

ABSTRACT

The National Children's Study Archive is an information, data, and sample repository for the National Children's Study, designed to make data and samples freely available for scientific research, with an approved research request. The study, which was active from 2009 to 2014, collected data and samples from over 12,000 mothers, fathers, and children across the US at 40 study locations. These participants are currently represented by nearly 14,000 variables, over 200,000 biological samples, and over 4,000 environmental samples in the Archive.

In order to facilitate meaningful evaluation and consumption of this wealth of data, the Archive has developed new exploratory tools. The Protocol Browser allows users to flow through the visit progression and visit instrumentation and identify available study datasets. The Participant Explorer allows users to investigate study participation by participant type (woman, child, father), demographics (e.g., education level, marital status), and data collection point. The Sample Explorer allows researchers to use demographics and study visit information to explore the available biological (blood, hair, nails, saliva, urine, vaginal swab, breast milk, cord blood, meconium, and placenta) and environmental (air, dust, water) primary and derivative samples that were collected from a subset of NCS families. Both the Participant and Sample Explorers provide researchers with the sample sizes available that fit the types of participants or samples of research of interest. The Variable Locator allows users to search the available NCS datasets for questions and variables of interest, returning specific data elements available. With these research tools, researchers targeting unique populations and topic areas can quickly establish if a particular topic area is represented in NCS data and samples. Once identified, researchers can use the Archive's proposal submission and review process to begin using NCS resources to pursue their scientific objectives.

BACKGROUND

The National Children's Study (NCS) Vanguard Study was a pilot for the planned Main Study cohort and was started in 2009. The NCS was conceived to be a nationally representative longitudinal cohort study of 100,000 children from before birth through age 21. The study aimed to evaluate environmental influences on child health and development. The Vanguard Study tested different recruitment strategies in 40



locations across 31 states throughout the U.S.

The study enrolled over 5,400 birth families and followed them through 2014. At protocol-specified study visits, questionnaires and interviews, neuro-psychosocial and cognitive assessments, and physical examination data were collected, along with nearly 19,000 biological and 4,000 environmental primary samples from which a sample repository of over 250,000 items was created.

PROBLEM

Faced with the problem of sharing complex data and study documentation, the NICHD identified the need to create a single data archive and sample repository with exploratory tools to facilitate ongoing scholarly endeavors around the NCS. From this charge, the NCS archive (www.ncsarchive.s-3.net) was developed to present the available study data and samples, provide supporting documentation, monitor and track proposed research efforts, and provide a secure environment where researchers, new or seasoned, could collaborate with support staff.

One of the challenges for the NCS is that the study was conducted with two main protocols—the Initial Vanguard Study (IVS) Protocol (2009–2010) and the Alternate Recruitment Substudy (ARS) Protocol (2011–2014). Each of these protocols included participant visits across multiple life stages (pre-conception, prenatal, perinatal, postnatal) each with different instrumentation. For a user to quickly assess how a participant progressed in the NCS, the Protocol Browser was developed. The Protocol Browser (Figure 2) shows points of participant contact within each protocol and displays those points of contact by life stage, participant visit, and instruments administered. Additionally, for each administered instrument a research dataset is identified if available. n addition to the flow through an individual protocol, the browser allows users to compare protocols (Figure 3) to understand how and if participants in each protocol experience the study differently.

Exploratory Tools for National Children's Study Data

Peter Gilbertson¹, Jodi Cisewski¹, Cora MacPherson¹, Christina Park², Anita Johnson¹, Jack Moye, Jr.³ ¹Social & Scientific Systems, Inc., Silver Spring, Maryland, USA; ²National Institutes of Health, Office of the Director; ³Eunice Kennedy Shriver National Institute of Child Health and Human Development

> Recognizing the core NCS Archive user will shift over time from NCS experienced to NCS inexperienced users, it was identified that systems or tools need to be developed to quickly convey study information. To facilitate the meaningful evaluation and consumption of this wealth of data, the Archive developed four new exploratory tools: (1) Protocol Browser, (2) Variable Locator, (3) Participant Explorer, and (4) Sample Explorer (Figure 1)

Figure 1. The Archive support tools



The Protocol Browser

Acconium Collection

Prenatal, L & D Medical Records Abstractio

Neonatal Medical Records Abstraction

igure 2. Protocol Browser protocol flow through dataset



IVS PRENATAL MED ABSTRT V2

IVS_NEONATAL_MED_ABSTRT_V2

The Variable Locator

The NCS Archive currently has information available from nearly 14,000 variables collected over the course of the NCS. Because many researchers tend to focus on one area of expertise, there needs to be a way to quickly evaluate what information was collected from NCS participants and how many data points, or responses, are available in that area of interest. To support this discovery the Variable Locator was created. The Variable Locator is a free text and variable name search for all data currently available in the NCS Archive. For example, a user can type in "sleep" and will be provided with a list of all variables and questions on sleeping, along with the number of participants who responded to the question and the number of records in the dataset (Figure 4). The Variable Locator includes the Study Phase, name of the corresponding dataset, variable name, variable label (question text), number of valid responses, number of legitimate skip responses, number of other responses, and total number of variable records. The locator allows users to use Boolean search terms AND, NOT, and OR to further filter their results. With this tool, a researcher can quickly identify the potential of NCS Archive data for their future research.

Figure 4. Variable Locator

question and the n	tor is a free text and variable name search for all da number of records in the dataset.	ata available in the NCS Archive. F	or example, a user can typ
The Variable Loca	tor includes the Study Phase, name of the correspo	onding dataset, variable name, vari	able label (question text),
Note: You can imp filter for results cor may also sort the c	prove your search results by using AND, NOT, and ntaining either the first term or the second. This is a columns by clicking on the table headers.	OR to further filter the results. Use also the default result if you enter m	AND between two terms to nultiple search terms witho
Note: Any change List in the show en	to the count of records to be displayed per screen tries box is applied immediately upon your selection	via the Dropdown on.	
Show 10 • ent	ries		
Study Phase	Dataset	Variable Name	Variable Label
ARS	SIX_MTH_CHILD_V1	SLEEP_PLACE_2_OTH	Description of what ch
ARS	SIX_MTH_CHILD_V1	SLEEP_PLACE_1	Does child usually slee
			What position does Ba
ARS	SIX_MTH_CHILD_V1	SLEEP_POSITION_NIGHT	what position does Pa
ARS ARS	SIX_MTH_CHILD_V1 SIX_MTH_CHILD_V1	SLEEP_POSITION_NIGHT	What position does Pa
ARS ARS ARS	SIX_MTH_CHILD_V1 SIX_MTH_CHILD_V1 SIX_MTH_CHILD_V1	SLEEP_POSITION_NIGHT SLEEP_POSITION_NAP SLEEP_ROUTINE	What position does Pa Does child have a regu
ARS ARS ARS ARS	SIX_MTH_CHILD_V1 SIX_MTH_CHILD_V1 SIX_MTH_CHILD_V1 SIX_MTH_CHILD_V1	SLEEP_POSITION_NIGHT SLEEP_POSITION_NAP SLEEP_ROUTINE SLEEP_HRS_NIGHT	What position does Pa What position does Pa Does child have a regu Approximately how ma
ARS ARS ARS ARS	SIX_MTH_CHILD_V1 SIX_MTH_CHILD_V1 SIX_MTH_CHILD_V1 SIX_MTH_CHILD_V1 SIX_MTH_CHILD_V1	SLEEP_POSITION_NIGHT SLEEP_POSITION_NAP SLEEP_ROUTINE SLEEP_HRS_NIGHT SLEEP_TIME_NIGHT	What position does Pa Does child have a reg Approximately how ma On a normal day, what
ARS ARS ARS ARS ARS	SIX_MTH_CHILD_V1 SIX_MTH_CHILD_V1 SIX_MTH_CHILD_V1 SIX_MTH_CHILD_V1 SIX_MTH_CHILD_V1 SIX_MTH_CHILD_V1	SLEEP_POSITION_NIGHT SLEEP_POSITION_NAP SLEEP_ROUTINE SLEEP_HRS_NIGHT SLEEP_TIME_NIGHT SLEEP_HRS_DAY	What position does Pa Does child have a regu Approximately how ma On a normal day, what Approximately how ma
ARS ARS ARS ARS ARS ARS	SIX_MTH_CHILD_V1 SIX_MTH_CHILD_V1 SIX_MTH_CHILD_V1 SIX_MTH_CHILD_V1 SIX_MTH_CHILD_V1 SIX_MTH_CHILD_V1 THIRTY_MTH_ADULT_HOUSEHOLD_V1	SLEEP_POSITION_NIGHT SLEEP_POSITION_NAP SLEEP_ROUTINE SLEEP_HRS_NIGHT SLEEP_TIME_NIGHT SLEEP_HRS_DAY PET_BEDDING	What position does Pa What position does Pa Does child have a regu Approximately how ma On a normal day, what Approximately how ma Do any of the pets slee



Figure 3. Protocol Browser protocol comparisons

SOLUTION: RESEARCH TOOLS



The Participant Explorer

The Participant Explorer is a tool designed to allows users to investigate NCS participants and their study participation on a metadata level. Users can see participant counts by participant type (woman, child, father), demographics (race, ethnicity, education level, marital status), and data collection point. Additionally, child participant can be categorized by gestational age, birth weight, and quarter of birth between 2009 and 2014. Users interested in family populations can review the demographic and activity characteristics for the 3000+ women, children, and fathers involved in the IVS phase of the NCS. Researchers focused on child health can review information on the 5500+ child in all NCS recruitment phases of the NCS (Figure 5).

Figure 5. Participant Explorer

Participant Explorer: Initial Vanguard Protocol

Participant Explorer tool allows users to view and manipulate the S, PBS) operational dataset. After selecting a dataset you can ex	ne data from the Initial Vanguard Protocol operational dataset and an All Study Childrer xplore the demographic and instrument information available for the selected participar
	618 participants
Sampled Participant	Child - Date of Birth
Initial Vanguard Protocol	2009 Q3
All Study Children (IVS, ARS, PBS)	2009 Q4
Participant	2010 Q1
	2010 Q2
Vvoman	2010 Q3
 Dad Child 	Child - Demographics
	 Male
	Female
	Multiples
	Child - Activity
	Child's Mother had a birth visit
	Physical Measurements (Birth)
	Physical Measurements (6 Month)
	Neonatal Exam (Birth)
	Infant Feed Form (1 Month)
	Infant Feed Form (6 Month)
	Child Interview (3 Month)
	Child Interview (6 Month)
	Child Interview (9 Month)



	5608 participants					
Sampled Participant	Child's Date of Birth					
Initial Vanguard Protocol	🔲 2009 Q3					
All Study Children (IVS, ARS, PBS)	2009 Q4					
Nother's Recruitment Protocol	2010 Q1					
	2010 Q2					
Initial Vanguard Study (IVS)	2010 Q3					
Alternate Recruitment Substudy (ARS)	2010 Q4					
Provider-Based Sampling Substudy (PBS)	2011 Q1					
Child's Gender	2011 Q2					
Male	🔲 2011 Q3					
E Female	🔲 2011 Q4					
when Ohild Democraphics	2012 Q1					
other office beinographics	2012 Q2					
Multiple	🔲 2012 Q3					
Siblings	 2012 Q4 2013 Q1 					
Deceased						
Withdrew	2013 Q2					
Child's Race	2013 Q3					
White	2013 Q4					
Black	2014 Q1					
AI/AN	Child - Interviews					
Asian	⊟ Birth Visit					
Other Race	3 Month Visit					
Multiple Race	6 Month Visit					
Child's Ethnicity	6 Month SAQ					
Hispanic	9 Month Visit					
Contational Area at Dist.	12 Month Visit					
Sestational Age at Birth	12 Month SAQ					
22 to 27 Weeks	18 Month Visit					



The Sample Explorer

There are currently 219,000 biological specimens and 4,600 environmental available samples connected with NCS participants. This large collection positions the Archive as a strong resource for future analytical studies. The Sample Explorer tool provides users with a method of exploring the available samples in the context of the participants that provided them.

The Sample Explorer allows researchers to use demographics, participant type, race, and ethnicity, and study visit information to explore the available biological (blood, hair, nails, saliva, urine, vaginal swab, breast milk, cord blood, meconium, and placenta) and environmental (air, dust, water) primary and derivative samples that were collected from a subset of NCS families. With the tool, researchers can identify sample sizes available that fit the types of participants or samples of research of interest. With sample number, sample demographic, and sample size information identified, a researcher can then put forth targeted research proposals requesting the samples of most interest to their research.

Figure 6. Sample Explorer

The Sample Explorer samples of potential u	allows you to search t se in your research, y	the inventory of biological rou can select the Create	and environmental sar Pending Request butto	nples that were collected in to start the sample req	from women, children, a uest process.	and fathers who partic	ipated in the NCS. Use thi	is tool to discover whi	ch samples are available for additional scientific analysis. If you locate one or more
69982 Samples								Sub Types	
69982 (1654) Blood	49740 (2185) Urine	3677 (939) Vaginal Swabs	796 (795) Hair	443 (443) Nails	6055 (1319) Saliva	3186 (424) Placenta	774 (373) Umbilical Cord	4661 (286) Breast Milk	 Blood + 1% ascorbic acid - 1982 Samples (1530 Participants) Blood + RPMI/DMSO - 3039 Samples (828 Participants) Buffy coat - 618 Samples (427 Participants) Buffy coat + RBC - 687 Samples (541 Participants) DNA, 200ng/ul - 4555 Samples (779 Participants) Plasma - 23259 Samples (1531 Participants) RBC - 292 Samples (257 Participants)
* (*) Cord Blood	* (*) Meconium	133 (131) Formula	229 (229) Dust - Plate	2216 (923) Dust - Vacuum	303 (289) Dust - Wipe	326 (322) Air	874 (244) Water - TWF	525 (228) Water - TWQ	
Participant 🚱	Participant 🕢 Race/Ethnicity						 RNA, 25ng/ul - 921 Samples (512 Participants) Serum - 27770 Samples (1642 Participants) Serum + MPA - 216 Samples (216 Participants) Whole blood - 2686 Samples (1356 Participants) Whole blood spot, dried - 3957 Samples (1537 Participants) 		
✓ Woman ✓ Dad	Woman Hispanic Dad White								
Child @				Bla	ck 😧				Related Publications Select a sample type to view related publications.
Time Point @				Autoria Autori	an 😧				Genetics NCS Research Day Abstracts
Pre-Pregna	incy 😧			NH	/OPI 🔞 Itiple Races 🚱				The National Children's Study. Abstracts of the National Children's Study research day 2011
 Early Pregr Late Pregration 	ancy 🚱			Oth	ner Races 😨				Short-Term Stability of Biospecimens
Pregnancy Genetic Consent							Planned variation in preanalytical conditions to evaluate biospecimen stability in the National Children's Study (NCS).		
 Birth Ø 3 Months Ø 	9			Ge	netic Consent				Clinical Chemistry and Laboratory Medicine
🗌 6 Months 🌘	9								Success of Consent for Collection of Various Sample Types Success rates for consent and collection of prenatal biological
12 Months	0								specimens in an epidemiologic survey of child health
36 Months	0								Birth Derects Research Part A. Chnical and Molecular Teratology

CONCLUSIONS

With these research tools, researchers targeting unique populations, focused topic areas, and specific biological or environmental samples can quickly establish whether the data or samples of interest are represented in NCS Archive. These research tools empower researchers to explore and refine study information in their own areas of study. Once identified, researchers can use the Archive's proposal submission and review process to begin using NCS resources to pursue their scientific objectives.

For more information, please contact:

NICHD website NCS Archive website **Archive Email**

http://ncsarchive c 2 not/ NCSArchiv



https://www.nichd.nih.gov/research/NCS/



Eunice Kennedy Shriver National Institute of Child Health and Human Development