enrolled: children with other developmental disabilities and delays (DD group) and children from the general population (POP group).

It is expected that the six SEED 3 study sites will enroll approximately 3,000 children across the 3 study groups (ASD, DD, and POP) and approximately 2,100 children in all will complete the study protocol. Children must meet all of the SEED eligibility criteria (see Attachment 4.a) to be eligible to participate. Assuming a 70% protocol completion rate among enrolled participants, this requires enrollment of ~167 mother-child pairs in each study group (however, it is important to note that many children who will eventually be classified as an ASD case will initially be identified from data sources as children previously classified as DD because they don't yet have a ASD diagnosis).

Potential ASD and DD children are cohort children identified by the study as having a suspected ASD or other (non-ASD) developmental delay or disorder (please refer to Attachment 4.b for list of diagnoses). Potential ASD and DD children will be identified through sources serving or evaluating children with developmental problems; final ASD case status (i.e., confirmed cases) will be determined from a clinical evaluation using standardized developmental measures conducted as part of the study. Sources for potential ASD and DD children may include, special education programs, state early intervention programs, state autism registries, hospitals, and clinics. POP children will be identified from birth certificates on the basis of birth date range and maternal residence in the catchment area at the time of birth.

Response and Completion Rates

There are several challenges in accurately determining research response rates for SEED. Because SEED is designed as a case-control study to allow sufficient statistical power to examine risk factors for ASD, children are recruited for participation up to 5 years postpartum. Additionally, while a key strength of SEED over other epidemiologic ASD studies is the focus on enrolling children from diverse (often understudied) population subgroups, this complicates recruitment procedures as children must be identified from multiple health and education sources at each site. Given the fairly high residential mobility rates of the US population in general and women in the child-bearing age range in particular[1] and the increasing move to cell phones instead of landlines particularly for young adults[2] it is often challenging to locate individuals identified as potential study participants through tracing procedures.

The above issues render it very difficult to calculate accurate response rates for a study such as SEED where by design eligibility criteria are that participants need to be both born in defined study areas at each site and also continuing to reside in these areas at the time of the study. This latter

[1] Geronimus a, et al. 2014. Residential mobility across local areas in the United States and the geographic distribution of the healthy population. Demography 51(3):777-809. https://www.ncbi.nlm.nih.gov/pubmed/24781651

[2] Blumberg S, et al. 2013. Wireless substitution: state-level estimates from the National Health Interview Survey, 2012. Natl Health Stat Report. 2013 Dec 18;(70):1-16. https://www.ncbi.nlm.nih.gov/pubmed/24467831.

eligibility requirement is necessary so in person child developmental assessments can be conducted. As mentioned above, those in person assessments are necessary for case classification.

Recognizing that recruitment, and in particular getting participants to respond to the invitation mailings and/or follow-up phone calls made at some SEED sites was one of the greatest study challenges, we added several enhancements and supplements to the SEED 3 recruitment protocol. All sites actively increased outreach activities to make SEED more visible in their communities; some sites enhanced their tracing activities; and the SEED 3 protocol was streamlined and thus less burdensome for participants in SEED 3 compared to SEED 1 and 2 participants. Additionally, for SEED 3, all sites are asked to work more closely with their data sources to develop mechanisms to increase response to the initial invitation mailing, i.e. through active follow-up via telephone calls, email messages, or both.

Among participants in both SEED 1 and SEED 2, study completion rates among participants who enrolled was fairly high, with approximately 70% of enrolled mother-child pairs progressed through the study to the in person developmental assessment (which is the final data collection step) across all sites.

Based on data collected thus far in SEED 1, 2 and 3, we have made the following assumptions:

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/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).

SEED 3 Sample Size Justification

The final sample sizes in each of the study groups (ASD, DD and POP) for SEED 1 and the expected SEED 2 sample sizes for those same groups are presented in Table B.1. Altogether over 1,400 children with ASD are expected to be included in the combined SEED 1 and SEED 2 dataset. The numbers in the DD and POP comparison groups will be slightly higher.

Table B.1. Sample sizes from SEED 1 and expected from SEED 2

Final Classification	SEED 1	SEED 2	Target, SEED 3	Total
ASD Group	707	773	702	2,182
DD Group	995	810	702	2,507
POP Group	898	753	702	2,353

Note: All sample sizes pertain to the number of children who enrolled and completed the study protocol.

We expect this sample to allow us to address many important research questions within our primary and secondary research domains. However, the sample size and corresponding statistical power will not be adequate for many analyses of rare exposures, ASD subtypes, and genetic associations.

Table B.2 provides data on the range of prevalence estimates we expect to observe among the SEED POP group for the types of exposures and child health conditions that are assessed in SEED; these data are based on past population prevalence estimates. The examples highlighted in the Table illustrate that both rare and common maternal exposures and child health conditions will be examined in SEED.

Table B.2. Prevalence of select SEED candidate exposures and child health conditions and expected sample sizes.

POP exposure prevalence estimates	Expected number in SEED 1+ 2 + 3 ASD group (N=2,182) with exposure or health condition if NO association (null hypothesis is correct)	Example exposures and child health conditions included in SEED
1%	22	Assisted reproductive technology use for index pregnancy; Select maternal chronic conditions such as pre-existing diabetes; Select child conditions such as epilepsy
2%	44	Select maternal infections in pregnancy such as pyelonephritis or sexually transmitted diseases; Select maternal chronic conditions such as thyroid deficiency.
5%	109	Select past infertility conditions such as polycystic ovary syndrome (PCOS); Use of hormonal infertility treatments (all types combined) for index pregnancy; Select child health conditions such as asthma
10%	218	Any maternal autoimmune condition/; Select child conditions such as ADHD; GI dysfunction.
20% or more	436+	Maternal labor induction or stimulation with Pitocin; Cesarean delivery; Maternal fever in pregnancy; Any maternal infection in pregnancy (all types combined)

Table B.3 provides sample size estimates for various types of ASD subgroups of interest. These data highlight that even though we expect to achieve a sample size of >2,100 per study group by the end of SEED 3, our sample sizes for some important ASD subgroups might be as low as 300 or less.

Table B.3. Sample size estimates for example ASD case subgroups of etiologic interest.

Proportion of full ASD group	Expected number in SEED 1+ 2 + 3 ASD group (N=2,182) with exposure or health condition if NO association (null hypothesis is correct)	Subgroup
20%	436	Complex autism
30%	655	Nonverbal; with regression
40%	873	With intellectual disability
60%	1,309	No intellectual disability
70%	1,527	Verbal; without regression
80%	1,746	Essential autism

Table B.4 presents estimates of sample size needed for analyses based on various scenarios of exposure/health condition prevalence and the strength of the association between ASD (or ASD subtype) and exposure/health condition. All calculations assume 80% power, 5% alpha error, and a 1:1 ratio cases and controls. The odds ratio values used in these calculations – 1.50, 1.75, 2.50, and 3.00 – are based on typical findings from previous studies of various risk factor-ASD associations. The exposure prevalence values used in these calculations match those expected for many maternal exposures and child health conditions being assessed in SEED (see Table B.2). The findings indicate that the combined SEED 1+2 sample with >1,400 expected ASD cases would likely have been sufficient to examine associations between ASD and very prevalent exposures/health conditions (10% or higher), even when the magnitude of the association is fairly low (OR=1.5). Additionally, the SEED 1+2 sample would have been sufficient to examine associations between ASD and exposures/health conditions that are strongly associated with ASD (OR=2.5 or 3.0), even when the prevalence of the exposure/health condition were low (1%). However, the sample size from the first two SEED phases would not have been sufficient to examine all exposures with a prevalence of 1–5%, that are more modestly associated with ASD (OR<2.5). Moreover, analyses of associations with ASD subgroups, with much smaller sample sizes than the ASD group overall, would have been fairly limited for nearly all studies of SEED 1+2 data.

Table B.4. Sample size calculations for analyses under various assumptions of exposure prevalence and odds ratio.

Exposure Prevalence (%)	Odds Ratio	Sample Size Needed in ASD Group or ASD subgroup
1	1.50	7,964
2	1.50	4,041
5	1.50	1,690
10	1.50	912
1	1.75	3,900
2	1.75	1,983
5	1.75	834
10	1.75	454
1	2.50	1,247
2	2.50	637

5	2.50	272
10	2.50	151
1	3.00	804
2	3.00	413
5	3.00	177
10	3.00	100

All calculations assume 80% power, 5% alpha error, 1:1 ratio cases and controls.

Overall, the expectation is that SEED 3 will add 702 children to the ASD group and each of the two comparison groups. Thus, after SEED 3 we expect to have a sample of >2,100 or more children in the ASD group. This will expand our ability to analyze rare exposures and/or modest (yet scientifically important) associations. For example with the SEED 1+2+3 sample we expect to be better able to assess associations between ASD and maternal exposures/child health conditions in the following scenarios:

- Exposure/health condition prevalence 5% and expected odds ratio 1.50;
- Exposure/health condition prevalence 2% and expected odds ratio 1.75.

Also, we will substantially increase our ability to examine ASD subgroups in depth and to explore potential effect modifications between various ASD risk factors.

B.2. Procedures for the Collection of Information

A complete discussion of SEED 3 sample size estimates including justification for this sample based on statistical power analyses is presented in B.1.

Data Collection Procedures

Recruitment and enrollment contact with study participants (mothers of children identified as possible participants for the ASD, DD and POP groups) will consist of: 1) mailed invitation packet via United States Postal Service (USPS) and, 2) follow-up recruitment phone calls and/or emails, (Please see **Attachment 12c** for a full study flow chart).

The Invitation Packet **(Attachment 10)** sent via USPS will include:

- A letter introducing the study (**Attachment 10a-c** specific to study group [ASD, DD, or POP]):
- A study brochure (**Attachment 10d-f** specific to study group [ASD, DD, or POP]) which provides a brief overview of the purpose of SEED and specific information about the participant activities and incentives.
- A response card and pre-paid envelope that the invitee can return indicating interest or non-interest in future contact to learn more about the study **(Attachment 10g).**

To the extent possible, each site will cross check all lists of potential participants against each other in advance of the invitation mailing, such that individuals are not sent duplicate invitations.

If the respondent does not contact the study site within two to six weeks of the invitation mailing (depending on individual site protocols), sites (in concert with their agreements with their data sources) will follow-up with the potential participant. The exact protocol for non-response follow-up will vary by site. At some sites, follow-up phone calls and/or email contacts will be used. At other sites repeat mailings might be included as part of the follow-up protocol. Across sites, a maximum of 9 phone calls/emails will be made in an attempt to follow-up with the potential participant. Telephone calls will be attempted at various times during the day and different days of the week to maximize the potential of contacting participants.

If the potential participant indicates at any point that she is not interested in further contact, no further contact will be attempted. A negative response includes: returning the response card, with "No, I am not interested in learning more about SEED" checked; sending a text or email to the study site indicating that she does not want to be further contacted; calling the study site number and leaving a similar message; indicating to a staff member during a follow-up phone call that she does not wanted to be contacted again.

For those individuals who agree to partake in the enrollment process, the next step is the screening and invitation call. The call will include, verbal consent for a brief autism screen, the Social Communication Questionnaire (SCQ), and administration of the SCQ (**Attachment 3**).

The SCQ will be administered to mothers of ***all*** children regardless of how they were initially identified (i.e.as potential ASD, DD, or POP participants). Each child will be assigned to a data collection workflow (ASD, DD, or POP workflow) based on the results of the SCQ screener and whether or not they have a previous diagnosis of ASD. All children with an SCQ score above the threshold set a priori by the study will be considered a presumptive case (even if they do not have a previous diagnosis); additionally, all children with a previous ASD diagnosis at entry into the study will be presumptive cases (even if their SCQ score is below the study threshold).

Presumptive case children (and their mothers) will be assigned to the ASD workflow, which includes 2 in depth standardized autism assessments: (1) direct observation of the child using a standardized scored instrument and (2) an extensive standardized maternal interview which is also scored. The final case classification for these children will be based on the results from these standardized developmental assessments. Children identified as potential DD participants (that is, children who were identified from a clinical or school source has having a non-ASD developmental disability or delay) who screen negative on the SCQ and children identified as potential POP participants who screen negative on the SCQ will be placed in the DD and POP workflows, respectively.

The data collection protocol will vary according to workflow group. In brief, data collection for mother-child pairs in all 3 workflow groups will include, the Pregnancy Reference Form (PRF) questionnaire **(Attachment 5a,b)**, Maternal Interview (MI) **(Attachment 4)**, Maternal Medical History Form **(Attachment 6a)**, Paternal Medical and Occupational History Form **(Attachment 6b)**, Child Health History Form **(Attachment 6c)**, and the Maternal and Child Residential History Form **(Attachment 6d)**. The data collection protocol for the DD group will end upon completion of these instruments. The POP and ASD workflow data collection protocols will both additionally include the Child Behavior Checklist (CBCL) **(Attachment 6e)**, child Social Responsiveness Scale (SRS) **(Attachment 6f-g)**, an in person developmental assessment **(Attachment 7)**, collection of maternal, paternal, and child saliva specimens and anthropometry measurements **(Attachment 8)**. Although both the ASD and POP workflow protocols will include an in person assessment, it will be more extensive for the ASD than POP workflow. Children in both the POP and ASD workflows will be administered the Mullen Scales of Early Learning **(Attachment 7a)**. Children in the ASD workflows will be additionally be administered the Autism Diagnostic Observation Scales (ADOS) **(7d-f)** and their mothers will be administered the Autism Diagnostic Interview (ADI-r) **(Attachment 7g)**, the Vineland Adaptive Behavior Scales (VABS) **(Attachment 7b)** and the Services and Treatment Questionnaire **(Attachment 6h)**. The ordering, mode of administration, and other relevant field activities for each of these instruments is described below.

Upon completion of the Invitation and Screening call, the Enrollment Packet **(Attachment 12)** will be mailed to participants. This packet will include a cover letter **(Attachment 12a-b** specific to workflow [ASD, DD, or POP]), a participant-friendly study flow diagram **(Attachment 12c),** consent document to review (**Attachment 15a-c** specific to workflow [ASD, DD, or POP]), the study "Bill of Rights" **(Attachment 12d)**.

As part of the subject recruitment and retention process, study sites will maintain a record of written and telephone contacts with invitees and subjects. Once a family is enrolled, each site will track the family's completion of data collection steps and receipt of incentives using tracking software created by the Data Coordinating Center (DCC).

At least one week after the Enrollment Packet has been mailed, study staff will phone the participant to answer any initial questions about the study and the enrollment packet materials and to complete the Pregnancy Reference Form (PRF) questionnaire **(Attachment 5a,b)** (Note; some sites might opt to administer the PRF questionnaire at the end of the Invitation and Screening call, depending on participant availability). The PRF questionnaire is a very brief instrument that ascertains information on the index pregnancy timing that will be referenced during the Maternal Interview (MI); thus, the PRF questionnaire must be completed in advance of the MI.

Once the PRF questionnaire is completed, the participant is informed that the pregnancy calendar, the Maternal Interview Preparatory Guide **(Attachment 13)** and a $30 money order or cash card for completing the call will be mailed in advance of the interview.

The MI will be administered via telephone with a Computer-Assisted Telephone Interview (CATI). Interviews may also be conducted face-to-face for Spanish-only speaking participants. In these cases, a hard copy of the interview will be used initially, and information will be transferred to the electronic database. Once the MI is complete, the participant will be informed that a packet of checklist-type forms plus a $40 money order or cash card for completing the MI will be mailed to her next. For participants in the ASD and POP workflows, the in-person clinic visit might also be scheduled at the end of the MI call.

For all 3 workflow groups, the forms packet mailed to the mother will include:

- Maternal Medical History Form **(Attachment 6a)**.
- Paternal Medical and Occupational History Form **(Attachment 6b)**.
- Child Health History Form **(Attachment 6c)**.
- Maternal and Child Residential History form **(Attachment 6d)**.
-

For the for the ASD and POP workflows, the packet will also include 2 standardized developmental assessments:

- Child Behavior Checklist (CBCL) **(Attachment 6e)**.
- Social Responsiveness Scale (SRS) (**Attachment 6f-g** specific to age).
-

All of the above forms are designed to be self/parent administered. However, at all sites, staff will work with participants to complete the forms for them over the phone if preferred.

In addition to the foregoing forms, this packet will also include materials to assist the mothers and children in the ASD and POP workflows to prepare for the upcoming clinic visit (**Attachment 14**).

If the forms are not returned via mail and/or if the participant indicates she prefers to complete the forms with assistance, SEED study staff will telephone the participants. Additionally, even for forms that were completed and mailed to the SEED site, study staff will review any ambiguous or illegible responses during a follow-up phone call. For mothers in the ASD and POP workflows, the follow-up call will also include a review of the materials sent on how to prepare for the in person visit and a discussion of what to expect during that visit. Once the forms in the packet are completed, the data collection for DD workflow participants is complete. Mothers in this group will be mailed a thank you letter and $40 for the time spent completing the forms.

The last step of the data collection protocol for in the POP and ASD workflow groups is the in person visit. The visit will include four main components: answer questions and obtain written informed consent; complete any forms from the forms packet which have not yet been returned and sign the pre-

filled medical release forms (time varies and sites will strive to have nearly all of the forms completed in advance of the in-person visit – see above description of follow-up phone call); conduct developmental evaluation; anthropometry and collect saliva samples. The mother will also be provided with the final study incentives during this visit - $40 for completion of the forms packet and $75 or $200 for the visit depending on whether the child is in the POP or ASD workflow.

The visits will take place at predetermined locations at each of the study sites. At some study sites this may include participants' homes.

The developmental assessment will begin once the child appears comfortable with the diagnostician and staff. Parents will be encouraged initially to sit with the child during the assessment in order to make him/her more comfortable. Children will be able to take a break if needed. If for any reason the child cannot complete the assessment, the family will be encouraged to schedule another appointment to complete the evaluation.

The measures administered during the child developmental assessment are listed below. Copies of the score sheets for each of the measures can be found in **Attachment 7.**

- Mullen Scales of Early Learning (MSEL, POP and ASD workflows) (**Attachment 7a**): standardized assessment with the child that assesses cognition in five developmental domains,
- Vineland Adaptive Behavior Scales (VABS, ASD workflow) (**Attachment 7b-c**) Interview with mother that includes items on child's communication, daily living skills, socialization, and motor domains.
- Autism Diagnostic Observation Scale (ADOS, ASD workflow) (**Attachment 7d-f**): a standardized instrument in which the researcher observes the child and tries to illicit social interaction and communication through the use of structured play activities. The examiner implements the module that best corresponds to the child's expressive language level in order to prevent language aptitude from impeding accurate diagnosis. Generates data that are scored to determine final study group classification – ASD case or control.
- Autism Diagnostic Interview (ADI-R, ASD workflow) (**Attachment 7g**) a semi-structured interview with the mother that includes questions about child's current and past social and communicative behaviors.
- Services and Treatment Questionnaire (ASD workflow) (**Attachment 6h**) Captures information related to the child's current use of developmental services.

With appropriate consent, at least 10% of the overall sample of children who receive the full developmental evaluation will have the ADI-R and ADOS and/or Mullen and Vineland videotaped for quality control purposes. All data from the developmental assessment is recorded on hard-copy record forms that will subsequently be transferred into an electronic database.

After completion of the developmental evaluation, a qualified examiner will take height, weight and head circumference measurements of the child and height and head circumference measurement of the mother (both POP and ASD workflows). The examiner will also ask the mother if the child has ever been diagnosed or been examined for a major birth defect or genetic syndrome (**Attachment 8a**).

Saliva samples from all children in the ASD and POP workflows, their biological mothers and if available and willing their biological fathers will be collected. We anticipate compliance using this approach will be higher than blood sampling, because of the greater convenience and less discomfort to the participant.

If the biological father lives in the child's household and is not available for the clinic visit, but is willing to provide a saliva sample, then the father's saliva collection kit will be given to the biological mother at the clinic visit. Likewise, biological fathers not living with the mother and child may be mailed a saliva collection kit if they are willing to provide a specimen. Contents of the saliva collection kit are described in **8b-d**.

Training and Quality Control of Data Collection Activities

For each data collection instrument, staff training and quality control procedures have been developed. All staff members will receive a general overview training with a standardized training protocol and instruction manual. Each staff member will receive additional training with respect to the specific data collection components for which they will be responsible. Additionally, both initial and ongoing quality control procedures have been developed for each data collection activity. A summary of the quality control requirements for each study instrument is presented in Table B.6.

Table B.6. SEED quality control procedures summary

Study contacts and instruments	**Type of QC assessment(s) and requirements**	**Specific QC training requirements[1]**	**Ongoing QC requirements (frequency of QC)**
Invitation phone call, including eligibility screener and Social Communication Questionnaire	*Intra-site:* Semi-qualitative call rating form -- a priori criteria established for acceptable score.[2]	Acceptable scores on 3 role playing (mock) calls and first 2 "live" calls.	5% per interviewer
Follow-up call, including structured Pregnancy Reference Form interview	*Intra-site:* Semi-qualitative call rating form -- a priori criteria established for acceptable score.[2]	Acceptable scores on 3 role playing (mock) calls.	5% per interviewer
Maternal Interview	*Intra-site:* 1) Semi-qualitative call rating form -- a priori criteria established for acceptable score.[2] 2) Quantitative inter-rater reliability assessment of selected interview items. Acceptable score is ≥95% concordance.	Acceptable scores for both assessments on 3 role playing (mock) interviews and first 2 "live" calls.	5% per interviewer
Parent/self-administered forms[3] (forms sometimes administered by SEED study staff via phone or in person)	No specific systematic QC requirements, but continual supervisor oversight and all forms reviewed for missing or illegible data or contradictory entries. Participants are re-contacted as needed.	None. General training provided on forms and appropriate responses to participant queries.	NA
Autism Diagnostic Observation Schedule (ADOS)	*Inter-site:* supervising clinicians establish reliability by scoring the same ADOS exam videotapes. Acceptable score is ≥80% concordance on algorithm items. *Intra-site:* all clinicians establish reliability with supervising clinician. Acceptable score is ≥80% concordance on algorithm items.	Both inter-site and intra-site reliability established in advance of study start.	quarterly inter-site and intra-site reliability exercises
Autism Diagnostic Interview - revised	*Inter-site:* supervising clinicians establish reliability by scoring	Both inter-site and intra-site reliability	quarterly inter-site and

(ADI-r)	the same ADI-r interview videotapes. Acceptable score is ≥90% concordance on algorithm items. *Intra-site:* all clinicians establish reliability with supervising clinician. Acceptable score is ≥90% concordance on algorithm items.	established in advance of study start.	intra-site reliability exercises.
Mullen Scales of Early Learning	No specific systematic QC requirements, but continual supervisor oversight and all forms reviewed for missing or illegible data or contradictory entries.	None. Supervising site clinicians monitor initial assessments until competency determined.	NA
Vineland Adaptive Behavioral Scales	No specific systematic QC requirements, but continual supervisor oversight and all forms reviewed for missing or illegible data or contradictory entries.	None. Supervising site clinicians monitor initial assessments until competency determined.	NA
Anthropometrics: Height, weight, and head circumference	*Intra-site:* All examiners establish reliability with project coordinator or other qualified examiner on age appropriate individuals. Acceptable reliability defined as agreement in 95% of instances (agreement within 0.5 cm for height and head circumference and 0.1kg in weight). Equipment calibrated periodically as needed.	Acceptable measurements on 5 role play (mock) individuals and on first 2 participants.	5% per examiner
Biologic specimens: Saliva specimens (child, mother and father if available) (ASD and POP workflow only)	*All*: central lab staff processes specimens upon receipt and performs preliminary QC (gross visual inspection). s	None. Extensive staff training on study protocol for obtaining and processing biologic specimens	2% sample of duplicate specimens

[1] Training QC requirements include requirement for staff to pass formal reliability or other QC assessment on mock exercises in advance of "live" field work and initial QC requirement on first instruments/exams once in the field. For each instrument, if a study staff member does not meet criteria for acceptable score during ongoing QC, retraining and training QC requirements are instituted.

[2] Semi-qualitative call rating forms for invitation, follow-up, and maternal interview calls include items such as use of call script, coverage of essential points, ability to respond to participant questions, probing on unclear or neutral responses, professionalism, and delivery and response recording for applicable study instruments (Social Communications Questionnaire, Pregnancy Reference Form, or Maternal Interview). For each item, QC supervisor rates interviewer as “good”, “fair” or “poor”. Criteria for acceptable score include: no item rated as “poor” and 20% or less rated as “fair”; and mandatory ratings of “good” for select items (dependent on call type).

[3] Parent/self-administered forms include Maternal Medical History Form, Paternal Medical and Occupational History Form, Child Health History Form, Maternal and Child Residence History Form, Child Behavior Checklist (ASD and POP workflows only), Child Social Responsiveness Scale (ASD and POP workflows only), Child Services and Treatments Questionnaire (ASD workflow only).

B.3. Methods to Maximize Response Rates and Deal with No Response

In the preceding sections (B.1 and B.2) we described barriers to recruitment, response rates from SEED 1 and SEED 2, comparisons to response data from other ASD risk factor studies, and our procedures for following up with participants we fail to reach after the initial invitation mailing. Here we describe our analyses of non-response bias and revisions to the SEED 3 protocol to enhance response rates and better assess non-response.

Analyses of Non-Response in Previous Phase of SEED

We assessed potential non-response bias both theoretically and empirically. As in most research studies that require active consent, there is a possibility that the study participants will not be representative of the population. We reviewed key articles on non-response effects. In a case-control study, differential participation between case children and control children can potentially lead to biased results. However, as noted by Bartlett et al. (2015), "…logistic regression complete records analysis can provide asymptotically unbiased estimates of the association of an exposure of interest with an outcome, adjusted for a number of confounders, under a surprisingly wide range of missing-data assumptions…. Specifically, exposure odds ratios are estimated without bias (asymptotically) provided that missingness does not depend jointly on exposure and outcome, and even then, special cases exist where bias does not result."[3] Based on this well-grounded theoretical framework for nonresponse bias, we conducted empirical analyses to better understand which factors were and were not associated with non-response. Most SEED sites were not able to gather detailed information about non-responders. However, GA SEED has access to birth certificate data on all mother-child pairs sent an invitation packet whether or not they were contacted and whether or not they enrolled in SEED once contacted. These birth certificate data included several demographic factors and several perinatal factors. We conducted a series of analyses to better understand how non-response using the SEED data collection protocol might impact results from subsequent analyses of SEED data. Key findings from our assessment are summarized as follows:

- 3 demographic factors were associated with non-response in POP group: maternal race-ethnicity, maternal age, and maternal education.
- 2 of these factors - maternal age and education - were associated with non-response in the ASD group.
- None of the perinatal risk factors we examined – preterm delivery, very preterm delivery, Cesarean delivery, induction of labor – were associated nonresponse in either the POP or ASD groups.
- We estimated odds ratios for associations between ASD and each of the aforementioned perinatal risk factors in 2 samples: our final sample (i.e. complete case analysis) and the full GA SEED sample of invited participants. We found that both unadjusted and adjusted odds ratios were similar across the two samples. This indicates that analyses of these types of factors using our final study sample are not biased.
- To further explore factors not available on the birth certificate, we created sampling weights whereby we weighted the final sample to more closely match the initial invited sample in terms of maternal demographics. We then performed weighted and unweighted analyses for associations between ASD and several maternal factors ascertained from the maternal interview, such as maternal infertility. Our estimates from weighted analyses matched well with the estimates from the unweighted analyses. This again indicates that these types of risk factor analyses are not biased in SEED.

[3] Bartlett et al. 2015. Asymptotically unbiased estimation of exposure odds ratios in complete records logistic regression. Am J Epidemiol;182(8):730-6. https://www.ncbi.nlm.nih.gov/pubmed/26429998.

In sum, using GA SEED data we demonstrated empirically that while select demographic factors were directly associated with study response, other biologic factors were not. Moreover, we demonstrated that analyses of associations of such biologic factors – both perinatal factors on the birth certificate and preconception health history factors captured via maternal interview – were not impacted by the low SEED 1 response rates.

SEED 3 Enhancements

We have added several enhancements to the SEED 3 protocol to further address recruitment issues and our ability to assess response impacts. In SEED Phase 3 all sites are asked to work more closely with their data sources to develop mechanisms to increase response to the initial invitation mailing, i.e. through active follow-up via telephone calls, email messages, or both. Additionally, in SEED 3, sites are asked to develop a mechanism a priori to obtain necessary permissions to allow for linkage of birth certificate data on ***all*** participants (ASD, DD, and POP groups) invited including those who are never successfully contacted, those who refuse contact, and those contacted but found to be ineligible such that SEED investigators will be able to directly assess characteristics of respondents vs. non-respondents. Thus, SEED 3 data will be invaluable not only in increasing the SEED sample which will allow investigators to conduct detailed analyses of ASD risk factors and the health of children impacted by ASD, but also in providing an unprecedented opportunity to understand response impacts not only in SEED analyses but in many case control analyses of ASD, for which there is currently a clear dearth of data.

B.4. Tests of Procedures or Methods to be Undertaken

In SEED 1, several standardized developmental assessment instruments were used that were previously developed and assessed for reliability by other groups; we carefully reviewed the data from those past instrument assessments and chose instruments that were both shown to be valid and reliable and best fit our study objectives. We also developed several study instruments such as the maternal interview, and many of the parent/self-administered forms. In constructing these forms we used questions previously developed and tested in other studies when feasible.

All instruments developed as part of the SEED 1 protocol were also pilot tested on small numbers of mother-child pairs who met the eligibility criteria for participation in SEED (<9) to detect ambiguous questions. Moreover, during staff training and project planning at each site, multiple mock (role-playing) exercises were conducted to develop efficient processes for data collection and identify any gaps or problem areas. During these exercises, further question verbiage problems were sometimes identified (often we asked the individuals acting as mock participants in role playing exercises to be challenging by asking difficult questions and creating scenarios in which they had complex reproductive histories). Additionally, as previously mentioned, we established very comprehensive quality control standards for all data collection components of SEED 1 and upon initiation of field work we carefully monitored data collection practices and made some adjustments early in the study.

This careful planning was essential to ensuring we were collecting high quality data; nonetheless, by the end of SEED 1, we became aware of a few questions that didn't work well and a few instruments that needed to be revised. We also sought to streamline data collection in SEED 2; this involved dropping some instruments altogether and consolidating others. Before beginning field work for SEED 2 we again conducted a few very small pilot tests of select instruments and again had study staff go through extensive training and mock data collection exercises. We again made a few revisions based on issues identified before field work was initiated.

In SEED 3 we have eliminated two instruments altogether from the data collection protocol. All other instruments will be retained without revision in order to ensure efficient data harmonization.

During SEED 2 data collection, quality control exercises and ongoing general monitoring conducted at all sites indicated that all study instruments worked well.

B.5. Individuals Consulted on Statistical Aspects and Individuals Collecting and/or Analyzing Data

SEED is a collaborative effort between the CDC, NCBDDD and the extramural CADDRE Centers (Note: CADDRE is the consortium of Centers for Autism and Developmental Disabilities Research and Epidemiology, these are the sites funded to conduct SEED). Staff members in NCBDDD involved in SEED include the following:

- Karen Pazol is the Team Lead, Epidemiology Team, DD Branch and the Science Lead for the CADDRE Network. Dr. Pazol is responsible for scientific oversight of the CADDRE sites and SEED project activities overall, providing direction to the collaboration.
- Seema Gupta serves as Project Coordinator of CADDRE activities
- Daisy Christensen is the CDC Science Lead for the CADDRE-SEED Data Coordinating Center [located at Michigan State University and funded through a separate cooperative agreement] and GA SEED co-investigator
- Aimee Alexander is the CADDRE-SEED Data Coordinator and Biomarker Analysis and Laboratory Coordinator
- Lisa Wiggins is the CADDRE lead clinician and GA SEED co-PI and supervising site clinician
- Norbert Soke is the GA SEED PI
- Charmaine McKenzie is the GA SEED Project Coordinator
- Stuart Shapira is a GA SEED co-investigator
- Patricia Dietz is a GA SEED co-investigator

The PIs for the CADDRE network sites are:

John Constantino, M.D.
Washington University
School of Medicine at Washington University Medical Center
Phone: (314) 747-6772
Email: constantino@wustl.edu

Maureen Durkin, Ph.D., Dr.Ph.
University of Wisconsin-Madison
Waisman Center and School of Medicine and Public Health
Phone: (608) 263-7507
Email: maureen.durkin@wisc.edu

M. Daniele Fallin, Ph.D.
Johns Hopkins University, Bloomberg School of Public Health
Phone: (410) 955-3463
Email: dfallin@jhsph.edu

Julie Daniels, Ph.D.
University of North Carolina at Chapel Hill
School of Public Health
Phone: (919) 966-7096

Email: Julie_daniels@unc.edu

Cordelia Robinson (Corry), Ph.D., R.N.
JFK Partners/UCHSC
Phone: (303) 864-5261
Email: robinson.cordelia@tchden.org

For SEED 3, CDC funded five sites through open competition. Three of the sites (University of Colorado/Anshutz Medical Campus; Johns Hopkins University; University of North Carolina at Chapel Hill) had been included in SEED 1 and 2 and the PIs remained the same. Two new sites were funded (Washington University; University of Wisconsin, Board of Regents). The list of PIs for SEED 1 and SEED 2 is in **Attachment 9b**.

NCBDDD also funded a Data Coordinating Center (DCC) and a Central Laboratory and Biosample Repository (CLBR) for SEED. Michigan State University established and manages the Data Coordinating Center and Johns Hopkins University houses the CLBR for SEED 1 and SEED 2. Both a DCC and CLBR will again be funded for SEED 3 activities.

All SEED principal investigators worked collaboratively to develop the original SEED protocol. Multiple working groups were established to organize and develop this large and multi-faceted case-control study.

In addition to the working groups, some of the original SEED sites (California, Pennsylvania, and Maryland) had advisory boards, including parents of children with autism, to review the original study materials and the study design. The SEED sites also conducted focus groups during the original planning phase and planning for SEED 2 with parents of children with and without developmental disabilities. The purpose of the focus groups was to obtain additional feedback on the study design and feasibility of the study.

Additionally, CDC, along with the CADDRE partners, established a five person peer review panel during the original study design phase. This panel consisted of experts in clinical research, epidemiology, genetics, immunology, and advocacy (see **Attachment 9c**). Each of the panel members reviewed the SEED protocol and appendices and provided feedback to the CADDRE group. The CADDRE PIs identified changes that were required of the protocol based on the panel's feedback and these changes were incorporated into the protocol prior to submission to the IRB. Also, in preparation for SEED Phase 2, CDC invited a panel of outside experts in epidemiology and population-based research field methods to review the SEED Phase 1 recruitment and enrollment methods and provide recommendations for improvement. These recommendations were incorporated into the SEED Phase 2 methods.