EPA/600/R-03/073
August 2003

Report of the Workshop on Innovative Technologies for Remote Collection of Data for the National Children's Study

U.S. Environmental Protection Agency
Boston, MA
May 12–14, 2003

National Center for Environmental Assessment
Office of Research and Development
U.S. Environmental Protection Agency
Washington, DC 20460

NOTICE

This document has been reviewed in accordance with U.S. Environmental Protection Agency (EPA) policy and approved for publication. Mention of trade names or commercial products does not constitute endorsement or recommendation for use.

This report was prepared by Westat, Inc. under Purchase Order # 3W-0123-BBLX as a general record of discussions during the Workshop on Innovative Technologies for Remote Collection of Data for the National Children's Study. As requested by EPA, this report captures the main points and highlights of discussions held during this meeting. The report is not a complete record of all details discussed nor does it embellish, interpret, or enlarge upon matters that were incomplete or unclear. Statements represent the individual views of each workshop participant; none of the statements represent analyses by or positions of the EPA.

CONTENTS

EXECUTIVE SUMMARY vi

1. MEDICAL AND BIOLOGICAL MEASUREMENTS WORK GROUP SESSION SUMMARY 1
   1.1. GROUP ORGANIZATION 1
      1.1.1. Group Participants 1
      1.1.2. Group Discussion Format 1
   1.2. NCS THEMES AND OVER-ARCHING TECHNOLOGY AREAS 1
      1.2.1. Themes of the NCS 1
      1.2.2. Technologies Related to NCS Theme Areas 2
      1.2.3. Over-arching Medical Data Collection Processes 2
      1.2.4. Cross-Cutting Technology Issues 3
      1.2.5. Flexible Technology 3
   1.3. PHYSIOLOGICAL MEASUREMENTS 3
      1.3.1. Remote Physiological Measurements 3
      1.3.2. Remote Sensor Technology 4
      1.3.3. Embedded Sensors 4
      1.3.4. Wearable Sensors 4
      1.3.5. Types of Remote Physiological Measurements 5
      1.3.6. Types of Biological Specimens 5
      1.3.7. Technologies for Specimen Collection 5
   1.4. NEUROLOGICAL FUNCTIONING 6
   1.5. INJURY-RELATED MEASUREMENTS 6
   1.6. MEDICAL RECORDS COLLECTION 6
      1.6.1. Electronic Records Collection 7
      1.6.2. Settings for Data Collection 7
      1.6.3. Pharmacy Records 8
   1.7. COMPUTER-ASSISTED TELEPHONE INTERVIEWS 8
   1.8. PARTICIPANT DIARIES 8
      1.8.1. Types of Diary Information 9
      1.8.2. Electronic Participant Diaries 9
   1.9. MEDICAL IMAGES 9
      1.9.1. Types of Image Information 9
      1.9.2. Technology Issues Related to Medical Images 10
   1.10. GENETIC DATA 10
   1.11. DATA ISSUES 11
      1.11.1. Data Comparability 11
      1.11.2. Data Access 11
   1.12. TECHNOLOGY VALUE VS. EFFORT EVALUATION 11

CONTENTS (continued)

2. DATA COLLECTION WORKGROUP SESSION SUMMARY 15
2.1. GROUP ORGANIZATION 15
2.1.1. Group Participants 15
2.1.2. Group Discussion Format 15
2.2. DATA COLLECTION OBJECTIVES 16
2.3. STUDY DESIGN 17
2.3.1. Standard Operating Procedures (SOP) 17
2.3.2. Standards vs. Flexibility 18
2.3.3. Study Model 18
2.3.4. Human Subject Protection 19
2.3.5. Intervention and Study Bias 19
2.3.6. Burden on Participants and Staff 20
2.3.7. Training 20
2.3.8. Interpersonal Communications 21
2.3.9. Security Strategies 21
2.3.10. Audit Trails 22
2.3.11. Quality Assurance 22
2.3.12. Database Maintenance 23
2.3.13. System Support 23
2.4. SYSTEM EQUIPMENT 23
2.4.1. Evaluation Criteria for Technology Selection 24
2.4.2. Validation and Verification of Instrumentation 26
2.4.3. Incorporating New Technology 27
2.4.4. Multiple Technology Modes 28
2.4.5. Equipment Procurement 28
2.5. DATA CHARACTERISTICS 29
2.5.1. Data Format 29
2.5.2. Data Ownership 29
2.5.3. Data Sharing 29
2.5.4. Data Quality 30
2.5.5. Other Types of Data 31

3. ENVIRONMENTAL MEASUREMENTS WORKGROUP SESSION SUMMARY 32
3.1. GROUP ORGANIZATION 32
3.1.1. Group Participants 32
3.1.2. Group Discussion Format 32
3.2. ENVIRONMENTAL MEASUREMENTS TOPIC 34
3.2.1. Remote Monitors and Sensors 35
3.2.2. Air Monitoring 35
3.2.3. Miniaturized Sampling Devices 36

CONTENTS (continued)

3.2.4. Pesticide Monitoring 36
3.3. QUESTIONNAIRE AND INTERVIEW INFORMATION 37
3.3.1. Standardization and Quality Control 37
3.3.2. Household Activities and Dietary Consumption 38
3.4. LOCATION AND ACTIVITY PATTERNS 38
3.4.1. Time-Location and Physical Activities 39
3.4.2. Geographic Data and Relationships 40
3.5. TECHNOLOGY VALUE VS. EFFORT EVALUATION 40

EXECUTIVE SUMMARY

The Workshop on Innovative Technologies for Remote Collection of Data for the National Children's Study was held on May 12-14, 2003 at the Hyatt Harborside Hotel Boston, MA. The National Children's Study is a large longitudinal study of environmental influences on children's health and development. The study will examine about 100,000 children across the US and follow them during prenatal development, through birth, childhood, and into adulthood. The workshop was structured around three breakout groups: medical and biological measurements, data collection, and environmental and exposure-related measurements.

The Workshop was produced by the Environmental Protection Agency (EPA). The workshop sessions were structured around three breakout groups: data collection; environmental; and medical and biological measurements. Through a series of telephone conference calls, Westat and EPA discussed the general plan of the workshop and the mix of relevant experience that was necessary to balance each breakout session.

Expert participants were selected for the workshop based on their appropriate expertise in the three areas of data collection, medical and biological measures and environmental measurements. Experts were contacted to confirm their interest and availability to participate in the workshop. Their expertise was verified reviewing publications authored by the expert, their professional record, and their resumes. Three chairpersons were chosen to lead each of the three break-out groups.

The format in each of the workgroups varied somewhat. The *Environmental Measurements* and the *Data Collection* workgroups had workgroup members give presentations based on their expertise. The *Medical and Biological Measurements* workgroup had open discussion, guided by a moderator and a general outline of discussion topics. Both the *Environmental Measurements* and the *Medical and Biological Measurements* workgroups developed separate "Quadrant Charts" to characterize 'value' versus 'effort' for existing and future technologies.

THE MEDICAL AND BIOLOGICAL MEASUREMENTS GROUP

The Medical and Biological Measurements Group reviewed the five themes of the NCS, in order to focus the discussion of technology that could be applied to the study. The five themes identified were: asthma; obesity; neurological outcomes; injury; and pregnancy outcomes.

The Medical group determined that there were several data and specimen collection processes that would be over-arching among the themes of the NCS. The group labeled these

processes “big picture items” and determined that each of these items would need to be considered when developing the application of technology for data collection within each of the five theme areas. The “big picture” items identified by the group were: remote physiological measurements; specimen collection; health records collection (from physicians, pharmacists and image information); from other clinical sources; and genetic information.

The group evaluated the different technologies that could be developed to obtain medical and biological measurements during the NCS. The table below illustrates only those technologies that have high value and low burden. The other technologies, not shown here but summarized in a larger table in the Workshop Report, were thought to have either too high a value with high effort and higher burden, and therefore less desirable.

Key Technology Findings of the Medical Workgroup

Technology	Time to Develop	Value Quadrant	Burden to Patient	Burden to Provider
Physical Parameters				
Diaper sensors	Long	2	1	N/A
Clothing sensors	Short	2	1	N/A
Game boy (child focused)	Long	2	1	N/A
Home spirometry game package	Current	1	1 or 2	N/A
Accelerometer	Short	3	1	N/A
Accelerometer in helmet	Current	1	1	N/A
Accelerometer in shoe	Short	1	1	N/A
Accelerometer in temporary tattoo	Current	1	2 or 3	N/A
Diary	Current	1	2 or 3	N/A
Specimen Collection				
Filter paper blood sample	Current	1	2	N/A
Home chemistry kits e.g. cartridge analysis of urine	Current	1	1	N/A
RF ID tracking	Current	1	0	N/A
RF ID urine analysis	Short/Long	2	1	N/A
Blood draws by experienced	Current	2	2 or 3	N/A
Swab it and mail	Current	1	1	N/A
Hair - at home	Current	1	1	N/A
Semen	Current	2	2	N/A
Milk	Current	2	2	N/A

Technology	Time to Develop	Value Quadrant	Burden to Patient	Burden to Provider
Primary Care Physician Medical Records and Pharmacy Records				
Electronic medical records	Current/Long	2	0	2 or 3
Web	Current	2	0	2 or 3
PDAs	Current	2	0	2 or 3
Photocopy and fax/scan	Current	2	0	2 or 3
Abstraction by study personnel	Current	2	0	1
Interface to pharmacy database	Current	2	0	1 or 2
Scannable barcode on prescriptions	Short	2	1	1 or 2
Record to smart card for patient history record	Current/Long	2	2	3
Images				
Pregnancy ultrasounds (2-D)	Current	1	1	N/A
Pregnancy ultrasounds (3-D)	Current	2	1	N/A
Digitized spirometry signals	Current	2	1	N/A
Non-invasive images of arteries to identify blood plaques in youngsters	Current	3	1	N/A
Digital photos of neighborhoods	Current	1or 2	1	N/A
Images of birth defects	Current	1	1	N/A
Images for bone age	Current	1	3	N/A
DEXA fat/lean mass	Current	1	3	N/A
Bioimpedance assessment	Current	3	1	N/A
Genetic Information				
Buccal cells	Current	1	1	N/A
Blood	Current	1	2	1
Cord Blood	Current	1 or 2	0	2
Nails	Current	1	1	N/A
Mouthwash	Current	1	1	N/A
Tissues as available	Current	1	0	2
Computer-assisted telephone interviewing (CATI)	Current	1	1 or 2	N/A

THE DATA COLLECTION WORKGROUP

The goals of the Data Collection Workgroup were to: discuss and discover innovative approaches to the collection of data that can benefit the NCS; determine data collection architectural directions that will facilitate the adoption of new technology; establish criteria for evaluation of innovative data collection technologies in the NCS; define approaches to leverage technology to reduce respondent burden and improve retention of participants; and define parameters to improve and assure the quality of data.

The presentations given by the invited workgroup members covered the following topics:

- Criteria for evaluating innovative data collection technologies;
- Considerations in data architecture and data management;
- Lessons learned from multi-site studies;
- Hybrid communications architecture networks;
- Secure centralized portals;
- Quality assurance considerations;
- Automated remote data collection;
- Leveraging technology to improve participant retention.

Key Findings of the Data Collection Group

The key areas identified include: the data collection objectives; study design; standard operating procedure; standards versus flexibility; study model; human subject protection; intervention and study bias; burden on participants and staff; training; interpersonal communications; security strategies; audit trails; quality assurance; database maintenance; system support; system equipment; evaluation criteria for technology selection; validation and verification of instrumentation; incorporating new technology; multiple technology modes; equipment procurement; data characteristics; data format; data ownership; data sharing; data quality; and other types of data.

In addition to the key areas to be considered in the remote collection of data, outlined above, there are other considerations that were discussed: general technology; network technology design; and data storage. Data collection using sensors was another area discussed.

THE ENVIRONMENTAL MEASUREMENTS WORKGROUP

The format of the Environmental Measurements Workgroup was a series of prepared presentations, followed by open group discussion of each topic. The environmental measurements breakout group was structured around three general topic areas: environmental measurements and samples; questionnaires and interviews; and location and activity patterns.

A listing of all of the technologies identified by the group was compiled. This included: sensors for triggers, data capture, screening samples/field labs, PDA's sampling/questionnaires/bar codes, GPS/Accelerometers/GIS, RFID products, standards (open architecture), and portable instruments were ranked as being high in their potential value to the NCS and could be developed with relatively low effort. Portable instruments in multiple homes, bar code (burden/compliance), and PDA/Diary (compliance) were considered to have lower value and low effort. Pre-pregnancy enrollment, sensor data capture and developing new sensors (for PM, chemicals) are potentially high value, but with a high level of effort required. The group recognized the characterization of the technologies is based on many assumptions. The table below illustrates those technologies that yield the highest value with the least effort.

Key Technology Findings of the Environmental Workgroup

Technology	Value	Effort
Sensors for triggers, data capture	High	Low
Screening samples and field laboratories	High	Low
PDA for sampling information, questionnaires, and bar codes	High	Low
GPS, Accelerometers, and GIS for location and activity information	High	Low
RF ID for consumer product and medicine use	High	Low
Standards (open architecture)	High	Low
Portable Instruments	High	Low

1. MEDICAL AND BIOLOGICAL MEASUREMENTS WORKGROUP SESSION SUMMARY

1.1. GROUP ORGANIZATION

1.1.1. Group Participants

The group was chaired by Richard Shiffman, M.D., of Yale University. The participants in the Medical and Biological Measurements workgroup were:

Participant	Affiliation
Richard Shiffman, Chair	Yale University
Charlotte Andersen	Cincinnati Children's Center
Dan Ewert	North Dakota State University
Debbie Hillard	NHANES/Westat
Sarah Keim	NICHD/NIH
Alex Lu	University of Washington
Penny Manasco	First Genetic Trust
Marsha Marsh	EPA
Yechiam (Ami) Ostchega	CDC/NHANES
Jennifer Peck	Texas A&M
John P. Pestian	Cincinnati Children's Center
Kathy Schneider	Iowa Foundation Medical Care

1.1.2. Group Discussion Format

The format of the discussion of the Medical and Biological Measurements Workgroup during the workshop was an open discussion, guided by a moderator and a general outline of discussion topics.

1.2. NCS THEMES AND OVER-ARCHING TECHNOLOGY AREAS

1.2.1. Themes of the NCS

The Medical and Biological Measurements Group reviewed the five themes of the NCS, in order to focus the discussion of technology that could be applied to the study. The five themes identified were:

- Asthma;
- Obesity;
- Neurological outcomes;
- Injury;
- Pregnancy outcomes.

1.2.2. Technologies Related to NCS Theme Areas

The group identified types of remote technologies that could be utilized to collect data for the five theme areas of the NCS. The group discussed the use of online developmental screening to assess neurological development. Home pulmonary function testing was discussed as a means of assessing asthma in children. For measurements related to obesity, the use of non-invasive cardiac monitors (e.g., Holter monitors), the use of computer-assisted telephone interviewing (CATI) to obtain dietary and caloric intake data, and the use of monitors to assess the amount of TV watched by children were discussed. To collect data related to pregnancy outcomes, the use of remote ultrasound to date pregnancy, and remote glucose testing were discussed.

1.2.3. Over-arching Medical Data Collection Processes

The group determined that were several data and specimen collection processes that would be over-arching among the themes of the NCS. The group labeled these processes "big picture items" and determined that each of these items would need to be considered when developing the application of technology for data collection within each of the five theme areas. The "big picture" items identified by the group were:

- Remote physiological measurements;
- Specimen collection;
- Health records collection;
  - * From physicians;
  - * From pharmacists;
  - * Image information;
- From other clinical sources.
- Genetic information.

1.2.4. Cross-Cutting Technology Issues

The group discussed technology issues that apply to all areas of biological and medical measurements that would be undertaken by the NCS. The group felt that the NCS would need to assess whether technology should be chosen on the basis of its being available in the public domain or as a commercial product; whether the NCS would utilize technology specifically for the study or adopt commonly-used general products; and whether the technology would be acquired on an open-source basis or from a sole-source, stable vendor.

1.2.5. Flexible Technology

The group discussed the need to align the medical measurements acquired by the NCS with the study timeline, and determine which technologies might be available at future time points. The group emphasized that enough time would need to be built into the adoption of a technology to provide adequate pilot testing. The technology also would need to be standardized with any technologies that were used previously to collect similar data.

1.3. PHYSIOLOGICAL MEASUREMENTS

The group discussed the range of physiological measurements that might be utilized during the NCS to assess the health of children in the study. In addition to the basic physiological measurements, such as height, weight, blood pressure, and body mass index (BMI), the group identified other types of measurements that may be desired. With one of the themes being obesity, the group felt that it may be desirable to assess activity levels of children. For neurological outcomes, some type of assessment of central nervous system (CNS) functioning would be desirable. For pregnancy and developmental outcomes, fetal and uterine activity measurements may be desired, as well as Tanner staging of sexual maturity indicators.

1.3.1. Remote Physiological Measurements

The group moved on to discuss the types of technological processes that could be used remotely to provide data for physiological measurements. The use of low-powered lasers to measure height was proposed. To assess neurological functioning, a "developmental gameboy" concept was discussed. In this scenario, NCS participants would receive a pre-loaded gameboy device, with "games" specifically designed to assess CNS functioning and administer memory and attention tests. As the participant played the game, the device would record performance on various measures. The same concept could be applied to test hearing. To assess hearing,

participants could be given a headset with a pre-loaded game that would be designed to assess hearing and record results. The group noted that automated visual acuity tests are now commonly used to test adults and children in a clinical setting. A variation of this technology could be developed to test visual acuity in children remotely.

1.3.2. Remote Sensor Technology

The group discussed types of innovative technology that could be implemented remotely to collect various types of physiological measurements. The application of "RF ID" chips, or sensors, was presented. With this technology, an environment is "painted" with the radiofrequency (RF) spectrum, which collects and transmits the information that is specified for the particular application. An example of the application of this technology is to collect inventory information using barcodes that exist on each piece of inventory on a shelf. Instead of scanning each item individually, an RF ID chip would paint the shelf with energy, collect the barcode information, and transmit the barcode information back to the chip.

1.3.3. Embedded Sensors

The group discussed the application of the RF ID technology to various types of physiological measurements. The idea of embedding various types of clothing with these sensors was discussed, and several different applications to collect physiological measurements were proposed. For children, these sensors could be embedded in diapers and directed to collect basic physiological measurements (such as such as heart rate, respiration rate, and body temperature) as well as perform urinalysis and transmit the results. It was noted that these sensors were already being adapted for use in "Sensatex™" clothing by the military, with fiber-optic transmission mechanisms also embedded for transmission of data. Other applications provide readouts via a wristwatch, Personal Data Assistant (PDA), or voice. The physiological data collected could be wirelessly transmitted to a personal computer and ultimately, the Internet.

1.3.4. Wearable Sensors

The discussion of RF technology spurred a discussion of "wearable sensors," and their general application to data collection during the NCS. The idea of sensors that could be worn by children to collect various types of data, with minimal set-up or burden to the participants, was explored. Several applications were suggested. Sensors in shoes that could measure activity, weight, and body temperature were proposed. Several different types of sensors that could measure physical activity were proposed and discussed. The idea of a waterproof watch that

could measure activity was discussed. The use of pedometers, which have already been developed to measure several parameters of walking and running, was discussed. Also proposed was the use of sensors in bracelets or earrings to collect physiological measurements. Since accidents affect many children, the group was interested in the use of accelerometers in bike helmets that could assess parameters like speed, velocity, and direction of movement.

1.3.5. Types of Remote Physiological Measurements

A focused discussion on the types of physiological measurements that could be performed remotely was held by the group. The group listed types of physiological measurements that could easily be performed remotely, including metabolic activity, glucose levels, carbon-dioxide, and bilirubin. The group discussed existing remote medical measurement technology that could be applicable to the NCS, including home fetal monitors and bio-impedance scales. General issues that need to be considered with remote measurements were raised, including calibration of devices, adherence to protocols, and centralized versus remote data analysis.

1.3.6. Types of Biological Specimens

The group next listed the types of biological specimens that would probably be collected during the NCS. These included blood, urine, hair, nails, buccal cells, saliva, cord blood, placenta, milk, feces, meconium, amniotic fluid, semen and other tissues as available.

1.3.7. Technologies for Specimen Collection

The group then centered their discussion on the technologies that could be employed for collecting these biological specimens. For blood and urine, the group discussed the use of filter paper, laser-etched tubes, or cartridges for remote analysis, similar to home pregnancy tests. The group discussed using swabs for taking samples for culturing microbes, or for collecting DNA from cells. Collection mechanisms that involve swabs that perform chemical reactions that could provide data directly in the field were proposed, such as those used to test cholesterol levels or the presence of HIV.

To collect breast milk specimens, the idea of sensors in bras or nursing pads was raised. The use of specialized diapers to collect feces and meconium was discussed.

1.4. NEUROLOGICAL FUNCTIONING

The group moved on to discuss measurement and transmittal of data related to neurological functioning. The point was made that developmental outcomes are usually the indicators used to assess neurological functioning. The group discussed several ways that developmental outcomes could be recorded during the NCS. The idea of providing a "baby book" to participants to track developmental milestones was proposed, such as first word, first walked. The use of a "smart board" on which participants would record developmental events and that would provide electronic storage and transmittal of data about these events was proposed. The use of a web-based baby book, where participants could record developmental information, was discussed. The group discussed providing the respondents with pre-loaded PDAs that contain applications presenting checklists of developmental milestones. Respondents would record events on these checklists, and mail the media (e.g., smartcard or memory stick) to the data collection center; alternatively, data could be transmitted via phone line. To provide a positive incentive to participants, the group explored the idea of a "milestone exchange" in which user-friendly feedback about development is sent back to the participant that is tailored to the child's developmental event that was recorded. The use of video images as a means of recording developmental information was explored. It was suggested that the NCS could provide digital cameras to participants, with instructions about events or time points during which the child should be filmed by the parents.

1.5. INJURY-RELATED MEASUREMENTS

The group moved on to discuss the measurement and transmission of injury-related measurements. Video applications were discussed that could record teen-driving behavior and risk-taking behavior. Sensors in bike helmets, that could provide data about behavior, were discussed. Technology that could track eye movements, to explore behavior preceding an accident, was also discussed.

1.6. MEDICAL RECORDS COLLECTION

The group discussed various ways that medical records could be collected for use in the NCS. The group explored the idea of deputizing clinicians as part-time researchers during the NCS, to facilitate records collection. The idea of integrating the NCS with clinical practice was explored. The idea of providing incentives to physicians, such as furnishing them with systems or equipment that would streamline records collection in their practice in exchange for providing data to the NCS, was discussed.

1.6.1. Electronic Records Collection

The group discussed the current state and use of the technology to collect medical records electronically. The group agreed that electronic records could be a rich and easily-accessible pool of information for the NCS; however the use of electronic medical records is not yet widespread. The importance of standardization of medical records data collection to the NCS was emphasized. To facilitate standardization, the creation of electronic templates for records collection, such as web-based forms, laptop forms, or pre-loaded PDA applications, could be useful for the NCS. Technology to scan paper-based forms could be employed by the NCS to capture medical records data. The use of personal health records, in the form of a smart card that the respondent would present during any clinical visit, as a source of data for the NCS was proposed.

1.6.2. Settings for Data Collection

The group recognized that there were many types of settings that would be of interest to the NCS as data collection points, in addition to pediatrician's or family practice offices. The group listed these, which included:

- After-hours clinics;
- Hospitals;
- Emergency rooms;
- Specialists;
- Social workers;
- Day care providers.

In addition, the group listed the types of professionals that could provide records that would be useful to the NCS. These included professional found in school settings, such as teachers, nurses, psychologists, and counselors; as well as religious ministers, and police officers.

1.6.3. Pharmacy Records

The group thought that information related to prescription drugs would be useful data for the NCS. The group discussed various ways that this information could be collected. The use of a diary, in which parents record information about prescriptions, was discussed. The advantages included the capture of non-routine prescriptions such as physician samples and hospital prescriptions. The idea of collecting prescription information directly from the insurance companies who provide medical insurance for participants in the NCS was presented. The use of consolidated pharmacy records was explored. A pharmaceutical data research company, called Ingenix, was mentioned as an existing source of consolidated pharmaceutical records that could be utilized by the NCS. Several prescription-related incentives for participants in the NCS were proposed and discussed. The group discussed the issuance of a "NCS credit card" to participants that would pay for prescription purchases and track medications and amounts prescribed. The group also proposed implementing some form of special pricing for prescriptions for NCS participants, such as a discount or rebate that would be sent to the participant when the prescription was recorded as NCS data.

1.7. COMPUTER-ASSISTED TELEPHONE INTERVIEWS

The group discussed the applicability of computer-assisted telephone interviewing (CATI) as a means of data collection for the NCS. The point was made that CATI was already being used for diet recall and mental health assessment in several studies. The idea that new mothers might respond to CATI more readily than to methods that require more time and effort, such as web-based applications, was discussed. The implementation of CATI as an interactive data collection system for the NCS, providing feedback on events and symptoms, was discussed as a means of collecting data while providing an incentive to the respondent at the same time. Other advantages of CATI were mentioned by members of the group, such as the fact that the interview could be implemented in different languages, could be performed in a centralized location, is easily standardized, and could provide a pleasant and positive experience to the participants.

1.8. PARTICIPANT DIARIES

The group moved on to discuss the use of diaries in which the participants would record medical information.

1.8.1. Types of Diary Information

The types of information mentioned by the group that could be recorded in diaries included:

- Quality of life;
- Diet;
- Menstrual cycle tracking;
- Activities;
- Medications.

1.8.2. Electronic Participant Diaries

The group discussed the idea that an electronic implementation of a participant diary could be useful to the NCS. Several advantages of electronic diaries that would make them attractive for remote data collection were mentioned, including: the interactive nature of electronic diaries; branching checklists; the ability to attach a bar code wand for sample handling; the ability to interface to sensor and analysis devices; and the ability to synchronize information via the web, phone, or a cradle at a centralized location.

1.9. MEDICAL IMAGES

The desirability of collecting data in the form of medical images was discussed by the group.

1.9.1. Types of Image Information

The types of images that the group thought would be useful to the NCS included:

- Pregnancy ultrasounds (2 or 3-dimensional);
- Digitized spirometry measurements;
- Non-invasive images of arteries to identify blood placques;
- Digital photographs of neighborhoods;
- Images of birth defects;
- X-rays to assess bone age.

1.9.2. Technology Issues Related to Medical Images

The group agreed that there were several complex issues that would be encountered if the NCS utilized medical images as data for the NCS. The group discussed the process of data capture, and noted that the NCS would need to determine who is responsible for capturing the images, at what time should they be taken, how the images should be acquired, and in which setting the images would be captured. The storage of medical images also presented several issues. The NCS would need to determine the level of compression versus the acceptable level of degradation for images used in the NCS. The transfer process between clinical settings to the central image storage location would need to be determined. The NCS would need to determine whether or not it is necessary to retain original images, or retain images only in digital format, and whether or not back-up copies of imaged should be stored. The group proposed that the reading of images might be automated by technology during the NCS. It was noted that if the Internet had enough bandwidth to handle the transfer of images from approximately 40 sites during the NCS, the images could be stored in the field and interpreted at a centralized location. It was noted that the number of professionals needed to read and assess the images would need to be considered, since the NCS intends to collect images from 100,000 participants.

1.10. GENETIC DATA

Several issues were raised related to the collection of genetic information from NCS study participants were raised. The invasiveness of specimen collection procedures was explored, and it was noted that non-invasive procedures have been used to collect genetic specimens, such as cheek cell swabs, mouthwash, and cord blood collection. The issue of who might be sampled for DNA for the NCS was explored. It was proposed that the samples would need to be taken from the child participants, as well as mothers (perhaps during prenatal wellness visits) and from fathers. The group also explored the types of material that could provide genetic information. It was noted that depending on the type of information desired, the NCS might want specimens that provide data on DNA, RNA, or proteins; as well as data from specific types of tissue. The timing of sample collection was discussed. It was noted that samples should be collected early in childhood, for a baseline measure, and then over time, in order to correlate with environmental influences. As a practical matter, the relative ease of collecting of samples at baseline, versus collecting samples opportunistically, was discussed. Technology related to the long-term storage of genetic samples was discussed. It was noted that commercial efforts are underway to manufacture and store micro arrays of genetic information for future uses. The

group thought that while genome scans might be useful for the NCS, and commercial efforts may easily provide these 5-8 years in the future.

1.11. DATA ISSUES

1.11.1. Data Comparability

The objective of gathering comparable data from all participants in the NCS was discussed. The group noted that this could be achieved using validated methods, standards and protocols, and perhaps technology to check comparability of data. The group noted that the fact that the NCS is intended to take place over approximately 20 years presents data comparability issues. The group anticipated that data collection methods would change over the time that the study would be conducted. It was suggested that the study would need to store meta-data about how data were collected in order to provide comparable analytic results. The trade-off between compression, summarization, and abstraction of data against loss of data detail for future analysis or readings was discussed.

1.11.2. Data Access

The future use of data from the NCS was discussed by the group. The group noted that there would be privacy issues with providing public access to NCS data. Should data be released only to the research community, the NCS would need to define the individuals who are eligible to access the data. The group suggested that the data access system should be flexible enough to accommodate different levels of access to the NCS data.

1.12. TECHNOLOGY VALUE VS. EFFORT EVALUATION

The group evaluated the different technologies that could be developed to obtain medical and biological measurements during the NCS. The group categorized the technologies using the following framework:

- The availability of or time to develop the technology:
  - Currently available;
  - Short development time;
  - Long development time.
- The value quadrant of the technology to the NCS:
  - 1 = high value and low effort;
  - 2 = high value and high effort;

    * 3 = low value and low effort;
    * 4 = low value and high effort.
- The burden that using the technology would impose on the participant or the provider:
    * 0 = none;
    * 1 = very low burden;
    * 2 = low burden;
    * 3 = high burden;
    * 4 = very high burden;
    * N/A = not applicable.

The group's ranking for each technology appears on the following pages.

Technology	Time to Develop	Value Quadrant	Burden to Patient	Burden to Provider
Physical Parameters				
Diaper sensors	Long	2	1	N/A
Clothing sensors	Short	2	1	N/A
Game boy (child focused)	Long	2	1	N/A
Home spirometry game package	Current	1	1 or 2	N/A
Accelerometer	Short	3	1	N/A
Accelerometer in helmet	Current	1	1	N/A
Accelerometer in shoe	Short	1	1	N/A
Accelerometer in temporary tattoo	Current	1	2 or 3	N/A
Diary	Current	1	2 or 3	N/A
Specimen Collection				
Filter paper blood sample	Current	1	2	N/A
Home chemistry kits e.g. cartridge analysis of urine	Current	1	1	N/A
RF ID tracking	Current	1	0	N/A
RF ID urine analysis	Short/Long	2	1	N/A
Blood draws by experienced	Current	2	2 or 3	N/A
Swab it and mail	Current	1	1	N/A
Hair - at home	Current	1	1	N/A
Semen	Current	2	2	N/A
Milk	Current	2	2	N/A
Primary Care Physician Medical Records and Pharmacy Records				
Electronic medical records	Current/Long	2	0	2 or 3
Web	Current	2	0	2 or 3
PDAs	Current	2	0	2 or 3
Photocopy and fax/scan	Current	2	0	2 or 3
Abstraction by study personnel	Current	2	0	1
Interface to pharmacy database	Current	2	0	1 or 2
Scannable barcode on prescriptions	Short	2	1	1 or 2
Record to smart card for patient history record	Current/Long	2	2	3
Images				
Pregnancy ultrasounds (2-D)	Current	1	1	N/A
Pregnancy ultrasounds (3-D)	Current	2	1	N/A
Digitized spirometry signals	Current	2	1	N/A

Technology	Time to Develop	Value Quadrant	Burden to Patient	Burden to Provider
Non-invasive images of arteries to identify blood plaques in youngsters	Current	3	1	N/A
Digital photos of neighborhoods	Current	1or 2	1	N/A
Images of birth defects	Current	1	1	N/A
Images for bone age	Current	1	3	N/A
DEXA fat/lean mass	Current	1	3	N/A
Bioimpedance assessment	Current	3	1	N/A
Genetic Information				
Buccal cells	Current	1	1	N/A
Blood	Current	1	2	1
Cord Blood	Current	1 or 2	0	2
Nails	Current	1	1	N/A
Mouthwash	Current	1	1	N/A
Tissues as available	Current	1	0	2
Computer-assisted telephone interviewing (CATI)	Current	1	1 or 2	N/A

2. DATA COLLECTION WORKGROUP SESSION SUMMARY

2.1. GROUP ORGANIZATION

2.1.1. Group Participants

The group was chaired by David Songco, Chief Information Officer of the National Institute of Child Health and Human Development (NICHD.) The participants in the Data Collection workgroup were:

Participant	Affiliation
David Songco, Chair	NIH/NICHD
Arthur Bennett	NIH/NICHD
Lew Berman	CDC/NHANES
Greg Binzer	Westat
Margo Brinkley	RTI
Rick Chestek	Booz, Allen, Hamilton
Carry Croghan	EPA
John Menkedick	Battelle
Judie Mopsik	Abt
Michael Rozendaal	Iowa Foundation Medical Care
Brian Smith	Penn State
Paul Swidersky	Quality Associates
Jeffrey White	Miami Children's Hospital

2.1.2. Group Discussion Format

The format used by the Data Collection Workgroup during the EPA National Children's Study Workshop was to have short presentations on key topics as a lead in to more extended discussions. These presentations were given by the invited workgroup members and centered on their experiences with different aspects of data collection. The topics formally covered were:

- Criteria for evaluating innovative data collection technologies;
- Considerations in data architecture and data management;
- Lessons learned from multi-site studies;

- Hybrid communications architecture networks;
- Secure centralized portals;
- Quality assurance considerations;
- Automated remote data collection;
- Leveraging technology to improve participant retention.

These presentations enabled the group to identify issues that the National Children's Study should consider during its design of data collection systems and processes. The issues may be grouped into four areas: data collection objectives, system design, system technology, and data characteristics. Some examples of issues for NCS to consideration included:

- What does NCS need to collect and measure?
- What is the effective trade off between innovative technology and operational stability?
- How should NCS validate and verify collection instruments, technology and processes?
- How should NCS accommodate the need for the use of multiple modes of technology to meet differences in geographic conditions?
- How does NCS insure the use of standard operating procedures across sites?
- How should NCS be configured and operated to minimize the burden on participants and staff?
- How can NCS utilize the potential of communications techniques and innovative technology to help retain participants?

2.2. DATA COLLECTION OBJECTIVES

The workgroup felt limited in its abilities to suggest innovative technologies because the lack of clear objectives in terms of what the study needs to measure. The NCS needs to present a list of potential outcome and exposure data that they want to collect before a thoughtful discussion on the selection of equipment for collecting these data can be conducted. The group felt that the scope of the data involved is unclear.

For example, does all the medical information have to be collected through digital devices even if this technology has not been accepted into standard practices? Blood pressure is still measured with the hand pump and stethoscope in most hospitals and medical practices. Will the study enforce the use of digital blood pressure devices to the exclusion of older technology?

What is the state-of-the-art in digital blood pressure measurement? Is it sufficient for the NCS? What is the variability of measurements taken by individuals using current methods?

The group felt that clearer data collection objectives would facilitate the identification of the technology to be used in the study and the accuracy and precision required for each technology for each specific purpose. The group recognized that data collected may include images, physical specimens, information on forms of various kinds, and instrument-based digital data streams of different types. Does the study expect the final form of all information to be stored digitally, even if it may reduce the ability to make fine-grained distinctions? For example, the group discussed that the current technology for digital images of x-rays and scans may lack the amount of detail needed to make judgments, particularly for finer structures like the feet and hands.

2.3. STUDY DESIGN

Many issues related to system design and operational procedures were brought up by the workgroup. These issues included:

- Standard Operating Procedures;
- Standards versus Flexibility;
- System Architecture;
- Human Subject Protection and Privacy;
- Intervention and Study Bias;
- Burden on Participants and Staff;
- Training;
- Human Communications;
- Security strategy;
- Audit Trails;
- Quality Assurance;
- Database Maintenance;
- Study Support.

2.3.1. Standard Operating Procedures (SOP)

Standard Operating Procedures (SOPs) are typically written for routine, repetitive tasks and define how to perform specific activities or tasks; e.g., how to take a blood pressure

measurement. SOPs would insure that measurements are taken in the same manner across all NCS sites. However, with respect to the implementation of the NCS the issues related to SOPs would include:

- Transitioning from current SOPs that may already be in place at the NCS sites
- The difficulty in establishing, maintaining and enforcing standards across the sites to ensure data comparability
- Incorporating the required uniform and consistent security policies as part of SOPs despite current site practices

2.3.2. Standards vs. Flexibility

Given a study of this scope, the group felt that standardization will be essential. However, the issue of burden on participants and sites is one that made the group wonder how much flexibility (variation within standards) would be built into the system. Some felt that standards versus flexibility was not an issue. They pointed out that Standards don't necessarily reduce flexibility, since the adoption of new methods could be enhanced by standards.

Issues discussed under this topic included:

- Guidelines or standard methods for obtaining information;
- The need to establish standards at the beginning of design and development efforts;
- General versus local data coding standards;
- Standardization of support;
- The need for a standard for incorporating new equipment into the system (e.g. the NASA program requires any shuttle experiments to conform to NASA equipment-data standards);
- Platform standards and associated approved operating systems

2.3.3. Study Model

The group identified several data architectural models that would be possible. These included:

- Data driven;
- Operational procedure driven;
- Data ownership driven

These different orientations emphasized that the system may require a hybrid form of architecture that incorporates several of the "purer" forms. The architecture would evolve from consideration of the realities of the selected populations, their characteristics with respect to the use of specific technological approaches, and the infrastructure associated with the various geographic areas.

In addition, consideration should be given to peer-to-peer communications linkages as well as network centric approaches. This could improve communications for specific purposes as well as assist in overall reduction.

2.3.4. Human Subject Protection

The point was made that any study involving human participants will involve gaining the approval of OHRP and any governing IRB committees, as well as meeting the provisions of HIPAA. Assurances for all research studies involving human subjects must be obtained from OHRP prior to the use of Federal funds. Informed consent forms, protocols, etc., and any changes to these, have to be approved by an IRB before being submitted to OHRP.

Issues related to Human Subject Protection that were brought up by the group included:

- A study of the magnitude of NCS could have multiple IRB committees to work with and NCS would need to be aware of the number of committees and their effect on the study's design and schedule.
- The applicability of HIPAA requirements to this study needs to be considered. In addition, HIPAA may have some proposed changes that NCS may have to consider, such as the anticipated requirement for e-Signatures in 5 years.
- Informed consent will be obtained initially from the parents. However, at some time in the study this consent will have to be re-obtained from the child.

2.3.5. Intervention and Study Bias

The point was made that data collection would inherently involve significant ethical issues. In some cases of potential danger to the participant, an intervention may even be required. Examples that were mentioned by the group were lead hazards in the house, children who were not immunized, abuse or neglect of children, genetic abnormalities, and potential legal requirement (by State law) of disclosure of pregnancy teenagers.

The ability of NCS to remain strictly an observational study was questioned by group members. The impacts that NCS may have upon the households in the study was discussed in terms of the potential for the introduction of bias in the sample or in the outcome results. It is anticipated that just by being part of the study, participants may show more interest in their health and the potential impacts of environment. Also, if initial findings or reports are sent to the households or participants this may have an impact on their actions and subsequent outcomes.

2.3.6. Burden on Participants and Staff

The potential of innovative technology to reduce burden was discussed. Some suggestions to reduce burden by using technology included:

- Develop electronic reminder systems such as PDA timers and schedulers;
- Use technology to reduce the number of physical visits that may be required;
- Make all systems low visibility systems (in terms of intrusion in normal life activities);
- Minimize the security burden of signing onto a system;
- Use on-line chat rooms and instant messenger capabilities to let children or parents in NCS interact with their peer groups;
- Remember that readily available help and support encourages more acceptance of new technology.

2.3.7. Training

Training for NCS was identified as an important task that will be a continual process and not a one time event. It was felt that the use of innovative technologies could aid in this process.

The broad issues dealing with training identified by the group were:

- keeping the training consistent over time;
- providing training materials in multiple languages and literacy levels;
- being complaint with Section 508 requirements;
- providing just in time training, instead of training staff/participants weeks or months prior to implementing what they have learned;
- providing initial training face-to-face (perhaps regionally) and then provide refresher training via the web;
- dealing with turnover and training of new staff;

- anticipating training costs; sometimes training is the first thing cut;
- making auditing and testing an integral part of training.

2.3.8. Interpersonal Communications

The group felt that the success of NCS will greatly depend upon developing an effective system for interpersonal communications; one that allows people to interact with each other, send and receive instructions, transmit information of interest, remind participants of tasks to perform, etc. There are many types of technology that can be deployed to enhance communications across the NCS network. Data is not the only form of communications that needs to be carefully designed and encouraged, both for system effectiveness and for long term retention. It is important to establish clear lines of communications as well as relationships within the NCS organization and across sites.

2.3.9. Security Strategies

It was recommended that a standard security process should be established for NCS and that NCS would need to enforce an across the board security policy.

The group identified issues related to security, including:

- The need for a central system for secure file exchange, NCS web mailboxes, and workgroup discussions;
- Potential problems with wireless communications;
- The need to limit access to data based on a "need to know" basis (role-based access);
- The need to establish the role of local administrators;
- The challenge of establishing an "end to end" trust model;
- A system of registering users and assigning privileges;
- Limited availability of portals;
- Authentication of users;
- The desirability of an easy way for the participant to access the system;
- The adoption of a model for delegating security responsibilities.

2.3.10. Audit Trails

The importance of tracking changes made to the NCS data was discussed. Audit trails are essential for tracking changes to the original raw data. An audit system should record date and identity with the edits, so changes can be attributed. Also, it is standard practice to create a list of special codes to summarize the type of edits or errors discovered.

One of the participants described an experience using PDAs in a wireless peer to peer setting. He mentioned that audit trails were the biggest problem. It was difficult in this wireless setting to determine the original and final version of the document and who was responsible for edits to the document, once the record reached the central location. In addition, once the original record was transmitted to other medical professionals with PDAs, simultaneous editing of the original document would occur. That would cause problems in determining how to merge the edits into one document or knowing which version to accept as the final.

2.3.11. Quality Assurance

Quality assurance will be a major critical part of the NCS. It will be dependent on the collection processes established, uniformity of standard operating procedures, specific standards for electronic records, and incorporation of effective auditing mechanisms.

The FDA Regulation 20 CFR Part 11 and OMB data quality regulations present standards for automated data. NCS will need to determine the level of compliance it will need to meet in regards to these standards. It was mentioned that the 20 CFR Part 11 is a stringent standard usually associated with clinical trials, and that all aspects of this standard may not be applicable.

FDA Regulation 20 CFR Part 11 does provide definitions of electronic records and electronic signatures. Electronic signatures are used to identify an individual and require two distinct components (i.e., user ID and password).

The group identified several issues related to this topic:

- How can we build data quality considerations into the design process?
- How do we assure the validation and verification of processes, equipment and data collection methods are incorporated as ongoing activities?
- How can the NCS ensure e-signatures are used by genuine owners (both at clinical sites and by home users)?
- Will the e-signatures be considered legally binding signatures, and if so, how will this be certified?
- How will lost passwords be handled?

- How will the system be monitored for suspect transactions or unsuccessful attempts to logon?

2.3.12. Database Maintenance

Several questions related to management of the NCS database of field data were brought up by the group:

- Who will maintain the study database, both centrally and locally?
- How and when will the database and individual records be finalized (locked)?
- How will archiving of the database be handled? Where will it be stored? Who will be responsible for it?
- At what point in the study will archiving the database begin?

2.3.13. System Support

The workgroup felt that NCS needs to provide a user support system for all the participants and sites involved. The type of support offered may vary. Some support may be more formal, like a help desk with a single contact phone number, and some may be less formal, like user groups.

2.4. SYSTEM EQUIPMENT

The primary issue identified by the group with respect to innovative technology is the trade-off between new technology and operational stability. The group's opinion was that for a study of the scope of NCS, it would be extremely undesirable to use equipment that does not have an established record as to reliability or does not have the necessary accuracy or precision to satisfy NCS needs careful consideration must also be given to the field support that will be required for any technology employed. The group noted that many new technologies need to prove themselves before they can be considered reliable enough for a monumental study such as the NCS.

The workgroup discussed the following equipment issues:

- Evaluation criteria for selecting new technology and equipment;
- Validation and verification of instrumentation, both initially and on-going;
- Flexibility of incorporating new technology periodically over the life of the study;

- Multiple technology modes, from low-tech to high-tech, that may be needed to satisfy different population subgroups or operating environments/
- The possible uses of new technology:
  * To enhance recruitment, adherence, and retention;
  * To promote communications with participants or within the NCS organization.
- Equipment procurement:
  * Government funded equipment (GFE) or off the shelf (OTS);
  * Potential for Corporate and Foundation sponsorship;
  * Promoting the NCS "brand" using technology.

2.4.1. Evaluation Criteria for Technology Selection

A worksheet for evaluating technologies under consideration by NCS was presented to the group by one of the participants. A summary of the worksheet is illustrated below. Overall, the group thought it was a good starting point for summarizing various criteria to be considered in procuring equipment for NCS. It also provides a structured approach for comparing various technologies, which will be useful when selecting among alternative equipment for a common purpose.

The group suggested the addition of one new criterion to the worksheet titled "Support Structure", including both front end and back end. The support structure on front end may include work required to integrate new technology in existing clinics sites. Back end support may involve getting data from equipment installed at off-site locations.

Worksheet for Evaluating NCS Technologies

1. Technology Name
2. Description of use and role in NCS
    a. Does technology directly measure an exposure, outcome, or covariate?
        i. If so, over what period of time?
        ii. Applied to what percentage of participants?
    b. If the technology does not provide a direct measure (collect data), how does it impact NCS data?
        - Quality and Integrity
        - Security
        - Ease of Access and Use
        - Other
3. Criteria for Evaluation
    a. Performance Criteria (related to a "gold standard")
        i. Accuracy (Validation Studies)
        ii. Precision (Measurement Error)
        iii. Reliability (Frequency of missing data due to equipment failure)
        iv. Error rate (Frequency of wrong data due to mistakes/misuse)
        v. Other factors affecting performance (e.g. weather, locality, infrastructure, human factors)
    b. Cost (compared to traditional alternatives)
        i. Equipment
        ii. Training
        iii. Operations (Data transmission, processing, management, storage)
        iv. Data analysis
        v. Support
    c. Burden on Participants (compared to standard method)
    d. Indirect Effects
    - Cohort and study team morale
    - Cohort loyalty and retention
    - Reduction or increase in management costs

In addition to the evaluation points presented in the worksheet, the group also mentioned other considerations:

- Defining appropriate users for each form of technology;
- Understanding indirect effects on study team morale and commitment; cohort commitment and retention; effect on study management costs; important linkages (e.g. to other exposure databases);
- Interoperability;
- Infrastructure requirements to effectively utilize specific equipment;
- Proprietary technology;

- Regulatory or human subjects restrictions;
- Other performance criteria such as durability, scalability (e.g. regional area, number of users, capacity), longevity (e.g. data storage, change in precision of instruments, compatibility of data over the life time of the study);
- Data and retention;
- Quality and cost over time
- Life cycle costs
- Cost as a function of time, and consumer base. It may be possible to get custom devices.
- Standardization of support;
- Currently in place technology;
- Technology's role in other aspects of the study's success (in addition to data collection);
- Software requirements to tweak technology for special purposes of the NCS;
- Standards setting and adherence;
- Environmental consideration of recyclability.

The group felt that defining explicit data collection objectives will ultimately be necessary before any technology can be selected. The group agreed that NCS should not try to predict what the new technology should be, but should have a plan for upgrading technology. For example, it was suggested the NCS start off with established low technology solutions, which are cost effective, then move to new technology after several years. Many new technologies only have a lifespan around 14 months before the next model becomes available or the devices are no longer manufactured. The group noted that the manufacturer's life cycle might impact the study. However, if NCS was able to get a custom built device this could insure a standard device is available over a span of several years.

2.4.2. Validation and Verification of Instrumentation

The group agreed that any equipment used in NCS should undergo testing to verify that it is taking accurate/precise measurements and is a reliable instrument to use. Validation is the process of documenting evidence that the equipment and/or software does what it says it will do. Validation only occurs once, unless changes are made to the equipment or software that will require re-validation. Verification is a process done on a routine basis to demonstrate that the equipment and/or software are performing properly after validation has been completed.

Continual verification of the technology is required to insure collection of accurate measurements.

Three types of validation qualifications were identified by the group for NCS's consideration:

- Installation qualification – defines software;
- Operation qualification – installed software;
- Performance qualification – testing software and hardware;

Some issues mentioned under this topic were:

- Whether validation should occur when data are moved from component to component;
- The need for cross validation studies of equipment;
- For participant self-collected data, can NCS build in validation/verification into the process?
- Who will be responsible for the continual verification of equipment?
- Validation should be performed under protocol or SOP. Standards must be set for a successful validation.

2.4.3. Incorporating New Technology

The group noted that a longitudinal study of the NCS scope will have to anticipate the introduction and retirement of equipment throughout the data collection period. The selected technology may become obsolete or no longer manufactured. Many new technologies, like PDAs and Tablet PCs, are continually coming out with new and improved models every year to two years. If NCS utilizes off the shelf equipment, then the study will have to be prepared for handling different models of the same equipment. Maintaining a supply of one specific device may be a challenge unless NCS is able to convince a manufacturer to customize a device for them, which may be possible given the scope and magnitude of the study. It would be easier to control the phasing in of new equipment if customized devices are used.

2.4.4. Multiple Technology Modes

Given the national scope of NCS, the workgroup felt that participants will vary in the types of technology that is appropriate for them or that can be supported in their geographical locations. Therefore, NCS would have to support various modes of technology from low tech to high tech.

The workgroup felt that the technology with maximum coverage across the national sample would use telephone-based technology. Telephones are almost ubiquitous and are lower in cost compared to the alternative high-tech solutions. The existence of telephone technology also is essential to fax and low-speed (modem-based) Internet access, as well. The use of the web was discussed, but this would require all participants to have a computer, a local Internet provider, and some familiarity with new technology as well as access to an appropriate communications path.

In addition to multi-modes of technology, the group felt that NCS would need to incorporate think in terms of multi-channel approaches that would include telephone, fax, Internet, simple paper-based forms, and advanced techniques such as wireless or satellite communications. In addition, allowances would have to be made to effectively incorporate participants with different literacy levels, who speak various languages, or who may have disabilities or other limitations

2.4.5. Equipment Procurement

The issue of equipment procurement was brought up. New technology can have a high price tag which includes other costs in addition to equipment alone, such as software development costs and validation studies.

Some of the issues raised by the group related to this topic were:

- Who will be provided with equipment? Participants, sites, labs?
- What kind of standards will be set in terms of equipment?
- Would NCS provide equipment to sites and labs as a method of promoting participation and retention of their good services?
- Would the equipment be Government Funded Equipment (GFE) or off the shelf (OTS)?
- Could NCS get corporate and foundation sponsorship to pay for some of the equipment costs?

- Could the NCS be promoted using technology? By supporting certain products, can NCS get products free or at a reduced price? Would this be appropriate, despite the potential savings?

2.5. DATA CHARACTERISTICS

Part of the workgroup discussion focused on issues about the data itself. The following list summarizes the various data issues identified by the group:

- Data format;
- Data ownership;
- Data sharing;
- Data quality and quality assurance procedures;
- OMB data quality regulations;
- Other types of data needed for documentation and recordkeeping, such as, data from the validation process, technical information on the study instruments, copies of informed consent forms;
- Data harvesting from other sources.

2.5.1. Data Format

The primary issue regarding data is whether NCS will enforce data standards across all the multiple sites, and what data coding standards will be incorporated. To ensure consistent data across sites, a standardized data coding approach was recommended as well as standard forms for each type of data.

2.5.2. Data Ownership

The issue of who owns the data and where the data will be stored was discussed. Will it be center based or site based? Issues of confidentiality may play an important role in who manages the collected data. There is also a need to maintain accurate records of and protection for informed consent and other approval forms.

2.5.3. Data Sharing

Data will be collected because it is useful for the NCS research purposes. The data is not intended, however, to simply reside in a closed database. Analyses of the data will be done within the NCS Program centrally and at the participating sites. In addition, NCS data is

intended as a major resource for researchers outside the immediate program. Mechanisms must be established to provide access to NCS data under strict controls and only after the data has been properly scrubbed and validated. The utility of NCS data may be enhanced over time as it is used by external researchers and combined/contrasted with additional studies that may be performed outside the NCS.

In addition, data related to specific participants may be provided to them in summary fashion as a retention incentive or as an ethical necessity. Public summaries of progress and findings will also be published at regular intervals. The NCS data system must provide for all these uses.

2.5.4. Data Quality

This was one of the major areas of discussion for the workgroup with a presentation dedicated to the topic. The practice of quality assurance and control is a very mature process that in clinical research, and has been documented extensively. The group felt that NCS would need to establish a quality assurance program from the beginning, and determine at what points in the study QA checks would need to be done, and how they would be performed.

It was suggested that QA checks could be performed either at a central office or the individual sites. If the sites are to conduct the QA reviews, different QA standards may exist across the sites and perhaps across the various laboratories that may be involved. It was recommended that NCS evaluate practices currently in place as part of review process for selecting sites, and incorporate the results of this evaluation as part of qualification requirements to be included in the study.

With respect to quality review of the collected data, the group identified different challenges, depending on whether the data were collected in real-time or not. Real-time collection was strongly favored by the group. The advantages of real-time data collection centered around the immediacy of having the data available for review. The data review process can help check for odd data, bad data, ethnic bias, or protocol problems. This would allow any problems with the data to be spotted as soon as possible, allowing corrections to be made to the data collection process to improve quality.

The group discussed the keeping of audit trails to document changes and edits to the raw data. This is an established process; however, it was pointed out that the bigger the system, the harder it will be to audit the process. One group member mentioned a potential problem with audit trails in a wireless setting where a document might be copied to several different PDAs and each PDA user could edit the document before sending it on to its final destination or to another

user. The group agreed that document control needs to be considered in a wireless communications system.

The issue of implementing QA for self-collected data was raised. The group felt that if participants take self measurements for certain health outcomes, the study will need to develop a method to verify this process.

2.5.5. Other Types of Data

The workgroup felt that NCS should be aware of other types of data that need to be kept as part of any study. For example, all equipment used throughout the study should be documented, along with data about when it was used, calibration, maintenance activities, etc. Informed consent forms and other approval documents will need to be stored. A system for tracking the expiration of the consent forms will be needed, since most consent forms are only valid for 2 years.

3. ENVIRONMENTAL MEASUREMENTS WORKGROUP SESSION SUMMARY

3.1. GROUP ORGANIZATION

3.1.1 Group Participants

The Environmental Measurements Group was chaired by James Quackenboss of the U.S. Environmental Protection Agency (EPA). The participants in Environmental Measurements workgroup were:

Participant	Affiliation
James Quackenboss, Chair	EPA
Gerry Akland	RTI
Steve Bedosky	Levine Fricke
Bob Clickner	Westat
Kai Elgethun	University of Washington
Tom Dumyahn	Harvard School of Public Health
Lara Gundel	Lawrence Berkeley Labs
Joel Jorgenson	North Dakota State University
Juliana Maantay	Lehman College
Marcia Nishioka	Battelle
Anna Orlova	John Hopkins University
Haluk Ozkaynak	EPA
Charles Rhodes	RTI
John Terrell	Booz, Allen, Hamilton

3.1.2. Group Discussion Format

The format of the Environmental Measurements Workgroup was a series of prepared presentations, followed by open group discussion of each topic.

The main focus of discussion for this workgroup was the remote measurement, collection, and transmittal of environmental and exposure-related measurements. The environmental measurements breakout group was structured around three general topic areas (or sessions):

- Environmental measurements and samples,
- Questionnaires and interviews, and
- Location and activity patterns.

Examples of the types of exposure measures that were likely to be collected in the NCS were presented to the Workshop. These included:

- Environmental samples: air, water, dust;
- Bio-markers for chemicals: blood, urine, breast, milk, hair, tissue;
- Questionnaires, diaries, and health history;
- Serology and medical data;
- Housing and living characteristics;
- Family and social experiences; and
- Neighborhood and community characteristics.

The NCS core hypotheses, related to exposure needs and exposure-related information were also presented to the group. These included the following exposure-related hypotheses:

- Pesticides and risk of impaired neurobehavioral and cognitive performance;
- Early life exposures leading to neurotoxic effects and risk of injury;
- Indoor/outdoor air pollution & bacterial/microbial exposures and asthma;
- Early exposures to endocrine-active agents and risk of altered age at puberty.

Other, exposure-related information that is likely to be collected in the NCS was also presented, including:

- Dietary consumption;
- Physical activity;
- Time-location patterns;
- Demographic and housing characteristics.

The group discussed issues to consider for each type of measurement including the use of monitors in residential and other non-occupational settings, the feasibility and participant burden, the quality assurance and quality control issues, and potential future directions of the NCS. Presentations were given within each topic area, and were followed by discussion. Following each session, the group identified and discussed recommendations to be presented to the full workshop. The following summary combines issues raised during individual presentations with the recommendations developed by the breakout group.

The group also discussed cross-cutting issues that set the tone for how environmental and exposure information would be collected in the NCS. These included decisions to be made regarding:

- The use of paper versus non-paper forms and questionnaires. While all-electronic information capture is attractive, it was felt that an analysis should be done of the return-on-investment (ROI) for equipment needs and consideration of how acceptable the technology will be for study participants. The group suggested that, if electronic methods were used, then full-function equipment (PDA or laptop computers) should be provided, and participants should be encouraged to use them for other purposes.
- Self- versus Interviewer administration of questionnaires, etc. Considerations include ability to detect and correct reporting and entry errors, and honest reporting for sensitive issues.
- Tiered Monitoring. The group suggested that information from remote data collection (e.g., sensor data or usage information) could be used to "trigger" further detailed tests. Timing of sample collection may be important to understand the frequency and magnitude of intermittent exposure events. This could minimize data collection burden by collecting samples only when they are likely to provide useful information on changes in environmental conditions.

3.2. ENVIRONMENTAL MEASUREMENTS TOPIC

The first general topic area discussed by the group was environmental measurements and samples. The specific items discussed were remote monitors and sensors (e.g., devices that measure indoor air quality, store the data over time and transmit it to a central data location either by phone hook-up or wireless technology), air pollutant monitoring, samples collected by participants (e.g., miniaturized sampling devices), and pesticide monitoring.

The group suggested that the NCS may be able to use miniature mass spectrometers, gas chromatography equipment or other samplers, where one set of instruments could be used to sample in multiple homes. Trade–offs between two different types of approaches, van-based teams versus field sensors, were discussed. These included the need for trained technicians to set up and operate the equipment, the ability to routinely calibrate samplers with known sources, the availability of equipment with wider ranges of sensitivity. Sharing equipment between homes would limit the time periods covered in each home, possibly missing period of elevated concentrations. Time periods for sampling might be targeted based on sensor results, or reported

events in the home. It was also noted that using field labs or centers in a van or "hub" could facilitate quick turnaround on samples.

3.2.1. Remote Monitors and Sensors

The challenge identified by the group for the NCS would be getting quality data from the study participants into a data base for later analysis. The group noted that the more data that the study collected, the greater the value for researchers and identification of causal relationships not yet conceptualized. Microsensors could allow a more detailed data profile to be reliably collected. They could also provide a low impact method to collect data in a way that is transparent to the study participants. The sensors could be placed in the home or worn, and then not require further involvement by the participant. Some examples of existing sensors are acoustic, magnetic, seismic, environmental (pesticides and fertilizers), chemicals, and biological/physiological (e.g., sensors that monitor biological values and temperature, blood oxygen level, and blood sugar), and possible biological agents. The data captured on sensors could be relayed or downloaded to hub, and the hub could send data to the collection center by existing infrastructure. Microsensor memory processing could be stored and forwarded at night. The group agreed that collection of high quality environmental data directly from the subject could be some of the most valuable information collected during the NCS. For example, information on changes in the environment (e.g., pesticide or consumer product usage) could be used to identify time periods and locations where additional monitoring might be done to confirm the sensor readings or to determine the impact of residential concentrations on actual exposures or biomarkers. The point was made that the use of microsensors could automate and therefore improve the quality and ease of data collection.

3.2.2. Air Monitoring

A presentation was given on monitoring exposures to air pollutants, especially particulate matter (PM) for young children. Special consideration is needed for the mobility of children, especially when they start to walk, for the ability of the child to carry/wear a monitor or having an adult care-giver follow the child, and for acceptability of the monitor (weight and noise level). Healthy children (and parents) might be less likely to participate than those with asthma, which might lead to participant bias. Technical expertise is needed to operate active samplers (using a pump to collect the sample). Ethical and legal issues may be involved in informing participants about measurement results, especially for "high" levels.

The group discussed surrogates for personal exposure measurements including combinations of environmental monitoring, questionnaire data, application rates, diaries and other environmental information. The point was made that the NCS would have to measure where the child was being taking care of, such as, home, school, daycare, or outdoors, etc. If questionnaires are used, they would have to ask information about environment, contaminant or chemical sources in the environment, chemical usage, application rates, and usage rates. Diaries would need to collect information about activities in the home, activities of the child, and personal exposure information. Other information would need to be collected by the NCS and linked to environmental data. This information includes ambient air monitoring data, point source data, traffic volume, and land use patterns that could link to environmental data from other sources.

3.2.3. Miniaturized Sampling Devices

A presentation was given on miniaturized sampling devices which may be used to capture specific agents such as environmental tobacco smoke (ETS) and particulate matter (PM). A major problem raised about currently available real-time PM measurement devices is that these are currently inadequate because of their cost, size, and usability. A small portable sampler was described, which allows for collection over a one-week period, separated into three periods of the day. It provides information on total PM, and samples can be analyzed for ETS, diesel PM, and nicotine. Some lessons learned from the development of this sampler were shared with the group, including the need for lead time to develop and test new instrumentation, and the need for collaboration between state and Federal agencies to fund development of samplers that can be used in different studies. It was suggested that most current personal air exposure technologies for PM are probably prohibitive for NCS. That is, they may be too burdensome and/or analytical requirements too expensive without some developmental efforts to optimize them. Personal PM exposure monitors (and probably "remote" samplers) have poorer minimum detection limits (MDLs), precision, and accuracy.

3.2.4. Pesticide Monitoring

During a presentation on pesticide monitoring approaches, it was noted that the types of products now available for use in residential settings have changed from organophosphates (OP) to pyrethroids, and that formulations for products are rapidly changing. Pesticide monitoring for dislodgeable surface residues could be achieved using wipe, PUF roller, and press/scrub techniques. It was suggested that a scaled down version of a PUF roller would work well during

the NCS. The press/scrub was not recommended because of the low collection rate. Children could also wear socks for a period of time, and these could be collected and analyzed for pesticides and compared with other surface residue measures. Dust-bound pesticides might be collected using electro-statically charged clothes (which are sold for cleaning floors and surfaces).

3.3. QUESTIONNAIRE AND INTERVIEW INFORMATION

Presentations and discussions during this session were focused on three general topics: standardization and quality control, household activities, and dietary consumption.

The group identified the need for standardization but cautioned that the NCS needs to avoid being locked into an outdated technology. The group also suggested that a mature technology, such as web- and PDA-based data entry, and computer-assisted personal interview (CAPI) and telephone interview (CATI) techniques should be considered. Web- and PDA-based data entry could consist of drop-down menus, and most questions should be close ended check-boxes. The advantages of collecting data electronically are that information is captured at the right time, details are correct details, and data collection burden is minimized. For self-completion of questionnaires or diaries by study participant's consideration needs to be given to literacy and language barriers, and the acceptability of web- and PDA-based methods. For collection of sample-related information, the PDA can be combined with bar code readers or RF ID tag readers to minimize errors in entry of sample and subject identifier information. They can include form-based data entry, including drop-down menus to minimize coding errors. Including error trapping and detection features at the time of sample collection allows for problems to be identified and corrected.

During discussions about the use of PDAs for collection of dietary exposure information, it was suggested a full function PDA could be provided to the participant so that they can become familiar with the technology and as an incentive for them to provide the information requested. It was also noted that these could provide video capture of information, including portion sizes for diet.

3.3.1. Standardization and Quality Control

During a presentation on "informatics," the point was brought up that NCS will need to identify what data needs to be collected in planning for the multiple centers design. Standardization of data and forms will be important for the NCS. It was noted that there would be need to modularize for flexibility and upgrades. The simultaneous need to identify the

purpose (exposure – response) by specifying the output required, protect (reduce risk) for other (unknown) purposes, and work with limited specifications was discussed. The NCS could standardize the ways a questionnaire is administered, samples are collected and analyzed to provide consistency in the information that it captures This includes standardization of the forms used to collect sample-related information in the field (i.e., where, when, and how samples were collected). It was suggested that a model for the "data relationship" be developed by an expert team, and that this be followed by a usability study, and pilot testing of the data collection and processing system.

3.3.2. Household Activities and Dietary Consumption

There was a presentation on the use of PDAs to collect sampling data and information on dietary exposure to chemicals and contaminants. It was noted that aggregate dietary exposure to pesticides could have several components, such as pesticide exposure in drinking water; pesticide leaching to potable water supplies, and pesticide run off into reservoirs. It was suggested that household activities and dietary consumption data could be collected and transmitted through the use of personal data assistants (PDAs). An implementation could be designed that could read a two dimensional barcode to track samples. Transmitting data collected through PDA could accommodate multiple users, and maintain data integrity by transferring data directly to a database management system (DBMS). In such an implementation, data entry in the office would no longer be needed which would minimize labor cost associated with data entry and minimize the "human error factor" in the data cycle. Radio-frequency identification (RFID) technology could track consumer or pesticide products as they are being used. This could register when products are removed from a storage location, and help to identify where they are being used within the home. This could be done without the participant needing to manually record of what products they are using, or to record this using a PDA and bar code reader.

3.4. LOCATION AND ACTIVITY PATTERNS

Presentations were made for three topics in this session: time–location and physical activity data, geographic location data, and integrating geographic data and relationships.

Global Positioning System (GPS) units can provide information on geographic location over time, provided that the antenna is able to "see" the satellites. They are able to operate in single story wood frame houses, but may have signal loss in other types of structures. They can be worn on a vest or bandolier, and can be reduced in weight to less than one pound, including a battery sufficient for a 24-hour monitoring period. The resolution is about 3-5 meters

horizontally, and about 10 m vertically. They can be used to identify geographic areas where study participants spend time (e.g., relative to outdoor monitoring data), and can be compared with self-reported time-location diaries or used to help prompt for recall of locations and activities.

Other devices, including heart rate monitors and accelerometers, can be used to determine physical activity (exertion) level. Accelerometers can track changes in location in three dimensions, which provide a mapping of activities, movement, and path over time. These are relatively new technologies, so some adoption risk exists and they need to be evaluated relative to the information needed. It was noted that there are large differences in measuring adults versus children, and the location where devices are worn (legs, arms, wrists) may be important. The Robert Woods Johnson Foundation has sponsored studies of human activity patterns, and it may be useful for the NCS to contact them regarding their studies and methods.

The group discussed geographic information systems (GIS) that could link together large numbers of different pieces of data. This can include information on sources of environmental pollutants, such as the toxic release inventory (TRI), and information on the toxicity of different chemicals. The group discussed the possibility of integrating this approach with environmental monitoring and modeling information in order to develop surrogates or indices for possible exposures to multiple chemicals. There was some concern about the resolution of the modeling relative to the population data, and how to reflect daily mobility patterns (i.e., location of residence relative to work or school locations).

3.4.1. Time–Location and Physical Activities

Several examples of the type of time-location data that would be desirable for the NCS were discussed, including location, activity, path tracking, total distance traveled, and speed. Different tools were suggested that could combine multiple data collection activities being performed by one unit. Tools such as global positioning systems (GPS) units, heart rate monitors, or accelerometers can track activity, monitor movement and track the path. GPS has the ability to log path data, to store raw satellite code required to post-process differential corrections, and to import data for use with GIS software to overlay with aerial photographs or maps. But, there are limitation to GPS units such as locations with no reception (including concrete and steel structures), the need for an external antenna, acceptability and compliance with wearing the unit, equipment failure, differential signal base file errors, interferences (power lines and microwave ovens), and map errors. The point was made that the technology needs to match the activity being examined. For example, GPS will not measure physical activity

associated with stationary bike exercise. It was noted that some sensors would collect inaccurate data and that the NCS would need to correlate data from its proposed method with other methods. Also, it may not be feasible to have all participants wear these units, but rather these might be used on a sub-sample to help validate questionnaire and diary responses.

3.4.2. Geographic Data and Relationships

A presentation was given on the use of Geographic Information Systems (GIS) to relate sources of environmental pollutants to population disease rates. The NCS could develop GIS to map statistical data on health and disease rates and cases, conduct data exploration and data visualization, detect disease clusters, assess environmental hazards and exposure potential, analyze spatial diffusion and distribution of infectious diseases, determine access to health services, and an integration of modeling and statistical software with GIS. The NCS would need to identify population centers to monitor likely exposures using information sources such as traffic patterns, the toxic release inventory, existing monitoring data, or other sources. The point was made that GIS data can be leveraged to identify population centers. The NCS could develop a "surrogate capability" to link the population centers to individual participants to help estimate exposures for time periods when respondents that cannot be monitored during the NCS. The surrogate data from existing sources could then be related to individual participants.

3.5. TECHNOLOGY VALUE VS. EFFORT EVALUATION

The group agreed that the NCS will need to identify what methods will not be available immediately, but might be available later in the study. The group developed and discussed an environmental working group "Quadrant Chart" to characterize value versus effort for existing and future technology. A listing of all of the technologies identified by the group was compiled. This included: sensors for triggers, data capture, screening samples/field labs, PDA's sampling/questionnaires/bar codes, GPS/Accelerometers/GIS, RFID products, standards (open architecture), and portable instruments were ranked as being high in their potential value to the NCS and could be developed with relatively low effort. Portable instruments in multiple homes, bar code (burden/compliance), and PDA/Diary (compliance) were considered to have lower value and low effort. Pre-pregnancy enrollment, sensor data capture and developing new sensors (for PM, chemicals) are potentially high value, but with a high level of effort required. The group recognized the characterization of the technologies is based on many assumptions. The NCS will need to assess the value and effort of these technologies for this large longitudinal study of environmental influences on children's health and development.

Technology	Value	Effort
Sensors for triggers, data capture	High	Low
Screening samples and field laboratories	High	Low
PDA for sampling information, questionnaires, and bar codes	High	Low
GPS, Accelerometers, and GIS for location and activity information	High	Low
RF ID for consumer product and medicine use	High	Low
Standards (open architecture)	High	Low
Portable Instruments	High	Low
Pre-pregnancy enrollment	High	High
Sensors data capture	High	High
Sensors (PM, chemicals)	High	High
Portable instruments used in multiple homes	Low	Low
Bar code (burden/compliance)	Low	Low
PDA/Diary (compliance)	Low	Low