The National Academies

Workshop on

The Design of the National Children's Study

January 11, 2013

The National Academies
2101 Constitution Ave., NW
Washington, D.C.

Proceedings by:

CASET Associates, Ltd.
Fairfax, Virginia 22030
caset@caset.net

TABLE OF CONTENTS

Welcome and Introductions 1
Sara McLanahan 1

Welcome to the National Academies 2
Constance Citro 2

Welcome to the Workshop and Statement of Its Purpose 7
Steven Hirschfeld 7

Panel Discussion: Decisions about Environmental Measures 12
Moderator: Marie McCormick 12
Linda Sheldon 13
Melissa Perry 18
Nicole Deziel 21
Antonia Calafat 24
Floor Discussion 27

Panel Discussion: Composition of Sample: Alternatives for Cohorts of Women 60
Moderator: Barbara Carlson 60
Irwin Garfinkel 62
Naihua Duan 71
Nancy Reichman 79
Michael Bracken 87
Floor Discussion 96

Panel Discussion: Weighting, Imputation, and Estimation in Proposed Design 123
Moderator: Steven Cohen 123
Graham Kalton 125
Colm O'Muircheartaigh 132
Richard Valliant 141
Floor Discussion 160

Panel Discussion: Factors, Issues, and Values to Balance and Consider in Reaching Decisions about the NCS Design 182
Moderator: Greg Duncan 182
Ed Sondik 183
Roderick Little 190
Ana Diez Roux 197

Floor Discussion 209

Adjourn 217

P R O C E E D I N G S (8:35 a.m.)

Agenda Item: Welcome and Introductions

DR. MCLANAHAN: I am Sara McLanahan and I want to welcome all of you to this workshop on the design of the National Children's Study. So I'm not going to introduce people, I'll let the moderators of the panels do that. And there are also pretty extensive bios about all the people in your folders.

But I do want to give a few thanks and give a few comments about the organization of the workshop. So in terms of thanks, I want to thank the Steering Committee who worked really hard to put this panel together and they worked during the holidays, so thanks very much to them. I want to thank the panelists for agreeing to come and talk to us and at such short notice. And I want to thank Connie and Nancy from the Committee on National Statistics and the people from the National Children's Study Program Office for providing us with lots and lots of materials.

In terms of the comments, you'll notice we have four sessions. They are organized around measures, the sample, data and then the last session is sort of the big big picture. We've asked the panelists to each keep their comments to about 10 minutes and then there will be some discussion.

We've left plenty of time for questions from the

floor but I would ask that the panelists stick to their time and I would ask also that the people in the audience also make their questions short, keep them on target, and make sure that they share the time with other people in the audience and I'm asking the moderators to try to keep everything focused and on target.

So with having said that I think Connie's next.

Agenda Item: Welcome to the National Academies

DR. CITRO: Thank you. I am Connie Citro, the Director of the Committee on National Statistics and I want to welcome everyone in the room and I believe we have a number of people listening in on the phone, to the National Academies under whose auspices we're doing this workshop on key scientific issues in the design of the National Children's Study.

And a number of you are probably familiar with this but some of you may not be. The National Academies actually comprises several interrelated organizations of which the oldest is the National Academy of Sciences, an honorary, self-perpetuating, nonprofit organization of distinguished scientists founded actually 150 years ago with a Congressional Charter signed by President Lincoln in 1863. And the Charter stipulates that the NAS is to provide independent, pro bono advice on request by the government on matters of science and art, where art really

meant sort of technology and engineering.

The first assignment successfully completed was to figure out how to make a compass work in an ironclad warship. Now during World War I the assignments for the NAS increased to such an extent that President Wilson issued an executive order authorizing the NAS to set up the National Research Council as its operating arm with staff and all the facilities needed to enable the NAS to provide advice pursuant to its charter on a large scale.

Members of NAS consensus panels and steering committees, such as for this workshop, continue to serve pro bono for which we are very grateful to them but government grants and contracts pay for everything needed to facilitate their work. Except I'm obligated to make the following statement which is that although the National Institute of Child Health and Human Development (NICHD)is the one who requested this workshop and has generously supported it, neither NICHD nor any part of NIH is paying for the food that's available and going to be provided to keep our brains alert and our blood sugar up. And if anyone wants to know the whys and wherefores of this statement, catch me at a break or something or Alan Guttmacher who is here is the Director of NICHD.

The request from NICHD to organize this workshop came to two standing units of the National Academies, the

Committee on National Statistics, which I direct which is in the NRC Division of Behavioral and Social Sciences and Education and the Board on Children, Youth and Families which is under both DBASSE and the Institute of Medicine. Kimber Bogard, Director of BCYF, I don't know if I see her but I know that she joins me in welcoming you all.

We previously collaborated, the two units, back in 2007, 2008, to complete a fast track review of the NCS study design at that time and several people in the audience and who are participating served on that panel. And a principal recommendation of that review was to use the Vanguard sites as true pilots which we are just delighted to see happening and we are honored to have been again asked to look at the National Children's Study as it enters a critical phase of moving from piloting to implementation of the main study.

Now the request from NICHD this time was for a workshop on a fast track, not a consensus panel which even on a fast track tends to take longer, to hear from a range of experts - including those not previously involved with the NCS which is hard because almost everybody having to do with child health and development has touched the NCS at some point in time - and to discuss several, not everything, but several key design issues. We are transcribing the workshop and we will provide a verbatim

transcript to NICHD as soon as we get it from the transcribers.

We will also produce a staff authored summary of the workshop that will go through the standard National Academies review processes and then be made widely available. The summary will not and cannot under National Academies procedures attempt to infer consensus recommendations from today's discussions.

But the hope is that the workshop sessions will not only air a range of views on the topics to be discussed but also encapsulate for NICHD the most important tradeoffs to consider in reaching final decisions on the design of the main NCS study and to identify key elements of evidence that NICHD needs to seek from its piloting work to inform its decisions.

I want to add to the thanks that Sara has indicated. Nancy Kirkendall, the CNSTAT Study Director for this project, worked tirelessly to organize this workshop, get the Steering Committee together and then the participants working through the holidays. She's been ably assisted by other CNSTAT staff. You saw Agnes Gaskin out there I believe and Michael Siri right here.

Profuse thanks are due to the members of the Steering Committee and to those agreed, often on very short notice, to serve as panelists. And then the NICHD staff

were extremely helpful, gave a lot of time and energy to responding to the numerous questions of the Steering Committee. Steve Hirschfeld and I want a name specifically Jennifer Kwan who was just a tower of strength to us.

We have a busy day ahead of us. Hopefully all of you had a chance to look at the background materials that were made available on the CNSTAT and the National Children's Study websites. As Sara said, the moderators have been asked to keep their panelists on schedule and to keep the panel and floor discussion on topic.

So as not to interrupt the flow of the day, our phone participants are in a "listener only" mode. However, if any of you on the phone have a burning question or issue, I'm sure you know somebody in the audience, send them an email or something and they can get up and then pose your question.

But I indeed again want to think everyone for helping make this possible and I look forward to a stimulating discussion that will help NICHD reach evidence based decisions on key elements of the National Children's Study's design. Thank you so much.

And I believe Steve is now next to give us just some background on what NICHD is looking for from this workshop and what are the topics. Again, we're not covering everything but we are covering absolutely key

elements of the design, namely the environmental measures to collect, the sampling strategy and the implications of the sampling strategy for other statistical aspects of working with the data. So welcome and thank you.

Agenda Item: Welcome to the Workshop and Statement of Its Purpose

DR. HIRSCHFELD: I too add my welcome and my profuse and profound thanks to a lot of people. I will begin with the program office and Jennifer Kwan and then we'll extend to the National Academy of Sciences and Connie Citro and particularly Nancy Kirkendall who I don't think took 5 minutes off during the holidays. We were getting emails on Christmas Eve and all through the weekends and we want to acknowledge the dedication that the staff showed in putting this workshop together.

To the Planning Committee chaired so ably by Professor Sara McLanahan and all the members who not only got involved in planning but in thinking and our planning discussions turned out in some ways to be very compelling and sometimes exciting discussions on weighty matters and intellectual and theoretical questions which were no where on the agenda but that's the way the discussions flowed and it was a privilege to be able to participate in those discussions. And there are many other people who helped contribute to this workshop and I know we're all eager to

begin.

So I was requested to give the charge and the charge begins with the Children's Health Act of 2000 and I won't read all the language here and I assume these slides will be made available and anyone can refer to them. But in essence, what I wanted to highlight is that we are asked first and foremost to plan, develop and implement a prospective cohort study from birth to adulthood. So that was one charge from the law.

Then in that study we are to incorporate complete assessments, gather data from diverse populations and consider health disparities. And we may include the consideration of prenatal exposures but we took that, since it was mentioned twice, as something of a mandate even though there was a conditional around it and so these are the characteristics that we take as where we should go forward.

So the design had a period of discussion and planning over much of the last decade. That resulted in 2009 in going into the field with something which was a multistage modified national probability sample using selected geographic areas as primary and secondary sampling units. So the primary sample units were selected geographic areas, and within those selected geographic areas were then geographic segments and then using door to

door household recruitment.

And our field experience taught us early on that there was a divergence between what was expected and what was actually observed. So that led to a reexamination of our options because our projections were that we were going to go beyond the timeframe and exceed the resources that had been anticipated for this particular phase of the study and for the study as a whole.

And so in examining options, the NCS was informed by the pilot data, supplemented by participant, community and expert input. The input for how we went forward was developed through a data and consultative process.

We held eight structured workshops and conferences since 2011, five of which were open to the public. We had the direct involvement of about 30 statistical experts, in addition discussions with a wide spectrum of professional organizations and individuals. And of course we had the involvement of our NCS Vanguard investigators with weekly teleconferences, monthly executive steering committee meetings, semiannual face to face meetings.

And then we requested and received a written proposal in 2012 which we paid for and which we discussed for collectively hundreds of hours and incorporated in to our own thinking and evaluation of where to go forward and

multiple informal discussions.

And the general design principles which have emerged is that we are anchored in a national probability sample, that we will have recruitment through health care providers and we will have, within the recruitment cohorts, a birth cohort as we're instructed to do in the law. And that birth cohort will be recruited via hospitals and birthing centers because 97 to 98 percent of births in this country occur at hospitals and birthing centers. So that would give us from many perspectives, the least biased population to draw our sample from.

And then a prenatal cohort where we would use the hospitals and birthing centers which had been selected through this national probability sample. We would then extend out to the prenatal providers and clinics that refer to the selected hospitals and birthing centers to enroll the prenatal cohort which we know will not have the same lack of bias as a birth cohort because different women seek prenatal care at different stages of pregnancy.

So if we're interested in early events, we would get one type of cohort. As we look at later events, we would get somewhat different characteristics of the cohort. And so this is how we are thinking.

In addition, there are resource questions because there are more visits that are incorporated into prenatal

data collection than there would be from birth. Then we would have to ask what is it that we're collecting, what is it we are learning and what are the tradeoffs in how we achieve a balance between the birth cohort and the women that we recruit then and the women that we recruit prenatally.

And the other general design principal is that we have a sample size of 100,000 of which we would target approximately 90,000 for the birth and prenatal cohorts and then an additional 10,000 would be reserved for preconception cohort in additional targeted populations.

So this is the framework that we are then asking some questions on today and the focus is on scientific integrity and points to consider for making decisions. We are not seeking a specific decision or recommendation but want to hear a range of options and more importantly why.

We'd like advice on specific technical questions related to cohort proportions and the types of environmental data collected. We would like to examine the relationships among prospective prenatal data collection, the types of samples to be collected and relative proportions of the prenatal and birth cohorts. We would like to identify considerations for prenatal data collection such as generalizability, costability, and you provide and we receive some advice on how to prioritize

these considerations.

And I won't go through the specific questions, they are in your program but we have a series of questions for each session with the final session devoted to, from the discussion, can you identify and synthesize the tradeoffs among factors, issues and values that the NCS leadership and NIH leadership need to consider and balance.

So that is the charge for today and I thank everyone for their prior input and we're looking forward to an invigorating and stimulating discussion that will bring us to the next stage.

Agenda Item: Decisions about Environmental Measures

DR. MCCORMICK: I am Marie McCormick and I am the moderator for the first session and I would ask the panelists to come up to the table. As indicated in the slides, we're to address the question about the proposed measures, biomarkers, questionnaires and physical measures that are most appropriate to address the questions of interest.

We're going to organize this a little bit differently than Sara described. We're given actually two questions to deal with. The first question: Are the proposed measures, and there were a whole list of those provided us, the most appropriate to assess exposures of

interest and if not, what measures would be taken?

Each of the presenters will then have 5 minutes to address these questions and then we will have 25 minutes for discussion.

The second question about how the National Children's Study should prioritize decisions, again we'll have 5-minute presentations from each of the panelists and then 25 minutes of discussion.

And we're actually not going to introduce the panelists. That information is all in your packet. We're going to move straight forward into the content of the meeting. And so I'll begin by asking Linda Sheldon to provide a brief overview of the Exposure Workshop that NCI did for the National Children's Study.

DR. SHELDON: Good morning and thank you. What I wanted to do is to talk about a workshop, and actually we at EPA with NIHS put together. It was nearly 3 years ago looking at exposure metrics for the NCS. And I am the Associate Director of our National Exposure Research Laboratory. It was primarily focused on exposure and what the exposure community thought was a way to look at it. The report is available. I do recommend that you read this report.

The charge that was given to the people in the workshop was to develop innovative exposure metrics and

look at the minimum amount of exposure data that you needed to collect to be able to answer the questions. We felt that this was extraordinarily important, not to look at every way you could measure exposure but what was the minimum that was needed.

What we did is that we looked at three areas of health linkages with an exposure. We looked at air pollution and asthma, insecticides and neurological development, and endocrine disrupting chemicals in reproductive endpoints.

Three separate workgroups were formed before the workshop and each workgroup had an epidemiologist, a toxicologist and two people that worked in exposure. So that in fact we would have this cross discipline coordination as to what were the chemicals, what were the time periods of susceptibility, knowing that, what is it that you would be able to do?

There were a series of teleconferences for each of these committees and then we had a day and a half workshop, where about 50 people attended, where each of the workgroups presented their findings on chemicals of interest, sorts, routes and pathways of exposure, critical time windows, biological samples, environmental samples, non-measurement approaches, protocol recommendations and research recommendations.

And I think that actually the good news is that much of what we recommended is pretty similar to what the protocol is now. I think there's a lot of discussion background justification for what we were doing. I think that one of the important things that was discussed before we started was an exposure metric, what was it and what made a good exposure metric.

To us, an exposure metric was not necessarily just a measurement. It could be a measurement combined with other data to model that would give you the ability to estimate an exposure. We felt that, as all epidemiologists know, true exposure gives you the best chance of being able to find an effect.

When we looked at what would true exposure be, it would be biologically relevant exposure during the entire time window of susceptibility and that's a very difficult thing to get to. An exposure metric might be a biomarker in urine and it would be if in fact it related to exposure. A biologically relevant exposure metric is one where that concentration in urine will tell you something about the concentration at the target, the biological target, where in fact the effect would take place. Blood lead would be an excellent biologically relevant exposure metric.

And we discussed this because that was going to be the standard to which all other things were measured

against. I think there are two things that are important here. Does the measurement lead to biological relevance? But the really more important consideration for this study was what would a sample collected on one day tell you about exposures over the entire period of susceptibility? And that is going to be extremely important, especially when you're looking at prenatal exposures. And I think that when we discussed this as a workgroup this became a very important thing in terms of decisions.

Some of the recommendations that we came to in terms of time windows of exposure, first trimester everybody thought was very important. We understand you may not be able to do it, so how do you relate those measurements collected during the third trimester perhaps to the first trimester. Third trimester was considered important. First year after birth was important. Years 1 through 4 should be considered annually, if possible. And for endocrine disruption, samples should be considered at puberty.

For biological matrices, blood for the mother was considered important for the first trimester and third trimester. The child, there was a lot of discussion about this, everybody wants to use the blood sample from children. Recognizing that there's only very limited blood you can get from a child - I remember seeing a blood draw

from my son when he was a year old and it was horrible.

Everybody's going to want those blood samples. There is very little blood sample and so we felt it was going to be very difficult to recommend blood samples for children but they were very important if you could get just one but use it wisely.

Urine, they are easy to collect and we felt that for the mother and the child it was good. Breast milk may be a good substitution for a lot of things.

Environmental samples, house dust was given absolutely the highest priority. Single samples can be collected during the same visit. We need to measure not just the concentration in the dust but the dust loading. So that is what is the concentration per square meter or something, it's not just good enough to get the concentration, you need a standard method, vacuum methods appear to be most feasible and a protocol is needed to be developed to address collection of a single sample from multiple analytes. This is very important because you can get organics, you can get metals, you can get biologicals but you've got to think about how you could get all of those samples in one sample.

I think I've probably used up my 5 minutes. So I will stop there. We have a presentation that we have given. There is a workshop report. I really do highly

recommend that people look at this. Thank you.

DR. MCCORMICK: Melissa will reflect on Linda's presentation and ask questions to what extent the recommendations of the workshop have been incorporated into the current design.

DR. PERRY: I have the distinct luxury of being what I would call a critical cheerleader, critically-minded cheerleader in the sense that I don't have a history with NCS and so I'm coming in with a brand new pair of eyes and perspectives. It's been an interesting foray into all the work that has been done. I've been very aware of the wealth of expertise that has come in to contribute to where the NCS is so far.

So what I considered my job to be was to size up the current environmental sampling approach and give critical perspectives on how does it look and where could it be improved as someone who is experienced in environmental health exposure assessment.

And when I read the report that Linda just described, it seemed to me as though a number of scientists of good will and stature had contributed a lot of careful thinking as to what the plan should be. And when I looked at the background information that we were given, much was there, as Linda said, but a couple of other considerations should be made.

One has to do with the use of questionnaires and I know that has been actively discussed, especially as it pertains to environmental exposure assessment. Are they relevant? And one only needs to look at the Agricultural Health Study that relied largely on questionnaires and the wealth of information that it has produced over the years to say that questionnaires do have their place.

At the same time, we are all considerate of what kind of respondent burden that might create and therefore there are instances where specific exposures would be of concern and a limited set of questions could be asked in the questionnaire. So my message is to not completely disregard the use of questionnaires for exposure assessment but rather to maintain a minimum number of questions for specific key exposures that you may not be able to do using a biologically relevant exposure metric.

Secondly, looking at the chronology as it's laid out currently, what the plan is for sampling of biological matrices as we call it, urine, blood, blood from the infant and also cord blood. These seem to be very well put together and well timed to the extent that we understand windows of susceptibility.

I am appreciative of the challenge that the investigators have experienced so far when it comes to the prenatal cohort. But the reality is is those are critical

windows of exposure that in years to come, if we can collect the proper matrices, we will be able to shed light on what these mechanisms are and we won't be in a position of saying why didn't we collect it.

The issue of breast milk in fact was not prominently figured in the current exposure collection plan and yet again I would have to endorse that plan, understanding the challenges of breast milk collection. Linda had mentioned the single blood spot for the newborn. By all means, let's collect it and plan accordingly, plan carefully as to how that's going to be used but let's collect it.

Antonia's going to talk about the relevance of metals and how that currently is not one of the key contaminants that are going to be evaluated. But the fact that the sequencing chronology of possibly first trimester, when at all possible, certainly third trimester, at birth and then going forward, looks to me to be a logical and a well thought out trajectory of sampling over time. So I would certainly endorse that.

We're going to come back to opportunities for knowledge gaps. House dust, as Linda said, should not be minimized by any means. It could be so informative. If we think about how much time we spend in indoors, house dust is such an important biologically relevant matrix of

exposure. And I will mention when we circle back around, the relevance of personal monitoring and where we're going in that area into very affordable ways of doing personal monitoring.

So let me just use that as a preview of what we'll do in the second round.

DR. MCCORMICK: Thank you. Nicole Deziel will now discuss issues in the mode of collection and temporal variability.

DR. DEZIEL: Thank you. I come to this panel with the perspective of an exposure scientist and environmental epidemiologist who developed many of the environmental sampling protocols for the Vanguard study. And now, at the National Cancer Institutes, I evaluate methodological issues with exposure metrics such as comparing surrogate measures of exposure with actual biological and environmental measurements and surrogate measures could be questionnaires or GIS-based modeling.

So we're being asked to judge the appropriateness of the proposed exposure measures. The appropriateness of the timing and method of sample collection is really dependent on what the research question of interest is. We don't have a specific research questions under this current model.

Rather, the NCS is doing a new approach, this

broad-based approach where they are going to collect lots of detailed information to address numerous future research questions. So given this proposed model I would say that the repeated dust, blood, and urine measurements do seem to be appropriate and strong metrics to collect.

I'd like to echo some of the benefits of dust, as has already been mentioned. It's something that I work with quite a bit at the National Cancer Institute. It can be very useful in providing information for chemicals for which we don't have good interview or questionnaire questions.

I do want to also echo the value of questionnaires as Melissa Perry noted. Also with respect to pesticides, we've seen that very specific well-designed questions about pest treatments, for example, do you treat for termites, to you treat for fleas and ticks, asking those types of questions does provide good correlation with actual measurements of the expected active ingredients in dust.

But some things you just can't ask about. For example, you can't ask people if they have PCBs in their home or if they have polybrominated diphenyl ether flame retardants in their TVs or couches. So dust can provide useful exposure information for those types of chemicals where there aren't good questions.

In addition, we at the National Cancer Institute and folks at Berkeley and EPA and others have shown that even non-persistent chemicals tend to be rather stable over time once they are in the residential environment. For example, some of the work that we've done, we've observed relatively high interclass correlation coefficients, 0.6, 0.7, 0.8, 0.9 for pesticides, polycyclic aromatic hydrocarbons, PCBs, and even some PBDE flame retardants.

So if these are the analytes of interest for the NCS, I think that could provide useful information during critical time periods such as preconception and the first trimester when we're unlikely to actually have samples collected.

However, based on some of the tables and documents we've been provided, it's not clear to me what the method of dust collection is and I think that's very important to consider. Are we going to collect vacuum bags, HVS3 samples, a subtle dust plate, a dust wipe, an air sample. Almost all of those have been on the table as part of the NCS sampling protocol in the past so which one of these is going to be the method of choice and I would hope that some sort of bulk dust sample would be collected.

In some of our research at NCI, we've compared concentrations of chemicals from a participant's vacuum bag or vacuum canister with a more standardized vacuum

approach, the HVS3 and we've seen very good correlation between the two methods for a range of chemicals, like pesticides, PCBs, and PAHs. So I think a vacuum bag is a nice compromise between getting useful exposure information but also feasible and not too burdensome on the data collector or the participant.

The other point I want to make is that I didn't see anything in the documentation about GPS measurements or getting a good residential history and so I would really emphasize the importance of that. If you can get GPS coordinates at a house and just ask a few questions about how long someone has lived in that home, you can take advantage of the growing number of rich publicly available databases that can provide some useful information about exposure like EPA's TRI or NADA databases. Some states have pesticide use databases and again the advantage of this is that you can get some exposure information during critical time windows when you may not have actual samples collected. Thanks.

DR. MCCORMICK: Finally, Antonia Calafat will talk about some of the current experience that has emerged from the Vanguard sites.

DR. CALAFAT: Good morning, I am the last speaker of the panel now and then I just can put together a little bit, just fitting the biomonitoring part that is the part

that I'm most familiar with.

Linda had mentioned that we are looking for the minimum data that would provide useful information. We all need to be very mindful that not one approach is going to give us everything so it’s going to be very important to get a minimum set of environmental measures, a minimum set of questionnaire information, some environmental data, residential data and, last but not least, biomonitoring information so biological specimens.

In the sense of the biological specimens that we have been talking about blood, urine and breast milk, we need to remember that each one of these matrixes is appropriate or most appropriate for certain chemicals. We tend to measure persistent chemicals in blood. Many of these persistent chemicals are also lipophilic, they partition into fat, so breast milk would be an excellent matrix for assessing just prenatal and postnatal exposure to some of these persistent chemicals.

In the sense of the non-persistent chemicals, which unfortunately is the direction that the market is moving, just moving from persistent into non-persistent chemicals, we have a tremendous challenge. These are non-persistent. That means they just metabolize rather quickly and for the most part we are exposed to these chemicals through episodic events.

So when you combine a non-persistent chemical with episodic exposures such as the ones that you would encounter like through diet, and they're not through the use of personal care products, then you have this issue of variability. So concentrations of these chemicals, we measure them in urine, and concentrations really go up and down, that is tremendous variability.

And how can we say that one measure taken today is going to reflect exposure like the critical window of susceptibility or is it going to reflect exposure that you're going to have later this afternoon or you're going to have in 2 days.

This is something that there's really no perfect world, there's nothing we can do about it. That's the nature of the beast if you want the chemicals that metabolize quickly and they're episodic exposures.

So one thing that we could try to do is try to collect as many samples as we can and also collect information, not only on when the sample is collected or provided, but also the time of the last urination. That could be also important when we try to put everything together.

In tying this into our experience with some of the samples that we have already analyzed for the initial phase of the Vanguard pilot, we analyzed samples for about

500 women. This is nationally representative and these urine samples in particular have been analyzed for a suite of different chemicals including some phenyls, phthalates and metals.

We have data that are already being processed as we speak. And despite the variability that I have mentioned before, there we are already seeing very important differences in concentrations of some of these chemicals, depending on the demographics of the population. So variability is going to be important. We need to consider it but because these measures are variable doesn't mean that they are not useful.

Melissa talked about the metals and we're going to touch upon this later on the second question. But it is important to consider that when we're collecting all these samples, we need to keep in mind that in some cases these chemicals are very prevalent and they are everywhere in the environment. So we need to make sure that the materials that we use for the collection of the samples are not introducing the target chemicals of interest. So an issue of potential prescreening of materials needs to be considered. And these would actually be very critical if we want to measure metals later on. Thank you.

DR. MCCORMICK: Thank you. This first question about are the proposed measures appropriate to assess the

exposures of interest is now open for comments and questions from the audience. And I'm assuming you want people to use the microphones on either side.

DR. DUAN: I'm Naihua Duan from Columbia University. I appreciate the very nice comments. I'm on the second panel and we have many questions we are looking forward to guidance from your panel and I think the priorities that Linda commented on, like the first trimester, third trimester, is really important for our deliberations.

I have a question and would appreciate your thoughts on how well can the first trimester exposure be assessed through retrospective recall. So assuming that it will be difficult to get a good preconception sample and we have to rely on prenatal sample cohort and a birth cohort which largely miss the first trimester. I think Nicole commented on the interclass correlation over time.

And I would appreciate if you can elaborate a little more on that, whether you will be able to recapture the first trimester exposures very well through recall or whether there are some gaps that need to be filled in that also can be addressed. Thank you.

DR. PERRY: One initial response to that is, well, it depends on the contaminant that you're talking about. And Nicole had mentioned that oftentimes the more

ubiquitous invisible compounds around us, no one's fully aware as to when they are being exposed.

At the same time, in your first trimester, one can certainly imagine remembering a pesticide event or using a paint or solvent so that would be a very specific period, especially if you're asking the woman immediately postpartum, what was happening in the first trimester. There would be a real recency there. But when it comes to the persistence,that would be problematic. You want to comment more about that Nicole?

DR. DEZIEL: Sure, just to put a timeframe on the studies that I mentioned. In many of the studies we looked at samples collected months or even years apart and still saw interclass correlation coefficients of 0.7 or higher for many pesticides, PAHs, again PCBs, and the repeatability of these samples though will be dependent on the physical chemical properties of the chemical but also consistency of use. But we did see over a couple year timeframe that a single sample may be representative of a period of months or years for some chemicals.

DR. PERRY: How about recalling PBDEs for example?

DR. DEZIEL: We're actually looking at some questions about how well people recall pesticide use during different time periods of pregnancy and how well that

correlates with the dust but I think that could be challenging.

DR. SHELDON: So one of the things that was brought up in our workshop is that every time somebody moves, you need to recollect samples because they may be persistent in that one environment so you do need to know whether or not they were in the environment where you are collecting samples, that’s a minimum question.

DR. DUAN: If I can follow up on the point Linda, or if you wouldn’t mind me to comment, Dr. Garfinkel is going to comment on the importance about the first born in the next panel.

So you mention mobility and I'm wondering whether we know much about the mobility during this critical period of time in the family, especially for the first born during the time they prepare for the first child arriving, the mobility. Do we have some good information on how much mobility is that might compromise the persistency of the exposure?

DR. MCCORMICK: I think actually it’s fairly high. Our experience is that particularly a young couple, particularly for first born, moved into a house because they now have a child so I think mobility is high. Sara, is that your experience.

DR. MCCORMICK: Yes, it’s high.

DR. DEZIEL: Just to follow up on that, I think that also highlights the importance of taking a residential history or at least asking how long have you lived in this house, that way you know the sample what timeframe it's relevant for.

DR. BRACKEN: Michael Bracken, Yale University. Actually, in our studies, 30 to 40 percent of families move after a birth within 2 years so it is an issue in mobility. I wanted to make a comment and a question.

The comment is it's very interesting that 95 percent of what you've talked about is prenatal testing and obviously your panel is very interested in testing during pregnancy.

It seems when we talk about these issues, we're like generals. We're always fighting the last war when we talk about lead and air pollution. But of course in a 21-year study, we've got to be anticipating what questions will be asked and we have frankly no idea what they will be.

So the questions for me are to think about sample storage, collection, stability, certainly where they are being collected. GPS units are a great idea. But they are going to be looked at in terms of gene environment interaction or epigenetics, that's clearly where a lot of these samples are now going to be applied in thinking about

those models.

So it's crucial that these samples are preserved for the decades and I wonder if the panel has really given thought to how that's actually going to happen. Because we know many samples are not stable and actually deteriorate over time.

DR. CALAFAT: This is actually one we had another question that we were going to just cover this, the collection and storage of the samples in our answer to the second question.

DR. DUNCAN: I am Greg Duncan from the University of California Irvine. I ran a national study for many years and the mobility rates in the national sample were just under 20 percent per year. In general, they are higher among low income families than high income families. They are higher among younger families. So there will be a lot of geographic mobility. Most mobility is local within county but still it's a change of household residence.

My question, one of the design options under consideration is to recruit some of the children, either from a hospital or prenatal providers and then have subsequent births to the original mothers become part of the sample as well. The advantage of the subsequent births is that they provide preconception, as well as very early prenatal information on exposures for these subsequent

births. A problem is that these are all second and higher priority births so you would not be able to use this kind of subsequent birth technique to get preconception and early prenatal information about first births.

So my question is I didn't hear any discussion at all about different hypothesized affects of exposures for first births versus subsequent births. Is that an important distinction and are there reasons to think that it's very important to get exposure information on first births quite early in the preconception period as well?

DR. PERRY: Actually as we prepared for this panel, we did have a discussion about that idea, and that's probably why Dr. Bracken, you heard almost this consensus among us echoing the importance of the prenatal cohort. Because we understood how logistically challenging this is and at the same time critically important for generating a wealth of new information about early exposure, in utero exposure related to even the new findings about prenatal bases of adult disease.

So one flaw I saw with this notion of if we had to forego prenatal sampling in anticipation of that second or third born was the fact that those individuals that might have children that have health problems may not go on to conceive and reproduce again. So there's a lost opportunity there where you wouldn't be able to study those

affected by an immediate or a chronic disease. That would be problematic. So I think that was the major concern that we had - that we wanted to send the message let's not abandon the prenatal piece.

And I know there's been active discussion about preconception and how very challenging preconception sampling could be and I think there are some design ideas, missing by design as well as some validity opportunities where you could do subsamples for preconception sampling as well.

That's not to send the message that we don't care about early exposures in infancy and early development. We certainly do, as it relates to predicting adult disease but that early period few people have been able to realize in a large sample. That's where we are positioned to do something very significant.

DR. DEZIEL: I would just add that by recruiting the second child, not only do you have unbalanced exposure information on these siblings but also those siblings will be correlated, their exposures may be correlated. So they would have to be analyzed separately or you'd have to use different statistical techniques so I'm not sure it would give you the statistical power to really look at that prenatal or preconception period. It's just a challenge to think about.

DR. CALAFAT: One other thing there could be some differences maybe between the first born and the ones that are born later, in terms of persistent pollutants. If the first born children had been breastfed so then some of the body burden that the mom may have may have gone into the first born and is going to be just a little late going into the others. That's one thing that we would have to keep in mind as well.

DR. GARFINKEL: My name is Irv Garfinkel. I'm also on the second panel. So I want to sharpen the question of whether what evidence we have, if any, then I take the first point that there may not be second births. That's a very important point.

But the question that I want to sharpen and get the answer is what do we know about the effects of exposures, does it differ for first and second births? Do we have any evidence on that question? That's the key. Am I clear?

DR. CALAFAT: We would say no. That's a very good question but at least I don't have an answer.

DR. PANETH: Thanks, Nigel Paneth, Michigan State University. Thanks for talking about the prenatal period as a critical window of opportunity. I just want to emphasize, one, that a big chunk of what's really important in child health, particularly in many of the

neurodevelopmental disabilities, birth defects and preterm birth are determined by birth. So postnatal environmental measurements are not weakly relevant, they're utterly irrelevant to the causation of those central components and that puts the onus squarely on the prenatal period.

And another thing I'd like to underline is recently I've been involved in the analysis of what they are now calling MOBAND which is the combined Danish and Norwegian birth cohorts which total 200,000. All have prenatal collections, all have prenatal blood, all have prenatal urine. None of them, none of them, have prenatal environmental house exposures. There's no dust, there's no house air or water. This is a unique possibility - the prenatal exposure environmental information is unique to the potential National Children's Study. So I think that opportunity should not be missed.

DR. MCCORMICK: Are there any other questions, comments?

DR. KERVER: Hi, I'm Jean Kerver. I'm with Michigan State University. I'm a registered dietician and I have a PhD in nutrition and I'm a nutrition epidemiologist.

I would just like to add diet back into this conversation, both because I was concerned in the background piece of information I saw that it said

something like a metabolic screen of serum total protein, BUN, calcium, iron and cholesterol may be used as a proxy for nutrition and dietary exposure. And I want to echo concerns that dietary exposure, like other exposures, must be collected in real time and it's very important I think to get it prenatally.

One comment I'd like to make about the design overall is that if we move from the 110 PSUs that reflect the great regional variation in food intake in this country, we will vastly decrease the variation we see in not only nutrient intake and food intake among different communities in this country, but also the different pesticide exposures through food.

And one other comment about the subsequent birth is that if we're considering prenatal biomarkers in subsequent births in some cases, for example, vitamin D, other fat soluble vitamins, you will see a big difference by parody based on the interpregnancy interval because the woman will have a decreased nutrient source after her first birth if she doesn't have time to build those up before the second birth. That is a big consideration of mine in going to a design that would eliminate first births by design. Thank you.

DR. SHELDON: So I think on this question there are two reasons to collect dietary information. One is

nutrition and the other one is exposure to contaminants. I think that nutrition is very important.

All the work that we have done on dietary exposure to contaminants, there is so much variability day to day for any particular individual, other than breast milk, that at least this group in the work I've done really only shows that you can get the extremes. So the question was organic food, some other or other kinds of dietary behaviors that you would think would be extreme because any single day of measuring pollutants in your diet are just going to be hugely variable. But you are more talking about the nutrition part and I think that nutrition does need to be put in.

DR. MCCORMICK: Any further comments, questions, panelists? Well, moving right along ahead. We will go to the second segment of this session which asks how should the National Children's Study prioritize decisions regarding exposure assessments in both the strategies for collecting these as well as some of the other issues that will be discussed. We will begin with Linda Sheldon talking about integrated samples.

DR. SHELDON: I think that this is an interesting question because I think everybody looks at the strategy and says, at least I did when I started, that you need to have a hypothesis that you test and you work around those

hypotheses. And as you start to look at the different hypotheses and what's known about environmental pollution, I think that much of your scheme ends up being your sampling scheme of what is doable, what you can measure, what's there, is almost the same.

So I do think that you need to keep in mind what are health outcomes and again I think we've gotten that with the prenatal exposures and stuff. And then it's what can you measure and then what is going to be what we will call persistent or persistent for the time window that you are trying to estimate? Once again, you've got to remember that you can collect one sample in one day and it's got to represent an entire period and so we do need to be able to focus on those groups of chemicals that are going to have a reasonable correlation or interclass correlation with different time periods. I think that's very important.

Otherwise you are collecting data that doesn't necessarily represent anything. I think that that is true for some of the house loading measures. I firmly believe that we are getting better at developing new models for air pollution - we've got the home environment and then we've got the community and the ambient environment.

And I think that at least for air pollution for some other pollutants, we have ability to go back and retrospectively determine over periods with various

modeling techniques what exposure you get. But I think that this idea of what window and how can you estimate the exposure during that window just becomes absolutely critical in prioritizing what you're going to do.

I think that one other thing that we need to keep in mind is that it doesn't make any difference if you're doing exposure for epidemiology or exposure analysis for anything else - and somebody already brought it up - is that we always seem to be looking under the lamppost for those things that we have already looked at and we know how to measure and we know how to do.

I do think that there are very new analytical techniques that are coming on board that may allow us to at least screen for tens of thousands of chemicals in matrices like house dust. And I think that we need to be able to, as technologies move forward, we need to be able to look at those in some of our archive samples.

DR. PERRY: Hopefully the panel will forgive me because when we were preparing I didn't have a chance to vet this idea with you all and so it's a maverick idea. But Linda actually did provide a preface to it and that is the notion of technologies.

So when I heard you say thousands of chemicals in a single screen I thought of the exposure zone. And there is active progress in this country and certainly in other

countries in determining a way to consolidate multiple exposure measurements within one small sample of blood or urine and Dr. Bracken's point about many of the mechanistic opportunities here are dependent on our successful ability at collecting particularly those blood samples and cord blood samples as well.

But how about personal monitoring. We are on the cusp - this country is on the cusp of advances in personal monitoring in so many ways that have public health relevance. If you had asked me 15 years ago if I was sitting in a board room at Apple and the vision was that we would all be looking at these tiny, tiny screens and watching movies and television on them day in and day out, I would've been quite skeptical. Well, here we are and anybody that has a smart phone knows how powerful this is as an individual monitoring device. It's your own portable GPS.

So we're already at the place with respect to nutrition where individuals are wearing personal monitors, monitoring their physical activity, uploading it online or in real time, and then getting to see their data that's personally customized. We're not that far off when it comes to actual food intake and I certainly know that we're not that far off when it comes to exposure monitoring, particularly with respect to air.

So can we be rapidly enough thinking about personal monitoring. These are not going to be expensive devices. They are going to be cheaper than an iPhone. Can we be thinking about an individual monitor that our participants would be wearing to monitor at the very least indoor air going forward.

So I'm not suggesting that we spend the next 5 years studying that opportunity but we have to be fully aware of what that trend is and can we seize an opportunity.

DR. MCCORMICK: Antonia will talk a little bit about the tradeoffs among capable and perhaps current and future storage effort at our collection.

DR. CALAFAT: This is something that we had said several of us before that we are going to be collecting, we may be collecting, possibly and hopefully, a large number of specimens, both environmental and biological. Then it's going to be very important just to document exactly how these samples have been collected.

Because now we are thinking about these chemicals that we have the current capability of measuring but it is possible that maybe years from now there may be some other chemicals that we could say okay, we are interested in looking at whether these samples that we collected can help us assessing exposure to these chemicals.

This is true in some cases because the market moves frequently and there are chemicals that are being banned now and replaced by other chemicals - that we don't know, that are actually not on our radars because they may not be even here now yet. But we have 21 years to collect samples and could provide very important information at the same time, again just the documentation of how the samples have been collected because these samples may be analyzed, not only for environmental chemicals but also for biologicals or some type of just genetic information.

At the same time it is really important to keep in mind that we have to look at the tradeoff or the cost balance. We have to collect the samples, to store the samples, and we need to have some compelling evidence at least now that it is important to make the investment into collecting the samples and storing them.

One other thing is that maybe if the cost becomes really very high, one potential option would be just collecting samples or select samples or samples only for a subset of the participants, much as we do with, for example within NHANES. The National Health and Nutrition Examination Survey is a survey conducted by CDC that examines the general population but the chemicals, most of the chemicals and some of the other biomarkers that are only measured in a subset of the population that can also

be made representative of the whole US. So that's one other possibility, just collecting specimens in some cases for a subset of the population.

DR. MCCORMICK: Thank you. The issue of prioritization for exposure assessment is now open for discussion.

DR. DEZIEL: I would just add that the NCS should prioritize their exposure metrics based on some simple descriptive statistics like what's the percent detection in the population, do we have sufficient range of variability in the population so that we can have adequate statistical power to look at the questions of interest.

And I would point out that we have 1-1/2 to 2 years of pilot data available to try to inform the main study. But we have this very short turnaround now between the pilot and the main study and would just urge the NCS to mine that data as quickly as possible so that the pilot can really inform this prioritization scheme and those decisions.

DR. SHELDON: I would like to make a follow-up statement to what Melissa said about the personal monitoring. There are a lot of groups that now have very large NSF grants to look at how to do this and some of its really exciting.

A group we're working with at NC State is using

nanomaterials to generate the power to be able to collect the pollutant monitors and a physiological response which means when these are developed you just put a patch on somebody and you never have to go back. You don't have to do those repeated visits and the key is having this electronic transmission for long-term monitoring.

So there are really a lot of new things that will be coming on the market that again people should at least be looking forward to. Maybe not this year, but again I think you're right. There are some very exciting things that -

DR. DEARBORN: I am Dorr Dearborn from Case Western Reserve University and exactly on this point of both availability now on personal monitoring.

We have with NCS funding developed residential air monitoring parameters where we can wirelessly download through 3G continuous monitoring of eight different parameters of air quality and we're able to detect obviously when somebody lights a cigarette. We're obviously able to tell when somebody turns a gas cooking stove on. We can see the increase in hydrogen dioxide.

These are not the size you would think of personal monitors but we also collect air particulates with laser light scattering and we're about ready to put these into some NCS participant homes to get some field

experience with how they work.

The question that I have is with this sort of practical technology we could very easily add a microphone and collect sound, not for amplitude but for character and deconvolute the frequencies so we could get some sense of what the source and nature of the sound is again on a continuous basis.

The other thought is what about putting a photocell and collecting light. Now we're putting these in the child's bedroom and in the kitchen. So that's here now.

DR. PERRY: Yes, there's a wonderful quote and that is "the future is here, it's just unevenly distributed".

So I'll one up you on that about the idea of noise. What about tiny cameras? Now again, I'm just making the point that the technology exists and what kind of burden that places on issues of privacy are very real and I don't mean to dismiss that by any means but we certainly need to understand to what extent our participants are interested.

I heard the idea that again the fact that folks are willing to wear monitors for physical activity and have them uploaded, I probably would have pooh-poohed that idea as being infeasible and unlikely to get a big response and

we're seeing a tremendous response to that. So we really have to follow our participants' willingness to be part of this.

And the technology does exist by all means. And it relates not only to this notion of environmental contaminants and exposures but very much in the nutrition and physical activity realm as well.

DR. BRACKEN: Michael Bracken at Yale University. One of the things that struck me about the NCS is they have not really commissioned systematic reviews of things like technology. And rather than depend on the Vanguard data which is actually very limited, they could look at actually what's being done in other cohorts.

And I'll just tell you about two of my cohorts where we gave women monitors to wear in pregnancy at three different weeks, all using nested subgroups. As someone suggested, you can't do this in the entire cohort but you can do it in randomly selected subgroups.

One group of women wore a monitor to measure environmental tobacco smoke and in a second cohort it was to measure electromagnetic field exposure. And these are papers that have been out 15 years now so there is actually a wealth of data on how to actively monitor pregnant women throughout their pregnancies which is available. And I think some well commissioned systematic reviews of this

kind of exposure assessment would actually serve the NCS better than trying to rely on the Vanguard data where actually it doesn't exist.

DR. MCLANAHAN: I'm Sara McLanahan from Princeton University. I have a question - what's the tradeoff between the first trimester and the third trimester? What's the relative importance of measuring these exposures in those two periods?

DR. PERRY: So once again the ubiquitous answer is it depends. It depends on the outcome of interest because during the three trimesters the fetus goes through various stages of development. And so if we're interested in chromosomal abnormalities for example, we're almost interested in the preconception phase. If we're interested in neurodevelopmental outcomes, we're interested in a third trimester. So it is very much dependent on the outcome of interest.

And because we haven't been precise enough in capturing adequate comprehensive exposures, we cannot say measure this contaminate at week 20 or week 19 and therefore you will know exactly what predicts this congenital abnormality. We don't have that kind of precision.

DR. PANETH: Nigel Paneth. In the question you were given one says potential public health impact of the

outcome. And we haven't heard much discussion about public health impacts. So let me toss out a few thoughts on that in relation to environmental exposures in pregnancy.

Some of the conditions we're most concerned about with prenatal exposures are not very frequent so to take ones that are prevalent at less than one percent congenital heart defects, cerebral palsy, type 1 diabetes, the power in 100,000 according to the data we're given, even with 25 percent exposure, will pick up maybe barely an odds ratio of 2. So if you cut it down further to 40,000 exposures and you won't get measurements on everybody then I think the public health impact of what we could do in the Children's Study would be thus proportionately reduced. So if you would like to comment on the public health impact part, I would like to hear.

DR. SHELDON: Our comment is we all agree.

DR. MCCORMICK: I guess I would add one thing and that is that given some of these conditions it ought to be very explicit what ones you're going to be able to examine, what ones are just not possible. And so there may be conditions of high salients, my particular scars relate to autism and you just may not be able to address them. There are strengths and limitations to every study and I think if there's a limitation that should be very explicit upfront.

DR. DUAN: Naihua Duan from Columbia University.

I would like to follow on Melissa and Linda and also Sara's very thoughtful discussions about personal monitoring. I think that's definitely a very promising technology. I agree that it should be incorporated into the study as much as possible and I'd like to add a couple of statistical issues to it.

I think one is that having a stream of personal monitoring over a period of time would be the ideal way to address the variability question that Antonia raised earlier. And especially with some modern technologies that allow us continuous time, or nearly continuous time measurements, that would give us the variation both in short term and also over a period of time.

And another issue I think that's also important is to consider the validity of the measurements. Because other technologies otherwise are using environmental monitoring and residential monitoring, those miss part of exposure for some participants that might be important, like the occupational exposure and for some pollutants, like the exposure through the car exhaust during traffic might be important. And the personal monitoring in a sense is an automated device to sample the exposure across different activity patterns. So I would definitely like to second your thoughts to encourage using those technologies as much as possible.

DR. DEZIEL: I would just add I also support the use of these new technologies but before - but just because they're available and they're exciting, I think we'd have to really think carefully about how are we going to use them and what's the research question that they are going to address and make sure we adequately think that through. We can't measure everything. It's a balance between costs and feasibility.

DR. DUAN: Maybe if I can add a word. Those technologies are not all entirely new. So the EPA has conducted a variety of personal monitoring studies for several decades like the Total Human Exposure Study and the technology is advancing but there is a history of those older and new technologies being used and maybe we should try to learn from what has been done.

DR. PERRY: I completely agree that these technologies aren't new. What strikes me is that we're in an era of greater ubiquity and involvement in embracing these technologies. That's what surprises me, from crowd sourcing to voluntary uploading of your real time GPS data. It's as though it's not nearly as unheard of or unconventional now so that participation may be more broadly assured.

DR. SHELDON: I would like to make one other point and it's sort of contrary to what I said before. But

I'm all for personal monitoring and I'm all for being able to, especially for air pollutants and organics and those kinds of things.

The technology almost has to be cheap enough or easy enough to use so that you don't have to be going back and making multiple visits to the homes, to give them to people to pick them up. A visit to the home by a sampling team costs a lot and so one of the things that was very exciting to me was the idea that you have these band-aid type monitors which are self-powered and if you get that then it starts to become affordable. Or, if it's cheap enough, you just give it to them and they throw it away when you're done. But I do think that we have to think of the cost of deploying and un-deploying the instruments in such a large study.

DR. HENRY: Carol Henry from George Washington University. I think we haven't quite come to grips with priorities yet at least as I was reviewing what we've just discussed. There's been again a fair amount of input.

So I guess I'd pose the question Linda Sheldon presented and I think the results of a very thoughtful workshop from 3 years ago where in recognizing that the NCS cannot be all things to all people. They chose three health outcomes to look at and to better understand what the correlations for environmental exposures might be to

those health outcomes.

So the question I would pose to not only the panel but the group is how are we going to try and get to that same kind of endpoint. We are not going to be able to predict what we need in 15 years but the three outcomes asthma, neurological development and endocrine disruptors are certainly the buzzing topics of today. And if those three were then emphasized in the NCS and we started really collecting data, which I think we could start to do, it seems to me we would start getting some place. And maybe there's a different way to look at this but I think those three outcomes would really be critical to this so my question is why don't we do that?

DR. PERRY: Dr. Henry, excellent point. In fact, in looking at the current plan, ensuring that the prior workshop recommendations from June 2010, that's when these were published, were incorporated into the current sampling plan, I had to ask the question how well considered are we in being able to respond to those three outcomes. And my conclusion was - again based on the chronology including the prenatal with some discussion of attempts at preconception, and then the sampling plan over time for blood, urine, breast milk and cord blood and a spot blood for the newborn - that you would have the matrices necessary and be prepared to analyze the exposures of

interest as it pertained to those three specific outcomes.

I don't want to dismiss the question about the prevalence of outcomes and whether or not we're going to be adequately powered given rare events and I had to think through it for a moment. In fact, we may not have a critical mass, an adequate number of actual diseases, one disease to be able to study but what about precursors.

I study chromosomal abnormalities in sperm. These are potential precursors to congenital abnormalities if that sperm is successful in fertilization. There are a number of pre-disease indices from DNA adducts to chromosomal abnormalities that one could identify.

At the same time we've talked about genetic and epigenetic mechanistic studies that our blood in particular is going to afford, it would give us insight into mechanisms without having the critical mass of cases that you would require. So I would like to reinforce the point that we would have pre-disease opportunities for investigation as well.

DR. MCCORMICK: I guess as a developmentalist, I get a little queasy when people talk about neurodevelopment as a specific outcome because in fact you're talking about a fairly large number of relatively rare conditions, even when you talk about cerebral palsy, that's highly varied as well.

You're talking about a broad array of conditions and I suspect you would not have the power to look at the individual conditions and simply saying well is it an IQ test below two standard deviations or what, I don't see that as a terribly specific outcome for which to drive your analyses. I would be nervous.

DR. SONDIK: Ed Sondik from the National Center for Health Statistics. There was one comment that had to do with the geographic diversity in the sample, in other words, the 105 PSUs versus something else. I wonder if you could comment on the importance of that. In other words, what are the tradeoffs that are associated with that sample that's more clustered in terms of the diversity of experience, geographic experience in the country versus something like the original sample 105 PSUs which is quite diverse?

DR. CALAFAT: I think it depends, as Melissa has been saying several times, on the chemical that you're trying exposure that you're trying to assess. Certainly if you're trying to look maybe at some pesticide exposures and then agricultural pesticides then you would just want to make sure that you're covering agricultural areas versus some other areas in the US which are not.

At the same time for some chemicals and it depends also and we're what kind of moving around in a

loop, really what we're trying to assess and realizing that we probably cannot cover everything but if I go back to NHANES. NHANES is a survey that they just sample about 15 localities every year in the United States, yet they obtained like representative data for the whole US population. So it's going to be a tradeoff again thinking about how much and the cost that is going to be a large number of PSUs versus just reducing them and what are the outcomes or the exposures that you're looking for.

There are differences for example in the use of some particular chemicals that may be by demographics and it is going to be particularly - phthalates is one example that comes to mind. Again it depends on what you're trying to look for. If it is a chemical that is coming from exposure in a residential use that is very much driven by either socioeconomic status. And then you don't have enough localities to cover that particular one, you may be in trouble but if it's something that is more ubiquitous, exposure that is much more widespread then maybe the need to have this many sites is not that relevant. I hope I covered it.

DR. SHELDON: So I have got one other comment on this. For a long time - I agree with Nicole's comment about being able to do a good job sort of looking at the community level exposure. What are the various sources of

exposure out in the environment in terms of air pollution, water pollution, soil, those sorts of things?

I am not a statistician. I am not an epidemiologist. I don't know what the tradeoffs are between being able to do a good job of community site assessment versus the number of communities you would have to look at but I think it's a consideration that people should give.

DR. PERRY: I haven't thought about it statistically by any means but I can certainly say when it comes to environmental exposure assessment we can point to a wealth of information to show that environmental exposures are not uniformly distributed and that different subpopulations are affected and exposed in different ways with greater burden than others.

People living in public housing for example are more likely to be exposed to pesticides and a variety of other chemicals, fumigants and such - living in proximity to Superfund sites for example. So the importance of representative sampling when it comes to patterns of exposure is obvious.

DR. LITTLE: Rod Little, University of Michigan. This sort of relates to this point and that is one way of increasing power is potentially to increase the variability of the predictive variables that you're interested in. I

was intrigued that the current draft of the design does apparently hold over about 10,000 for something else, the something else being left dot dot dot as far as I can tell.

So I'm wondering if the panel thinks there is some promise there in focusing some of the 10,000 on areas where it might increase the variability of the predictive variable which might help with the power quite a bit.

DR. BRACKEN: Michael Bracken. In response to Dr. Sondik's question which I think must have been rhetorical because the answer is pretty obvious and Dr. Perry just gave it.

If you think of the environment exposures, they are certainly clustered. They are not uniform for the most part. And if you want to capture in this country as much of those exposures as possible then 105 counties will do it much better than a smaller number of counties. So I don't think there could be any doubt that you will look at more environmental exposures with a larger number of PSUs.

DR. PERRY: In the background information, I read about two useful ideas - missing by design and then the validation approach which is where I thought that 10,000 could come into play. That is let's say we wanted to study children in the Salinas Valley or parents of farm workers who were being excessively exposed to ambient pesticides and probably food residue pesticides as well. You could

take advantage of the fact that the exposure was much higher in that subpopulation and do a missing by design study.

Let's say you wanted to use two different mechanisms, one was more expensive and more invasive and perhaps more precise, and the other being easier to do simple spot urine sample lets say, cheaper but maybe not as reliable. That would be a perfect scenario for a validation study. So that's what I thought of for the 10,000. I also thought of the opportunity for the preconception opportunities given the challenges of recruitment.

DR. CALAFAT: And I just want to add that in some cases then we're pretty homogeneous if you want and then because there are a lot of exposures that we're having that are very similar. If depends on how you define exposure. If it's coming from like a use of personal care products that are also environmental chemicals but then they tend to be much more ubiquitous around the whole United States, just there may be some pockets of the population that they may use some products more than others but that you would capture it pretty much regardless.

Certainly, the more you can sample the better but if you're thinking about some other type of exposures that could be like suburban versus urban because we live very

much in those. And I think that a lot exposures that we're getting are from indoor environments. So there may be differences between populations like you have in the suburban area that is more likely spread out versus an urban population that you may be sharing a lot of exposures. Even though you're not using that particular product or that particular chemical but your landlord is. So I think that that could be an important distinction as well.

DR. MCCORMICK: Any more comments or questions for the panel, for each other?

Break 10:15

Resume 10:35

Agenda Item: Composition of Sample: Alternatives for Cohorts of Women

DR. MCLANAHAN: So the next session is going to be on the sampling design and Barbara Carlson from Mathematica Policy Research is going to be moderating that session and I'll let her introduce the people.

DR. CARLSON: Thank you. Welcome back. So this session title is about the Composition of the Sample: Alternatives for Cohorts of Women and I expect this will generate some interesting discussion.

As the NCS moves from a household based sample design to the one that's currently being proposed which is

to have a prenatal cohort and a birth cohort, possibly with the collection of subsequent births from each of those cohorts. So that would be the 90,000 and perhaps part of the 10,000 set aside for a convenient sample of pre-pregnancy women.

This session will talk about balancing the theoretical and practical issues in terms of the sample and the sample allocation. We did hear this morning about the environmental measures and that's obviously a factor in how the sample gets divided. There are also some initial findings from the pilot study informing some of the practical issues on the prenatal cohort especially. And they are just starting to work on the birth cohort.

So our four panelists will discuss the two questions for this session about the allocation of the sample between the two cohorts. One is how does one decide, how does one balance the two cohorts and then what should the allocation be? Should it be a 50/50 split, 80/20? Even the extremes of just one and not the other of the two cohorts.

And then what is the population that would be represented in each of those cohorts, especially the prenatal one, given that we're unlikely to enroll women in their first trimester or to enroll very many in their first trimester.

So the panelists have decided that these two questions are so intertwined that they are going to deal with them together rather than sequentially. And each speaker will talk for about 10 minutes. Then there's going to be a discussion amongst the panelists themselves and then we'll open it to the floor.

So our first speaker is Irv Garfinkel from Columbia University.

DR. GARFINKEL: I am going to read my comments. I've been working on this statement for a while so I'll read it even though even reading it it's not quite where it should be.

The National Children's Study is likely to be the most important birth cohort study in the United States in our lifetime. In retrospect it is not surprising that there has been a protracted struggle and very expensive pretesting over how to conduct it.

Different objectives and different disciplinary traditions need to be reconciled. In particular, it was not obvious how to conduct not just prenatal but even preconception exposure data from a population based probability sample of births at reasonable cost.

This is a very complicated problem. The recent evolution of the study design however has been very positive and we are now on the brink of reconciling the

conflicting objectives. Dr. Hirschfeld and his staff at NICHD are to be commended.

The last month for me has been an exhilarating learning experience. I feel like a young student, back in school but this time with an amazing set of teachers with vastly different views on some critical matters who are patient and kind even when I said stupid things. Hopefully the stupidity is gone. If not, my teachers are not to blame and I must say this morning's panel continued that. I learned so much and am grateful.

My prepared written comments include thanks to many of the people here including my fellow panelist Michael Bracken. Michael and I are on opposite ends about the appropriate balance between the prenatal birth cohort so it is really important to note an even more fundamental area of agreement shared by NICHD, as well as by Michael and I, that probability sampling is essential to the quality of the NCS.

NICHD now proposes to enroll 45,000 mothers and children at birth from probabilistic samples of hospitals and prenatal providers. The 50/50 split between hospitals and prenatal clinics is a huge step forward from sampling only from prenatal clinics in terms of cost and scientific value.

However, neither a prenatal nor a birth cohort

will produce excellent prenatal data. Only sibling data can produce that at reasonable cost. Close to 100 percent birth cohort that enrolled subsequent sibling births would save even more costs than the current 50/50 split and would immeasurably increase the scientific value of the study when it is completed 21 years from now.

Though neither of us is a biologist, Michael and I also agree that gathering prenatal data must be a critical component of the NCS. My understanding which has been reinforced this morning is that first trimester prenatal data is the most valuable part of the prenatal data for most of the questions we're interested in and that for some questions preconception data may be equally important. The prenatal data produced by the prenatal sample fails miserably on these grounds.

And I'm going to highlight what data we need by offering a design to get you to think about and I'll conclude with what data we need. So consider a probability design in which hospitals within each PSU, and I'm not going to talk about whether that's 50 or 100, and births within each hospital are sampled with a known probability and somewhere between, actually let's just say 60,000 births are enrolled in hospitals at birth. And the data on placentas and cord blood are collected at birth and from what I've learned we also later collect breast milk.

All subsequent sibling births over the course of the 21 years of the study would be enrolled in the study. This design provides nearly as large a sample of children with prenatal data and if first births were over sampled in the birth cohort, the sample of siblings with prenatal data would be as large as a pure prenatal sample. And I'll explain why in a minute. And the sibling prenatal data is dramatically superior to the prenatal data provided by the prenatal sample because the sibling data comes earlier in the first trimester and includes data on preconception conditions as well as data on a previous birth.

Consider the extreme case that if observations without prenatal data were worthless - I don't believe that but I think some people do - so consider that case. Observations with no prenatal data are worthless, a birth cohort in which data on the first birth were thrown away and data only on sibling births were followed, the birth cohort sibling data would produce far superior prenatal data at lower cost than a prenatal cohort.

If the fundamental biology of harm from environmental exposures is the same for first and subsequent births and we have, from what I've heard this morning no evidence that its not, and subsequent births and early prenatal data and preconception periods are critical, the design that I'm suggesting is nearly optimal.

The virtue of the assumption that the fundamental biology of exposures is the same is that I suspect it is right and even if not true, it enormously simplifies what is otherwise an extremely complex sampling problem.

A third virtue is that it points to the importance of finding out what we already know about the assumption or hypothesis and what we need to know. The vice is that it may not be true. It's possible the biology of exposures is different and the birth cohort provides no data, no prenatal data on first births and therefore cannot test this assumption. That justifies at least a very small prenatal sample of first births.

Now let me emphasize, even if we knew with certainty which we don't, that the biology of first births was different from the biology of subsequent births. With respect to exposures, a 100 percent prenatal cohort would be superior to a mix between a hospital and prenatal cohort only, only if there was no value to very early pregnancy data, no value to pre-pregnancy data, no value to having a much larger cohort of children to follow with half the sample having no prenatal data.

I submit that there's not a person in this room who believes any of these three. Consequently, because a prenatal sample costs multiples of a birth sample, the optimal prenatal sample will be much closer to zero than

100 percent.

Enrolling sibling births from a birth cohort has enormous virtues. Most important, it is the most cost efficient method of sampling births during preconception and very early pregnancy. Within 3 years of all births, nearly 30 percent of mothers have a subsequent birth. Within 5 years, the figure is about 44 percent.

Within 21 years, the overwhelming maturity of mothers will have completed their childbearing and assuming that completed fertility is about two children, a birth cohort of first births will have sibling births with preconception and prenatal data on the same number of births as 100 percent prenatal cohort.

A random sample of hospital births would yield nearly as large a sample. Each observation generated by the sibling sample will be superior to the prenatal sampled observation because it will contain data not only on preconception but also earlier prenatal data and data on mother’s previous births including placentas and cord blood.

This information will be invaluable for imputing missing exposure for first birth prenatal period. This is so counterintuitive what I’m suggesting and so central to the optimizing problem that it is worth repeating. So long as the biology of exposure is the same, the best data, not

just on preconception but also the prenatal period as a whole, comes from siblings not from births sampled prenatally.

Two other virtues of the sibling sample are worth noting. First, although sibling based estimates are less precise than corresponding non-sibling estimates, siblings allow researchers to control for or rule out confounding from genetic and environmental circumstances shared by siblings.

Second, collecting sibling data is cheaper from start to finish than collecting data from two children from different mothers and different household circumstances. Each birth enrolled in a prenatal sample cohort is de nova. Each sibling enrolled from a birth cohort is enrolled from a mother who has already been recruited and is a loyal member of the study.

A birth cohort is superior to prenatal cohort in terms of cost and sample size and this is so for two reasons. First enrollment costs of a birth cohort will be smaller than enrollment costs of a prenatal cohort because of economies of scale.

Second and far more important, the prenatal collected from the first births enrolled in a prenatal cohort is very expensive. NICHD estimates that the cost per child of enrolling a prenatal birth and collecting

prenatal birth data is at least three to four times and may be as much as 10 times the cost of enrolling a child at birth.

To simplify and illustrate how these ratios could be so big, assume that it costs $1000 to enroll a mother in either a prenatal sample or birth hospital sample and another $5000 to collect prenatal data for the mother. The ratio of total cost would be 6 to 1. If enrollment costs were $2000 and prenatal data collection costs $18,000, the ratio would be 10 to 1. In other words, for every child enrolled in a prenatal cohort for the same cost, 3 to 10 children can be enrolled in a birth cohort. Thus more sibling and first births can be enrolled in a birth cohort than a prenatal cohort.

It’s true that for every sibling birth enrolled the costs of the prenatal data collection will be incurred eventually but total costs are lower for four reasons. First, enrollment cost of already loyal members of a longitudinal study must be lower than enrollment costs of de nova prenatal mothers. Second, the costs of collecting data on family circumstances are lower for siblings. Third, the siblings will be followed for a shorter period of time. Finally, the enrollment and data costs of siblings comes later than the enrollment costs for a prenatal sample which means they are lower because the

later the costs is incurred the more it will be discounted.

The only parts missing from a birth cohort described are the prenatal and preconception data on first births. The prenatal missing part can be efficiently filled in by a relatively small sample of first time pregnant mothers drawn from prenatal proprietors. My guess is that 10,000 would suffice and might indeed be too high. Every additional birth of first time pregnant mothers drawn from the prenatal providers reduces the number of sibling births that can be enrolled in the study by between 4 and 10 children. This is a really steep price to pay for inferior data.

I'm just going to conclude with the questions. The analysis above identifies the key scientific questions underlying the choice between the size of the prenatal and birth cohorts. How important is early prenatal data? How important is preconception data? Is the fundamental biology of harm different for environmental exposures the same for first and subsequent births?

The key operational questions all relate to cost. Is the ratio of the cost of enrolling the prenatal sample as opposed to the birth sample 3 to 1 or 10 to 1? How costly will it be to collect placenta and cords on the first births?

Finally, time is important. The birth cohort

sibling design requires prenatal data in later years than would a prenatal cohort. Once these issues are clarified, a formal sampling design analysis of the kind described in a minute by Naihua will get you a precise optimal allocation.

I will end my comments where I began. The National Children's Study is likely to be the most important birth cohort study conducted in our lifetime. We need to keep our eyes on the prize and make sure that at the end of 21 years of study, subject to budget constraints, the data are as good as it can be. Many of us will be retired and some of us will be dead by then. But the scientific value of the study will be so great and the cost of following the sample so low as compared to the costs we're talking about for the first 21 years, the NCS unlike all of us, will live on. Thank you.

DR. CARLSON: Thanks Irv. Our next speaker is Naihua Duan from Columbia University.

DR. DUAN: Thank you. I would like to second Irv's compliment to the study. What has already been accomplished for a very, very complex and challenging study and I think we all look forward to the day when the main study gets underway, some time hopefully in the near future.

I would also like to second Irv's suggestion to

maybe incorporate sibling cohort into the study as a primary recruitment methodology to address the challenge in recruiting an early pregnancy or preconception cohort.

And I would also like to thank the first panel for laying out the groundwork for all panels because we're really looking to you to prioritize and let us know what are the important time periods and for us to help respond to the questions about sample design.

So, I will share some thoughts on the statistical issues involved. The first thing I would like to concur with what the Vanguard investigators recommended. The design for a good sample study should be made according to the study objectives. And this is a large study and it really has the potential to go beyond being a descriptive study and specific hypotheses will help us derive how the design decisions are made.

And I think that goes with Rod Little's comment that to the extent we know the study should try to maximize or optimize the dispersion of the exposure to get highly exposed and maybe also to get lowly exposed. Those are the informative cases for us to understand the exposure effect and the bulk of the majority of the sample population in the middle region I guess doesn't really give us as much information.

Now I guess for a large complex and multifaceted

study like the NCS -- note that I put the objectives with an ”s”-- we need to balance across the multiple study objectives. I would like to share with you some thoughts about the possibility of using the statistical methodology of optimum design to try to address that question. And I would appreciate from my statistician colleagues your thoughts and comments.

So the optimal design literature mainly started with Jack Kiefer’s 1959 paper and it has evolved into a major literature in statistical methodology, mainly in experimental design. The main idea is to use the methodology I guess as a big part of the idea, to balance across the multiple study objectives.

And the application is somewhat limited to sample design and personally I have used this approach a couple of studies I worked on over the years. One of them is the HCSUS - HIV Cause and Service Utilization Study that was sponsored by what used to be the HCPR with Marty Shapiro, Marty Franco, Paul Cleary and Sam Bozzette. It was a very challenging study recruiting HIV-positive patients through the providers, somewhat similar to the way the prenatal sample is being discussed here. And we are very excited that I think the study was known as the first ever national probability sample of HIV positive patients in care.

Another study I applied this methodology to

myself was the National Latino and Asian American Study. I do believe there is a potential for wider applications of this methodology in sampling applications and I will try to convince you of that. So I will give you a couple of examples of how this methodology potentially could be applied.

So first, making up a naïve simplistic model for what the study might want to accomplish. I'm sure that the investigators in the study can make this more realistic and more refined. But I just want a very simple model to illustrate the idea.

So Y stands for the outcome, say the cognitive function for children at age 5 and E1, E2, E3 stands for exposure at different times periods, like the preconception or the first trimester and E2 might stand for the third trimester and E3 might stand for postnatal and then regression coefficients, b0, b1, b2, b3 are the study objectives we try to assess.

So we like to try to get the best estimates for those parameters as possible. I want to make a technical note that I'm making an assumption that the exposure measurements are standardized to have mean 0 and standard deviation 1 so those are standardized effects and that also gives the intercept parameter, b0, the interpretation as the population mean.

So also exposure is very often scaled in distribution so we might think about those variables as the lack of the exposure measure. And then I think the background document listed 5 candidate designs being considered. So I call them D1, D2, up to D5, and so each Dk is a possible design.

So one of them might be to take 40,000 from birth cohort, 40,000 from prenatal cohort and 10,000 from the sibling cohort and maybe 10,000 from the hotspots. And then another design could allocate differently and it's conceivable we could go beyond the five designs currently on the table and consider a range of possible designs.

So the exercise in optimal design is to specify a performance criteria for the designs and then go through the exercise, somewhat similar to doing power calculations, to calculate those criteria for each design and then try to figure out for another design we're looking at which one has the best performance according to the criteria we specify. So the key is to specify a common criteria across multiple study objectives.

So one simple criteria in the optimal design literature usually called the "A Optimality Criteria" is the sum of the variances of those parameters for each specific design. And this might not be a very good criteria because it does not take into account the relative

importance of the different study objectives.

So the next possible design is to do a weighted version so the only difference here is I added those weight terms. So each W stands for the importance or weight, the investigators want to attach to each study objective. And so now we go through the exercise to try to calculate the weighted A optimality criteria and try to find the design that achieves the best performance.

So this does not answer the question. But I'm offering to transform the difficult question about how to allocate a sample to a question that might be more tangible for the investigators to think about -- how important it is that we want to reduce the uncertainty in each of the parameters. Then we can debate whether we want to put a ratio of 5 to 1 or 3 to 1 to the earliest exposure effect to the postnatal exposure but I hope this methodology makes it more tangible to try to carryout this exercise.

Another possible extension is to use this total survey error methodology so instead of considering the variances that are usually used in the optimal design literature. I think it's conceivable we could use a performance measure like the mean square error which is the variance plus the square of the bias to incorporate both the sampling error and also the non-sampling error.

I think this might be a methodology to

incorporate the non-probability sample strategies into this framework. If there is a possible serious consideration for non-probability sampling methodology for part of the study then if we can specify how the variance and the bias trade off then maybe that could be incorporated. I have never done this myself but if this sounds like a useful approach, I would definitely recommend the study to -- I'm speaking in self-interest for my profession -- to recruit a good, young, energetic statistician who is familiar with the issues involved to work on those issues that potentially might help save big bucks down the road.

So if I come back for a few minutes to talk briefly about a couple other topics. The next question about the cost, I would like to argue that we should not just look at the recruitment costs. The study does not stop with recruitment but there's a lifetime stream of cost.

So overall the ultimate product is what the study overall has accomplished and what did it cost. So I would recommend that we take the follow-up costs into consideration in choosing exposure sampling strategies and the future cost is counted in today's dollars and we have good economists who know how to do that.

Another quick point is the multi-cohort study uses the multi-frame sampling strategy. I might be

preaching to the choir but I want to make a note that the multi-frame sampling strategy does not require each cohort to be representative of the entire population. All that is required is that the cohorts combined represent the entire population. So it is possible to try to take a cohort of say prenatal sample that might or might not covered the entire population, or like Irv suggested, the sibling cohort would not cover the entire population but it would give good coverage for an important part of a population but then the rest of the population might be covered otherwise.

Another point I'd like to suggest is to integrate the special cohorts into the overall design deliberation. This 10,000 special cohort may be from the hotspots. Potentially it will be analyzed together with a main cohort comparing hotspot high exposed to the general population. And I think it would be advantageous to take the entire sampling strategy under one roof to make the best decision. Maybe we want 20,000 instead of 10,000.

Maybe there are other things that could be done. Maybe we want to go to warm spots also to recruit our samples, maybe in some geographical areas. I think Melissa commented earlier, maybe the multifamily housing might be highly exposed to pesticides.

So I hope that this integration might allow us to

maybe be more flexible in thinking about the sampling strategies. I think the response rate is a small part of it. I will write up the comments Nancy and share that with you later. Thanks very much.

DR. CARLSON: Thanks, Naihua. Our next speaker is Nancy Reichman from the Robert Wood Johnson Medical School in Princeton University.

DR. REICHMAN: Thank you, Sara, Irv and everyone else for asking me to be here today. So I'm an economist and I've been studying maternal and child health, in particular socioeconomics, determinants and consequences of poor child health for decades. I was Sara McLanahan's first research associate at her then new Center for Research on Child Wellbeing at Princeton back in 1997.

I quickly got sucked into a brand new birth cohort study that Sara and Irv were cooking up called Fragile Families in which new parents were to be interviewed in two cities, two cities, okay. Irv asked me if I could help them gain access to hospitals to conduct the study in Oakland, California. Well, I'm always up for a new challenge so I said sure, I'll try it. So I tried and it wasn't easy and somehow I pulled it off. Well, no good deed goes unpunished, they started adding cities. And they couldn't be stopped.

So the first two cities quickly became seven

cities and eventually grew to 20 cities and then I think they had the sense to stop. So I guess I'm here because I secured hospital access, with a good team obviously, to conduct the Fragile Families survey as well as medical record data collection of abstractions of charts at 75 hospitals in 20 cities across the United States with a success rate of over 90 percent of sampled hospitals.

So I have to tell you that the process probably took years off of my life. But - look at the data that came out of it. So in my 10 minutes or so today I'll cover two things. I'll briefly discuss whether I think it will be harder or easier to get hospital access now 15 plus years after Fragile Families for the NCS as it sounds like sampling from hospitals will definitely be at least part of the picture.

I will also raise a few general issues and questions related to costs and benefits of sampling from prenatal care providers versus hospitals that I think are very important in order to decide the allocation across the cohorts. Just one housekeeping issue - this is going to sound weird. But to be parsimonious with words, I'm going to sometimes use the term placental material when I mean placental material and cord blood.

So in terms of the first issue - whether I think it will be harder or easier to get into hospitals 15 years

later? I don't think it would be harder to get hospital access today because during Fragile Families the Institutional Review Board (IRB) landscape was constantly shifting under our feet.

When we started, IRBs were in their primitive stages at many places other than big research institutions as was NIH in terms of its own policies vis-a-vis human subjects protection. In fact, some of the biggest problems we ended up having were in hospitals that initially had the easiest application procedures and retroactively decided that the approved study was not acceptable. That was tough.

The issue of sand shifting under our feet should be much less of a problem today because procedures have universally become much more formalized. However, substantial resources, it's not cheap, and a well chosen team would be needed to get through the necessary processes and of course, and this is an important issue, for Fragile Families the issue of whether it would be logistically possible to collect placental material when mothers are consented after they give birth was not at play.

Irv, after consulting with docs and hospital administrator thinks it would be possible to collect placental material when sampling is done in hospitals and given the conversations and the emails I also think it's

possible. First there are zero risks to the mother from collecting the needed materials. Second, it's apparently not at all unusual these days for mothers themselves to take the initiative to have their placentas preserved and banked. We've been told by hospital administrators, research deans, different people, that if fairly compensated and the burden to the hospital minimized, the hospitals would likely agree to a system in which all of the placentas that might be needed for the study are preserved and stored, at the study's expense of course, with those of non-consenting mothers later destroyed.

And perhaps, and this is my own crazy idea, some of the placental material could be stored as a perk to consenting mothers and made available to them should they need it, for example if stem cells are needed to fight diseases in her child or other family members down the road and that's just an idea.

In terms of the how-tos of getting into hospitals, I'd be happy to share lessons I've learned along the way but I'd rather use my 10 minutes, or whatever I have left, to get into some of the bigger issues on the table today.

The two key issues I'll raise now are questions I think need to be answered in order to assess the cost and benefits of sampling from prenatal providers versus

hospitals. I echo Irv in saying we can't really answer the question number two what the optimal allocation should be, although Irv has an answer, until we get a clear accounting of projected costs and benefits of each type of sampling.

So first, my first question is what type, and this is just my ignorance maybe but I'm like a practical person, I need answers to these questions. First, what does administrative approval consist of at prenatal care providers? If the provider is in the hospital, such as a hospital clinic, I assume that the hospital IRB approval is what's needed.

Otherwise, what outside institutional approvals, if any, are required at private practices or other types of prenatal sites? How many different types of provider sites are there and on average how many mothers would be recruited per site? How costly will it be to maintain quality control and standardization of protocols across what I sense I would be a large number of small sites?

In particular, I worry a lot about keeping track of case dispositions and response rates, particularly the denominators which sounds like a logistical nightmare to me. We had 75 hospitals in Fragile Families and that's starting to sound like a small number after today.

So it's hard to compare the cost of securing institutional access and running the study without having

better information on these things.

My second question, and this might be really off the wall and maybe I'm misunderstanding what the prenatal sample was going to do but - my second question is if sampling is done at prenatal providers, and I'm not averse to that, is an encounter with the mother in the birth hospital actually necessary. I believe that was part of the plan, like you sample at prenatal providers and then you also get the mother after she's given birth, you go to the hospital. Is that necessary?

If sampling is done at hospitals, a hospital encounter would obviously be needed. If, on the other hand, sampling is done at prenatal care providers, would it be possible for mothers to request placental material, and don't laugh at me, placental material and medical records from the delivery hospital for purposes of the NCS? People request medical records all the time, you pay, and people request placentas all the time.

If this could work logistically - that is mothers could request through the study their placental material and medical records from the hospital - this could potentially eliminate the whole hospital encounter piece, this just came to me yesterday, for births that are sampled at prenatal providers unless there are other reasons to have a hospital encounter that I'm missing.

If a hospital encounter would still be needed for some reason, it seems to me, and I agree with Irv, that sampling from prenatal care providers would be enormously expensive compared to sampling from hospitals and it might be worth it, I don't know. Because of the added cost of the prenatal data collection encounters, the need to access both prenatal care providers and hospitals and the logistics of coordinating the study across so many sites.

So if a hospital encounter is needed when births are sampled from prenatal providers, the benefits of having prenatal data for the first birth, what you can get based on the panel earlier, so the benefits of having prenatal data for the first birth above and beyond what can be collected from the placenta. And I have no idea what the answer is to that question, I'm ignorant. What can you learn from the placenta and cord blood about prenatal exposures? And the things we talked about in the first panel.

I'll say this again. If a hospital encounter is needed when births are sampled from prenatal providers, the benefits of having prenatal data for the first birth that you could get above and beyond what can be collected from the placenta would need to outweigh the much higher cost of sampling in that setting. I just don't know enough to make this calculation, would love a clear rendition of how

important having prenatal information - this has come up over and over today - for the first birth really is in the scheme of things considering all aspects of the study, when placental material is available from the delivery and good pre-conceptional and prenatal data could be collected for the subsequent child.

The bottom line is, from what I know so far, hospital sampling seems to be much, much, much cheaper. Unless the hospital step could be skipped when sampling from prenatal providers but I still don't have enough information to know how expensive the prenatal provider sampling would be.

Key pieces of information including the following are missing, how expensive would it be to get placental material when sampling from hospitals? Could hospital encounters be skipped entirely when sampling from prenatal care providers? What are the projected costs of access and management under prenatal are versus hospital sampling?

Once we have the relevant information, good cost projections are needed for both approaches and the unique benefits of each must be factored into a full cost benefit calculation that I'm glad I don't I have to do, that by the way should include the sibling cohort. I want to underscore that we're really trying to get the optimal mix across three cohorts here.

So now I just want to take 15 seconds to make an entirely different point because I have an audience here. While the NCS will be truly pioneering by collecting prenatal and pre-conception data in addition to birth data and beyond, it seems short sighted to stop there. My colleague Julien Teitler and I think it would be possible in about one minute of survey time to collect invaluable information on the health of the prior generation, that is the mother's parents, information on death of a parent, the parent's name, the parent's date of birth and the year of death could be collected to obtain age, cause of death, education and other variables, from death certificates.

Brief information about the parent's lifetime smoking and drinking could also be collected from mothers providing additional valuable health information. As animal and human studies increasingly demonstrate, determinates of health can originate well before the parent's generation. Thanks for your attention. I look forward to an interesting discussion.

DR. CARLSON: Thanks Nancy. And our fourth panelist is Michael Bracken from Yale University.

DR. BRACKEN: Thank you. So, Irv set this up as a little debate and as we've gone down the table it's becoming increasingly more extreme. And I will be extreme in a different direction. Irv, as they say in Washington,

when I wasn't epidemiologist I was a biologist actually.

So I would say why should the majority of the NCS actually be recruited in pregnancy. I'm talking about 85 to 90 percent in pregnancy. I'll talk about the remainder in a few minutes.

Fetal origins of disease, fetal, is one of the dominant series in studying both childhood and adult disease. And it needs very large pregnancy cohorts. It's actually almost 40 years ago to the day when HRPS published a crucial paper showing how fetuses exposed to diethylstilbestrol when they grew up -- girls in their 20s and 30s -- developed vaginal adenocarcinoma. We know antibiotic use in pregnancy increases risk for asthma. We're concerned about mental birth defects and physical defects, 5 percent of all pregnancies. They need pregnancy cohorts to study.

We know that people born at low birth weight actually have higher risks of adult cardiovascular mortality. They also seem to have lower risks of cancer mortality. We don't know why. To follow up and find out why, we need pregnancy cohorts.

The influence of many drugs used by millions of women in pregnancy on physical and mental disabilities in their children may remain uncertain. The public health concerns, what's the effect of antidepressants,

antiepileptics, antiemetics, pesticides. The number of concerns is legion. We can only understand that by studying pregnancy cohorts.

We do know that the origins of autism, cerebral palsy, ADHD, and many other so-called perinatal conditions actually have origins earlier in pregnancy but we don't understand what they are. All of this research needs pregnancy cohorts.

We also know that exposure data is actually poorly recalled, even when asked at birth when you're asking about what happened in pregnancy. Infant mortality in the United States ranks 34th in world, it was 12th in 1960, it was 23rd in 1990. But the causes of it won't be found in birth cohorts. They are due to associations in pregnancy, including disparities in prenatal care. They have to be studied respectively in pregnancy.

Twenty-five years ago someone said we know that the vicissitudes in our own uterine existence may profoundly influence the rest of our lives, both physically and behaviorally. Actually, it was me. But this is not new science. It's not new information.

Another reason to study pregnancy, pregnancy itself merits study. Miscarriages occur in about 15 percent of clinically recognized pregnancies. Fetal death and stillbirth are all outcomes of great concern.

Obviously, they need pregnancy cohorts.

Many birth cohorts are being completed around the world but they're birth cohorts, they're not pregnancy cohorts and this is where the NCS could make a real contribution. The proposed mixed cohort, the layering sample, is far too cumbersome and is unnecessary and it misses the real scientific goals.

Is a pregnancy cohort more costly? Actually, there is no evidence that it is. It's no more costly to recruit in provider practices. The table in the document that's being supplied, showed evidence from 16 cohorts, three were mine, most of which collected biospecimens for an average cost including indirect costs of $2,000. Even with inflation, this couldn't possibly exceed more than $5000 and it's still two orders of magnitude less than what the NCS Vanguard cost.

Within a primary sampling unit, you need a list of providers and the hospitals in which their patients deliver. These form clusters. And then these clusters are sampled to form a probability sample. It's not that complicated. There is no cost to the sampling process itself and recruiting sample providers in hospitals should be no more costly than recruiting them when based on convenience. And sampling fractions and denominators can be obtained from birth certificate data where all this is

recorded.

Saying a blood sample collected prenatally is the same cost as saying a blood sample at birth except it's got much more information. There may be more additional costs in prenatal exposure assessment versus estimating prenatal exposures at birth but these costs are not related to recruitment. They are costs related to sample collection. And they are costs worth bearing because they relate to collecting more valid data.

Now how early can gestations be studied in a pregnancy cohort? Well, we've got four Yale cohorts, a total of almost 17,000 pregnancies and in one where we restricted gestational age to week 16, we have 30 percent at 8 weeks. We have 91 percent by 12 weeks, thus the end of the first trimester. In another cohort where we restricted to 22 weeks gestation, it was almost the same at 8 weeks, 29 percent, and 76 percent at 12 weeks, and same for the others.

So from a cohort of 100,000 pregnancies, you could have 30,000 women assigned for interview, which is what we use, by 8 weeks based on LMP and 75,000 to 90,000 by 12 weeks. So collecting first trimester exposures in pregnancy cohorts again is well documented and it's not that complicated.

Respective prenatal information on first births,

not just the subsequent births in women already enrolled, is crucial. First pregnancies are biologically different from subsequent pregnancies. Preeclampsia is a first pregnancy disease. Fetal growth restriction is more severe in first pregnancies. And as these children are followed up through childhood, we get into issues of birth order. Are we only going to be studying children who already have a sibling? That would be at tremendous detriment to this project.

Now biological exposures may not differ between first and subsequent pregnancies but the scientific interest is actually in the interaction between these exposures and the fetus. And the fetus is changing from one pregnancy to another. So there are biological effects and we're studying these already in gene environment studies and in epigenetics.

Now certainly these are areas where there is more hypothesis than fact but it would be a scandal if NCS could not study these questions because assumptions have been made, which you've heard about, at the sample design phase. An assumption free strategy places fewer constraints on the way pregnant women are sampled so they are representative of all pregnancies in the United States.

Now the preconception cohort, and this is a 10 to 15 percent, is a particularly interesting group. Many

hypotheses concern exposures at the time of conception or before. It's also a really difficult cohort to recruit. Women in fertility clinics who may know and plan the exact date of conception are highly selected exactly by virtue of their infertility and preconception probability samples are almost impossible to obtain and likely not worth the effort.

So I think here with some common ground, and in my view the sibling cohort should be used for preconception studies. They are using women who are already recruited to the NCS and based on an original probability sample of recruiting women already under surveillance and women who've shown a willingness to participate in research.

It is a disadvantage of selecting only preconceptions, prior to a subsequent pregnancy, after a prior pregnancy. Nonetheless, these will be preconceptions of choice I think for the preconception cohort.

The birth cohort I see absolutely no advantage in to recruiting at birth. It misses the unique opportunity offered by the NCS to study the most important scientific questions of our and future generations. The only worthwhile time to recruit in the hospital is for women who receive no prenatal care and this is an important group of women, often at high risk for poor health and problems in children rearing who should be recruited to the NCS. And

they will need special recruitment strategies.

Is recruitment easier in the prenatal clinic versus at the hospital? Well, in our experience, absolutely it is but you've got to consider both the provider themselves and the hospitals and the subjects. In terms of providers, they are easy to recruit. They don't have IRBs. I think Nancy asked the question what do you do? Well, providers don't have an IRB or your own investigator IRB covers it.

They are also actually more homogeneous in hospitals in the way they deliver care. Hospitals of course of IRBs and they vary considerable in their institutional practices and in what goes on in the delivery room. We've had more hospitals refuse to join research than providers. In fact, we've had no providers outside of our NCS experience refuse to join the study.

And refusal by a hospital to participate, and some do refuse, eliminates many more women from a sample than does refusal of a private practice, because you're knocking out a whole bunch of practices. In terms of the subjects, consent is more readily obtained prenatally. It can't be obtained when a woman is in labor, that's unethical. And women themselves are not very conducive to being asked to consent to research when they are in labor.

In hospitals, after a mom has delivered, she may

indisposed because of the delivery or the child may be indisposed. The child may be in the NICU or it might be going up for adoption. And 12-hour discharges, which are very common now in hospitals, will mean missing quite a large number of women. Obtaining consent after labor to get cord bloods and placentas, you can get consent after labor but the cord blood and placenta have probably disappeared. They will be lost. And the presence of families and the excitement of post-birth are other barriers to obtaining consent after labor.

In contrast, when you're recruiting in a prenatal practice, the medical records of the study subject are flagged when they go into the hospital so that in the delivery room they know this is a placenta we need to keep and this is cord blood we need to keep and it almost always works.

So in conclusion, the most sophisticated sampling design will fail utterly unless the practical details of how obstetrical care is delivered in this country are taken into account, both in the providers and in the hospitals. So thank you for the opportunity to share some thoughts about this.

DR. CARLSON: Thank you Michael. So we heard four different opinions here on the allocation. We have Irv who is arguing for primarily a birth cohort with

subsequent births. We have Michael arguing for primarily a prenatal cohort. We have Naihua who proposed a quantitative approach to optimal allocation and we have Nancy who provided a lot of practical issues, some questions to consider and her experience with Fragile Families.

So first I'd ask the four panelists to continue the discussion amongst themselves before we open it up to the floor.

DR. GARFINKEL: So I think it's first important to identify where we agree and then we can figure out where we disagree and what data can be brought to bear on where we disagree.

So I agree, I think we both agree that we need to get prenatal data, that that is very important. I'm not sure how - I guess Michael does agree actually that first trimester prenatal data is really important. And there I think we might do some comparisons as to which method gets you more first trimester prenatal data.

And I think when all is said and done, when you consider the costs, the upfront costs, you will see that you can get as much, actually I believe more, greater number of people with prenatal, early prenatal data from the birth cohort and I'd like to understand why you don't think that's true because I can't follow that.

DR. BRACKEN: Well, it's not just a matter of cost. It's a matter of what are the scientific questions. I mean if we're just doing this on cost, you wouldn't collect any biological specimens -- that would be really cheap.

DR. GARFINKEL: You get as much. I'm saying if you can get as much, you get more early prenatal data from siblings from the birth cohort than you can get from 100 percent prenatal sample because collecting that prenatal data on the first birth is so expensive. I believe it's also more expensive to enroll the women in the study but I've assumed, let's assume it's no more expensive, you could do enrollment equal cost, you're still collecting all that data prenatally for your first birth. When you sample from the hospital, you don't do that. You save all that money upfront.

DR. BRACKEN: Well, first I don't think you're saving that much money. In fact, I don't think you may be saving any money. But secondly, getting prenatal information on first births is important. If, at the end of this study, no information is available on pregnancies to women delivering for the first time - that would be an absolute scandal in terms of trying to answer important scientific questions. But that's what would happen in Irv's model.

DR. GARFINKEL: That is not what I heard this morning, let me just say. The panel this morning, when they were asked, I asked the question do we have any evidence that the biology of exposures differs by first and subsequent births. The biology of exposures, not biology in general, the biology of exposures and the panel said no.

DR. BRACKEN: But the biology of exposures is only half of the question. The question is how do these exposures interact with a developing fetus. And we cannot assume that the developing first pregnancy fetus is identical to subsequent developing fetuses. In fact, we know from many examples why there's a reason to believe that would not be the case.

DR. REICHMAN: Irv, there is one argument for getting - that I can think of, another argument for getting the prenatal data on the first pregnancy. The collection is structured as part of the mother's prenatal care so she's going for visits regularly and the different collections can be made but if it's for the subsequent sibling there's no connection to the provider so it might be a little more complicated to do the collections. I don't know. But it's just something that should be factored into the whole big mess.

DR. DUAN: If I could a couple words. I'm very good at being in the middle between two points of view as a

statistician. I think probably from, I guess from priorities discussed in the first session, we all agree we want to try to get the exposure data as early as possible. And there are different ways to do that.

And I would think that the prenatal sample and the sibling sample are not mutually exclusive. The question is really how to combine them and make the best use of them. If we go through some kind of formal calculation, maybe the type I suggested or maybe some alternative methods that my esteemed colleagues might suggest, you know we might come up with the conclusion that a large share should go to the birth cohort or a large share should go to the prenatal care. And I think that is a question I think we should really work out.

I do think, Irv, this missing first births is a pretty important question that we should try to address and the prenatal cohort might be as close as we can to get them unless there's a practical way to get a decent preconception cohort. So we might need to combine the strategies instead of trying to go exclusive one way or exclusive another.

DR. BRACKEN: Can I just respond to that thought -- to what you just said -- because you can create very sophisticated subsamples and this, that and the other, but please have some sympathy for the people who are trying to

run these studies. When you've got women in different subsamples and you're trying to manage what schedules they are on it becomes awfully complicated and the opportunities for making mistakes actually increase. Especially a study that's already really complicated.

I think a straightforward sampling strategy where you just decide you're going to get everybody, pre-pregnant would really simplify the whole thing and remove a lot of error that would occur in the field in trying to implement these really sophisticated subgroup study designs.

DR. DUAN: Yes, Michael, I fully appreciate that. As a practicing statistician I've worked all my career with challenging interesting studies with logistics and the HCSUS study I talked about earlier in many ways was a very challenging study. But those are things that are challenging but it is not something that is impossible to address, especially nowadays with advances in information technology what we talked about in the first session. Also what can be done in managing the field work, et cetera. So this is such a large study and I think we should not shy away from having some complexity in the design.

I want to add a note to the question that Nancy posed earlier about keeping track of response rates in this provider based sampling context and that's an issue that in the HCSUS study I worked on previously we had to deal with

and we have a couple publications. I'll be sending citations to Nancy.

One thing I want to note is that in a provider sample, the appropriate design needs to take into account both the response rate at the provider level and at the individual patient level, because when a provider refuses to participate all of its patients are automatically non-respondent to the study and this needs to be taken into consideration.

Another thing that I want to make a note of is that in some discussions about response rates sometimes there is a mix up between what statisticians call response rate and what is sometimes called a cooperation rate. So sometimes the response rate is calculated as the proportion of the candidates who agree after being contacted. That is usually called a cooperation rate in the statistical literature, like the AAPOR definitions. The standard definition that is usually used does take into account the candidates that we want approach but we were not able to approach and in our field of studies that is often an important component.

DR. CARLSON: I did have an opportunity to listen to one of the weekly calls earlier this week and I heard that for the prenatal cohort they are giving the practices four choices, whatever works best for them as a way to

sample and recruit women and they are finding that that's working out pretty well. Most practices are choosing a temporal type of sampling. There are cases though where they feel that they may not be completely keeping track of the denominator which is one of the concerns here that we've all experienced in field studies but that's something to consider as well. That's probably a little bit harder to keep track of in prenatal providers than in a hospital.

DR. BRACKEN: Actually it's not because the birth certificate will eventually give you the provider data and the hospital data and you can figure out what you should have expected.

Can I just go back to a comment that was made this morning as it relates directly to this topic? In the document provided by the Children's Study on power calculations, it's got these enormous groups, musculoskeletal defects and it shows what you could estimate, nervous system defects, major birth defects, neurocognitive development, neurodevelopment disability groups, developmental disabilities.

These are not ways that people study these things at all. For birth defects you want to know about congenital heart malformations and even in there subgroups. And you study these in pregnancy cohorts so by going from 100,000 to 45,000, you've totally wiped out even the bit of

power that's demonstrated here and even this is inadequate. So it just seems like not the way to go forward when we need more specification of disease, not more grouping and so you can't study these things at birth adequately any more.

DR. CARLSON: I would like to open the discussion up to the floor, to questions from the floor and I'm sure there'll be more discussion amongst the panelists for those questions.

DR. DEARBORN: Dorr Dearborn from Case Western, just a point of information. The Vanguard recruitment was not limited to first births. So if we do not limit to first births to a prenatal component, we've got a mix of primips and multips and won't that address the issue that we're talking about.

DR. GARFINKEL: I think we agree that we should have - I would say - given the objectives and the uncertainty that the biology of exposures is the same, it's worth testing and that should definitely be tested. In order to test that, you need to get really good data, early data, as early as you can possibly get in a small prenatal cohort.

And my only question is how big does that cohort have to be in order to test that question? I doubt it has to be more than 10,000, maybe it does but that's a

scientific question. That's one hypothesis. It's one hypothesis amongst lots of hypotheses. Michael is convinced he knows the answer to the question but I'm not.

The panel this morning, which has more expertise than I do, was not convinced it was true. If it is true, even if it is true, it would not be optimal to have 100 percent prenatal cohort. If you go back to Naihua's optimal sampling design, you would also have to say that preconception data is worthless and Michael in the end finally did concede that second births would be really good for preconception sibling births.

But you would also have to say that earlier prenatal data is not worth much. And most important, you would have to say that data for half your sample that's missing data, prenatal data on the first birth is worthless for all the different kinds of questions we would like to get. And what I heard this morning is it is far from worthless, even on the first birth. If you get breast milk or you get cord blood, you can get the continuing chemicals that stay in the body so the possibilities for imputation and the richness - Michael has dismissed that data as being completely worthless. I think that's silly.

DR. BRACKEN: No I haven't dismissed it at all. What I am saying is with your model, Irv, you don't have real time pregnancy data on first births and that is a very

dangerous position for this study to be in going forward. And I'm not saying we know for sure all of this. We don't. But we will never know if we go your route because we won't have the data to actually even look at.

And Dorr, you're right. Recruiting of pregnancy would include first, second, third and all other births but that's exactly what you want. You do not want to be controlling on anything at this stage, families which is another, as to Irv's point, is somehow you could control for family variation. He talked about confounding. You don't want to do that. It's a mistake.

You want to actually be able to study the effects of co-variants going forward and if you've controlled the study by matching and controlling and confounding, you severely limit your ability in the future to actually examine the role of those factors. So that's yet one more reason not to go down that particular road.

DR. PANETH: Nigel Paneth. This discussion, I have to be frank, reminds me of the tagline of my friend on email - okay, it's all very well in practice but does it work in theory because I hear we have experience, substantial, voluminous experience in prenatal recruitment.

And I don't know if Nancy or Irv have recruited in prenatal care settings, but Michael has just told you that he's enrolled 17,000 pregnancies in four cohorts, 30

percent of them as early as 8 weeks for under $2000 a person. I myself have been involved in seven studies, in fact that's all I've ever done really in my research career, is to enroll either births or pregnancies - four of them births, three of them pregnancies, six of the seven funded by NIH. It's much easier to recruit in pregnancy than birth. I don't where this hallucination comes from. You don't have IRBs to take care of.

Prenatal care providers, I've had one refusal of a prenatal care provider in some 100 different prenatal care settings. I've had, working for years in Wayne County where we have all the obstetric leaders we could not get 25 percent of hospitals to agree to even a protocol where the woman consented in advance to placenta collection.

As to the idea, let me be very frank here, as to the idea that a random sample of hospitals will agree to alter their protocol in the delivery room to do something different for the placenta, and the very difficult problem of collecting cord blood - by the way the amounts we're talking about have nothing to do with umbilical stem cells, that's not in the picture - is ludicrous. I'm telling you, it will not happen.

I'm predicting right now that if you start from hospital consent the mother protocol you will not get placentas and cord bloods from the vast majority of

randomly sampled hospitals in the US. Yes, some academic hospitals. It just doesn't happen. Not from theory, I'm talking from experience Irv.

DR. GARFINKEL: Just let me leave two things. I will leave you making great assertions about feasibility, I won't comment on that. I think Nancy can. I want to comment on the costs because it's been raised twice and you sent the document, the Vanguard that 30 or 40 of you signed on to the document on cost effectiveness.

I teach cost benefit and cost effectiveness analysis. I would not give a passing grade - you would not pass my course and let me explain why. You cannot from that document - it's very simple.

The first thing is you need to do accounting. You need to say we're going to do A, B, C, D, E and F, that's what you're going to do in the prenatal cohort and then you have to cost out A, B, C, D, E and F. I read your document. There is nothing on that. There are assertions about we did this study, the range of this cost was $5000, another one was $2000, no explanation of why those costs differed. No explanation of whether they were comparable costs. I don't even think you got the same years.

So for you to comment - the naivete on costs that permeates this discussion - that's my professional expertise. I don't know biology. I don't know that so I

have nothing there. But I do know how to measure cost and you don't.

DR. PANETH: May I defend what we did, fully stating that it was not intended for your course in cost effective analysis. Everyone can read it. It is a description of what actually happened in studies very much like this. It was from NIH documents, the numbers were added up. We described what we did. We didn't make the claim that this was a cost benefit analysis. We only were talking about the facts that we know and what we have done and I have done. And we have recruited in pregnancy, we've recruited pregnant women for under $5000 in studies in which we had interviews, we had birth collections, we had biological collections in pregnancy. And we know that to have actually happened over a period of years. Is that exact, we make no such claims as you have failed us for. But against what are weighing this, what do you have --

(Chair intervened in this discussion)

DR. MCLANAHAN: Stop, stop, stop. This is not what this is about, we're not arguing about this - no, no, no. Answer the questions that this panel was asked to answer. We will welcome comment against oppressions but we're going to talk about what information do we need in order to make these decisions (inaudible - off microphone)

DR. DUAN: I would like to make a clarification

around an important point that seems to have been lost in this discussion. The sibling cohort, which I do agree with Irv is probably the most practical way to get a good preconception sample, does not necessarily have to come from a hospital birth cohort. It could very well be coming from a prenatal cohort.

So we could recruit a prenatal cohort and then go on to recruit the siblings. The advantage of sibling cohort is we already have the agreement of the mother to participate in the study and relatively speaking we have some economy scale to recruit her for the next child.

So in some sense this question is not a question about hospital over prenatal care. The question is once we have a sample of the first child, what can we do to get additional children?

DR. BRACKEN: It is true. The sibling cohort can be developed straight from the prenatal cohort. That's not the issue. But I do have to respond to Irv's comment about the cost estimates because it's very simple. If someone gives you $5 million and they are all RO1 grants, what can you get for your money and what that document shows is you get cohorts in the order of 2000 to 3000 pregnancies with biological sampling. It's not complicated economics. This is the bottom line. That's what your grant costs, what do you get.

DR. KALTON: Graham Kalton from Westat. I just want to put on the table the existing provider based sampling methodology that's just going into the field right now and to draw your attention to one aspect of it. One of the ideas I think behind the birth cohort was you don't have complete coverage with the other.

And I want to describe the PBS because it actually gives you better coverage than a birth cohort. I'm making an outlandish statement just to get your attention. So what do I mean by that? The PBS is currently being put into the field takes as many providers as you can within an area and it recruits women from those providers.

Now how early can you recruit them and that's a critical issue. If you look at existing data on that something like 70-odd percent of women report that they have their first prenatal visit in the first 3 months. Now the question for the study is how quickly can you get to them? Is that really realistic for this recruitment and so there's that issue.

It turns out that very few of them don't have any prenatal care but the design is such that they are covered by a method which is to include a hospital as "the first prenatal care." So if they come in for birth and they haven't had any prenatal care, they are picked up there.

So you're getting that coverage.

So it gives you complete coverage. It gives you coverage also for the fact that if you're frame is likely deficient, and it may well be to some degree in the providers, they're picked up equally under the hospital mode. So what you're getting is data as early as possible for women who are being sampled through the prenatal care.

A question is it going to be early enough? But you're covering the whole coverage so you're getting a complete coverage at birth and so to my mind it takes away one of the arguments for the birth cohort.

And there are these issues about - another question you then have to face is which methodology is going to be more acceptable in practice. Is it better to recruit through prenatal practices that get the practice on board and the woman on board at this particular point or is it better through the hospitals?

I must say that Fragile Families did a remarkably good job with the response rate I think you said of 90 percent but as you point out it doesn't have any of the biospecimens and I think that's a really critical issue. It also didn't cover situations where the woman was ill or the baby and so all the birth defects and things.

You can imagine going to a woman after birth and asking will you join this study and the baby's got Down

syndrome or something. You're going to have some challenges there. And there's a remark about stillbirths in the documents we've got and I don't know what you're going to do about handling that particular question.

So there are issues that way but equally there are issues about getting into practices because they are busy and they have to make arrangements and so that's why we've worked out different methodologies with them to get a probability sample at first visits. Its first visit so you get them as early as possible and it avoids multiple counting. So that's the design that's in play at the moment and I think it has a lot of merit if it can be operated properly.

Just a quick comment on the siblings - I think the idea is a nice one but I think you have to start thinking about the operationalization of it. I'm talking about now for the preconception.

The plan with the NCS is to go to women fairly frequently after the birth, every 3 months in the first year and then it's every 6 months and you want to get to these women at a point immediately on pregnancy, on the point of conception. So how are you going to map those two things together and as you go further into time, the visits are less frequent and so on. So you have to work out the logistics of all of these designs to really establish which

one’s going to be best for the study.

DR. CARLSON: Didn’t the original NCS though have a pre-pregnancy data collection plan? The original NCS design had the pre-pregnancy component so I assume there’s something - data collection worked out.

DR. DUAN: I have a question for clarification Graham. So if I hear you correctly, the design that you mentioned as being carried out, seems to be in a sense a combination of the prenatal cohort and what has been called the birth cohort. So the births that did not go into prenatal is captured through the hospital in the births.

Yes, that might be a smart - we had some discussions previously that if scientifically it’s better to take the prenatal cohort instead of the birth cohort because we get at least some prenatal exposure data then the birth cohort could be used as a residual to try to recruit women who did not have births prenatal. Just clarification is that -

DR. KALTON: That’s the current PBS feasibility study that’s just in the field at the moment.

DR. DUAN: So that sounds like a very good approach. Another point I want to add about maybe the operation of the sibling cohort. In many studies many of my colleagues are doing in recent years, a good part of the study protocol uses information technology like mobile

devices. I guess it's because there is this notion about the ecological momentary assessment that tries to go beyond the kind of usual every 3 months contact but tries to encourage or invite the participants to do something and send some feedback to the study when an important event occurs. So I think with some careful planning and maybe taking advantage of technology it is a possibility that we might be able to try to get as close as possible to the exposure timing and that is really important.

DR. KALTON: Yes, these are pretty difficult things. I met this problem in a number of cases and it's pretty hard to implement that. But I think you actually - I've heard people wanting to get at least in the first months, so what you want to do is you want to get these women to contact you and say, I think I might be pregnant, she doesn't know at that point.

DR. DUAN: So I think that is the idea of the EMA that is being used in many studies now successfully.

DR. REICHMAN: What is the incentive for the prenatal care providers to participate in the study? This is a question for anybody. I mean these are busy practices I'm sure. They have a lot of paperwork with insurance. They don't need more work. I understand the incentive for the hospitals. I can tell you what that is. What is the incentive, why are they going to do this?

DR. MARKOVIC: I am Nina Markovic from the University of Pittsburgh and we have an NCS site and I can address that. We actually found with providers they like to have a plaque on the wall. We featured them in our brochures, the actual providers were in brochures with their own children and there was public recognition that they were supportive of the study and they felt that they were affiliated and providing back good science and contributing so that was really helpful.

DR. REICHMAN: So it's like a certificate on the wall, a degree or something.

DR. MARKOVIC: Just to comment then I have also participated in studies, like studies where we did recruitment in the hospital. And we found that buy in at the hospital was kind of top down and we did not get good cooperation in labor and delivery until we put a 24/7 research staff team in the hospital to collect the samples.

Whereas with our current cohort, we're paying the woman and/or her significant other whomever, 25 bucks to give us a call when the woman is headed to the hospital and then we show up and collect and it's a lot less expenses to pay 25 bucks for a telephone call than staffing 24/7.

I did want to comment on the first born issue. My mother says the first born is like a pancake. It's not quite right and you throw it away. And I feel like I'm

being thrown away here.

From a woman's perspective the life course I think that there's many significant changes that happen with the woman during that first pregnancy. She may be continuing to work. She may be smoking or drinking or have other exposures during that periconception time that don't occur with a second or third pregnancy because now she has a toddler in the house. And we did here that people were moving, 20 percent were relocating annually in one cohort and 40 percent particularly for a first birth people are looking to move from an apartment to a home or to a larger space. So don't throw out the pancake.

DR. DUAN: If I can echo as a first born at risk for being thrown away, I'd like to add a note completely outside of my area of expertise and I would speculate the issue here is not just the biology of the exposure health outcome but potentially maybe the sociology. The first born's parents are getting on-the-job training and later children are exposed to more experienced parents who might be better able to manage exposure or cope with the issues.

And Nancy's question, I'd like to add a note. From my experience with various provider based studies, one is definitely compensate them for the time and resources they had to devote. In some studies we do pay for a staff member to help us with the recruitment.

Another important device that I thought was part of the reason the HCSUS study was a very challenging issue of the provider based study was successful was we offered the providers opportunity to be collaborators as part of research teams and many of them did. Those are issues they care about. So that's in a sense an extension of the plaque and many of the collaborators are genuinely interested in what we're studying and they participated and they deserved to be collaborators and many of them did.

DR. REICHMAN: So you have 100s of collaborators on papers and things or it's just the study team.

DR. DUAN: We have a very - our papers are Franko, Shapiro, Duan, et al and HCSUS Research Team and that's a very long list and very appropriately so.

DR. BRACKEN: I would just add that the proof is in the pudding or the pancake. I mean the fact is providers do contribute to research. It's just not a problem. And how you manage it varies enormously between providers. Some you might put a few 100 bucks into their Christmas fund or they usually have a slush fund or you put something on the wall so you do all kinds of things.

But this is actually why you need local knowledge working with providers and I think why it's going to be very difficult for contractors to come in from the outside and to try and manage this process because a lot of it's

personal relationships. The people in the provider's offices have clinical appointments in obstetrics, in your own hospital so they are colleagues in that respect. They are much more likely to be receptive to you going and talking about research to them than they would be to an outside group. So I think this is an area, one of many, where losing the local academic centers will be a real detriment to recruitment.

DR. DUAN: If I could add a word to it, Michael. So the experience we had with HIV study was we went through a process to recruit prominent HIV providers in each sample geographical area as our site captain so it's a collaborator model and the site captain helped us identify the other providers, helped us recruit other providers. That tremendously helped the study to get into the door. I completely agree with the model you described.

DR. MCLANAHAN: So I just sort of want to say two things about this that, to me, we have to answer before we can answer the other questions. How well do the two sampling cohorts generate a good representative sample? I mean what's the real response rate under these two provisions?

My sense is that a lot of the stories about success with providers and people's own hospitals are based on convenience samples. So I'm sure that all of that's

true. I just don't know what the evidence suggests because that's what I care about the most is the integrity of the sample and the response rates at the beginning and at the end of the day. So we need to know the difference, is there a difference between the provider based cohort on that variable and the hospital based cohort?

And then the second thing it seems we really need to know how important are these 8-week measures because I think all else being equal if the provider can do as well on response rates it seems like it would be the sensible way to go because you're going to get more data on the prenatal care. But if it turns out that the most important data is in the first 6 weeks then we want to put more money into something that's going to give us the very, very early. And I don't know the answers but to me that's the information we need from the scientific community in order to make the decision about how to allocate the sample.

DR. BRACKEN: So it is true that most of the provider examples certainly that I'm getting are from convenience samples but when you've got 100 percent acceptance it's hard to believe that going to a random model you would actually get large numbers of defections. I just don't think that would happen.

And you can work out the probabilities of all of this and do the sampling in either route, hospital or

provider.

In terms of the 8 weeks, that entirely depends on the hypothesis. There are some exposures, such as cigarette smoking and the outcome being low birth weight, exert a lot of their effects in the third trimester so third trimester exposures are actually very, very important but the crucial thing is not to be able to measure early ones as well.

I mean you don't want to do a massive study like this and only be looking at late trimester exposure and you don't need to. In the provider model you will get, according to our data, about 30,000 women who are in by 8 weeks of their actually being found and prepared for interview at 8 weeks.

DR. MCLANAHAN: So I have to respond to say in Fragile Families, we did find that 70 percent got prenatal care but it wasn't in the first trimester and there's big issue for race ethnic minorities on this. There's a big difference in their access to early prenatal care and so I care a lot about the disparity piece of this. So I'd really want to make sure - and I'm not claiming I know what you would get if you were to try to do it this way. But I just think we need to know is there going to be consistency across the race ethnic groups, the income groups and what populations are we going to miss in the first 8 weeks by

just starting with a regular provider sample.

DR. GARFINKEL: It's wonderful you gave the example of smoking. So here is a question. Here's a question the study has to answer. If you did a birth cohort and I'm going to make an assertion and then say-- we need to verify this. My assertion is you can measure smoking throughout the pregnancy for the mothers that you interview at birth. And why do I say that? Because you can ask retrospective questions, you can ask questions within the second pregnancy and measure it and make the comparison. Your ability to impute really accurate data is, I believe, really high.

Now maybe I'm wrong about that but if it is, you have no advantage. Therefore, this is a critical question. If I'm right that you can address the prenatal smoking then you can't use that as an example as an advantage for enrolling the prenatal sample.

DR. BRACKEN: Okay, so my example was smoking and birth weight. But if I had talked about smoking and increased risks for some other conditions, you'd want to look at first trimester smoking.

So it entirely depends on the question that you're asking and you want to have the opportunity to look at exposures across the trimesters. Smoking, and actually for reasons that are unclear, women who smoke in the first

trimester, their children are at high risk for asthma. There are associations with smoking and some birth defects. And those are first trimester smoking exposures.

So why limit this study to just looking at exposures in certain trimesters. You need to look right across the board.

PARTICIPANT: I agree with that. Do you disagree that this is a critical issue for the design?

DR. GARFINKEL: I thought our charge was to point out what do we need to know in order to choose between the relative size of the birth cohort and the relative size of the prenatal cohort. And I'm saying a critical question is how good is our ability to impute missing data if we go with the birth cohort and if it's really good across all trimesters, first, second and third, on smoking - your example was smoking - I'm making an assertion I think it's really good, I could be wrong. But -

DR. BRACKEN: There's data on this and it's not good. If you ask them they will underreport smoking compared to, for example, measuring nicotine at different points in pregnancy.

DR. GARFINKEL: Definitely, I concede that point that does not mean it's definitely wrong, you can adjust for that. You seem to know the answer to the question. You're very confident. You're very confident about

feasibility. You're very confident about scientific knowledge. All I'm saying is I'm less confident than you on those questions. Shouldn't we address those questions in order for NICHD to make - isn't that a critical question.

DR. CARLSON: I think we're out of time, I'm sorry. I want to thank all the panelists for a very interesting discussion and I think we have a lunch break now.

Lunch 12:19

Resume 1:16

Agenda Item: Weighting, Imputation, and Estimation in Proposed Design

DR. MCLANAHAN: Okay, so we are going to start with session three which is on weighting, imputation and estimation. And Steve Cohen is the moderator of this session.

DR. COHEN: I just wanted to thank CNSTAT or the National Academies for a delicious lunch and hopefully your sugar highs will be brought down as we get deeper into this discussion.

This conversation is so linked to the earlier discourse that we had in terms of the underlying sample design that it's a really good continuum that we have these discussions on what the analytical fronts are in terms of

issues of if you are optimizing your sample in terms of the birth cohort, you would have one potential optimization scheme for the sample design that might not be in the similar dimensions in terms of stratification or the geographical multistage design that you would have from a prenatal sample.

And when you're considering composite estimation, these issues in terms of getting the same representation of subsets of the population, but that composite dimension is not necessarily the optimal design for an integration strategy. Naihua had a very nice slide up there that went in the direction of design optimization, if the NCS is going to go forward with the split of some sort between the birth cohort and the prenatal study.

But it begged for one dimension that clearly has to be specified a priori and that's the cost, the cost dimension not only for recruitment but for the whole longitudinality dimension. So when I saw that particular slide I was looking for like a cost optimization such like minimizing variance for fixed costs or minimizing costs for fixed variance.

But our panelists will at first tackle the issue of some of the complexities of a design that potentially is a dual frame design for estimation purposes and is longitudinal in nature with several stages of enrollment

and then survey attrition and over a 21-year period looking at strategies other than perhaps weighting strategies that are typical for adjusting for non-response to help minimize bias in estimates. We will also have some discussion of imputation and particular attention will be given to an issue of a situation where you would have say the birth cohort where you wouldn't necessarily have information, prenatal information, for the first births but you could conceivably do an imputation but that would be a total imputation for that population and what the caveats are on that venue.

So with that said, we have a very distinguished panel and we would start with each of the panelist's spending 10 minutes on these estimation issues. We would then have them react to their statements and then we would go deeper in some of those particular areas. So with that said, let me turn to Graham Kalton who is Chairman of the Board at Westat. Graham.

DR. KALTON: I am going to follow on a bit on the conversation I started this morning because my preference is not to think about two cohorts but a single cohort. And I want to lay out a plan for how you can do a single cohort and that then enables me to be able to present a method of handling all the problems that come out about missing data and so on.

With two separate cohorts, you've got to work out how you're going to fit them together and so on. If you can put the whole thing in a single structure you avoid that problem. How might one do that?

Well, there are different variants on the theme of doing such a design but the PBS design I described this morning is indeed just such a design where you take a sample of providers and hospitals are the provider of last resort and they are in the frame and therefore you've got the complete coverage at birth.

I failed to mention this morning why that's better than the birth cohort and the reason is because 2 percent of women give birth at home and you've got a potential of picking them up from providers - another big deal. But it does give you complete coverage so if you think about that, you could just go with that frame and go on from there.

Now, what are the downsides to that? Well, if you take a geographic area, you're going to have to list the providers and from that list you're then going to draw a sample and then if you need to go to the hospitals, you're going to spread your data collection across a large number of hospitals. And so is that a problem? So this is a practical issue that needs to be thought about. I'm not getting into those difficulties. I'm just pointing out the

different modes by which you can tackle these problems.

But you could go that route and that's essentially what provider based recruitment is doing currently. The study centers made lists of providers and we sampled those and then we go from there.

Now, let me now move to a different way of doing it which is being proposed by the program office also. And that is to say that rather than have that spread across all the hospitals in the way that was going to happen for getting the birth data, take a sample of hospitals first of all and get the providers associated with those hospitals. Now, if you do that then the births are all going to be in a small number of hospitals and that's a nice attractive feature.

There are issues because providers are not linked to just one hospital and there are a number of issues to sort out like that. And you have to face then the question about how much lack of coverage do you get out of doing that, not in the birth level but in terms of the prenatal data. So those are the issues that come up.

But whichever way you go about that, I put forward the following design which aims to be able to integrate things in a nice way. One of the problems as a survey statistician, the first thing you ask yourself is what's the defined population? And that's very nebulous at

the moment. It's spread over time in a vague sort of way so I'm going to define a population which is all the births in the United States in a given enrollment period. Now the enrollment period may be 2 years, the shorter the better but it has to be multiples of years because you want seasonality covered.

So let's suppose it's a 2-year period. So my sample is to be a sample of all the births that appear in this 2-year period and that means that if I'm going to get them prenatally, I'm going to have to start identifying them and enrolling them earlier and they are included in the study if they give birth in that period. And of course at the end of the period, some of them you may want to think about recruiting but they will give birth outside the end of the enrollment period -- so be it.

Now the neat thing about that design -- and then the others you pick up in the hospitals, of course you pick them up if they are in that particular 2-year period. The neatness about such a design is that it is all integrated but also you've got all the benchmark data from birth certificates that will enable you to assess how you're doing and make adjustments to the data because you can get from the birth certificate records all the information that you need. So it does add a little complexity because you've got to do the timing a little differently but once

you've done it everything falls into place.

Now if you then think about it at the point of birth then you should have complete coverage. Of course, there will be a few missed but let's say we've got complete birth coverage in some sense at that point or I've reweighted it at birth.

Now once I do that if I take that analysis from birth forward then I'm going to face attrition. There are going to be people who, kids that will drop out and for one reason or other we can't follow them. A standard panel survey problem, typically handled by a weighting adjustment.

You've got a nice benchmark starting point. You've got their weights and you adjust as you go forward using the data that you've got on them thus far. Now there's no reason why you can't reverse time in this particular modeling and look backwards. Because the prenatal data is of the same fashion except we are looking at time in a different direction.

So again we've got the birth cohort and now we can weight it back to say well we've got missing data if you like to think of it as a kind of attrition, we didn't get the first trimester for some of them, that's a kind of reverse attrition. And so we can think about the data geared around the birth which is when we've got the

complete coverage in both directions.

Now it's pretty messy to analyze, I'm not pretending otherwise but conceptually it gives you a framework in which to think about the problem which I couldn't grapple with without such a structure. Otherwise you're going to do little analysis on this group, a little analysis on this group, and how you put them together is very unclear. This provides that framework. And so to me that was an attractive approach to handling the problem.

Now, do we want two cohorts? Well, I'd need to hear arguments, which I haven't heard, as to why you want fewer cases with prenatal data collections than otherwise because if you want that, that can be accommodated in this design.

We can sample the prenatal data at a lower rate and then we add a woman - maybe we'd like to have a 1 in 50 chance of being in the sample at birth but only want her to have a 1 in 100 chance prenatally so we give her 1 in 100 chance. And then of those women who were in that category, we subsequently give them an extra chance to come in at birth.

Now I don't know why you'd do that but you could do it. If you really want to argue that, I want a big sample from birth on, then you could do such a procedure at the cost of giving up on the other. So I think the

argument is it's there if you want it and you just need to think about whether you really want to go that route.

What I said this morning was that you have complete coverage at birth through the provider based approach and that is correct in the conceptual way. And you do have an issue because you will have prenatal provider non-response.

And Colm and I were talking about that in the break and my first reaction was I needed to know for all the providers in the area as to whether they'd be respondents or not. And he said no you don't need that.

So I can classify all those that were - well, I'd have to work through how to do it actually, I'm trying to think this aloud because it was a casual conversation we just had - but non-respondents can be non-respondent providers, you have a chance at picking them up at the hospitals is the point I'm making. And so working out how you do that is another issue.

So, I'm trying to think of all the topics I'm supposed to be covering. I have no idea how long I've been talking for so you may -

DR. COHEN: Graham, you really covered the highlights. Maybe we can come back to the other issues in more detail but that uniform one design method that's integrated you've put out there so we can come back to that

in the discussion. So thank you.

DR. KALTON: I think he's telling me my 10 minutes are up.

DR. COHEN: So our next panelist is Dr. Colm O'Muircheartaigh who is Dean at the Harris School of Public Policy at University of Chicago.

DR. O'MUIRCHEARTAIGH: Thanks, Steve. I would like to say first that Graham and I had no consultation apart from the casual conversation we had during the break. Nevertheless, we arrived at a very similar conclusion. There are probably, as in the case of pre-pregnancy data, there are probably antecedent reasons for this rather than the purity of our rational thinking.

Again, I think the way I approach the problem is to think about what you're trying to represent in the study and what makes this study different from a convenience study of one kind or another. And let me say, I'm very encouraged by the fact that there seems to be a general acceptance of the principal of probability sampling when applied to whichever of these sample components people are talking about.

I've been involved in the National Children's Study I should say in terms of expressing my prejudices since I think 2002 to a number of conversations about the inferential basis of the study and I recognize there really

are quite different disciplinary approaches to the thinking about inference. I think the special characteristic of the National Children's Study is that it could allow for a population based inference for a study which is substantively much deeper and much more intricate than most such studies would allow.

If you think about it in those terms then what you really want to do is to think about how to represent the population in the sample. And I think everybody has been approaching that problem in one way or another and again I can't see any particular advantage for thinking about separate cohorts in the sense of thinking about different parts of the population or overlapping parts of the population and separating them at the selection point.

So if you really know what you're looking for in terms of coverage then it makes much more sense to think about the generality of the population and think about how you cover it in a unified way. Now there's a difference between a unified method of coverage and a uniform method of coverage.

And I think that one of the distinctions we want to make is between thinking only of one methodology as being the only way you can do it as a single study, or thinking about multiple methodologies within the same framework. And I think that's really what Graham was

saying, let me say what I think and you can decide whether they are the same or not afterwards.

So rather than think about this as two separate cohorts, you think about covering all of the parts of the population with which you're concerned. And if you accept that it's possible to cover a fairly large component, large part of the population through a provider based sample, there will be some faults in the frame so some provider's will be missed in the frame. There will be some non-response by providers once you go into the field but the remainder you can cover and cover respectively in that way.

You're still left with the part of population that hasn't been covered due to these two failures, the non-coverage and non-response. And that would suggest that even if you were a dedicated ideologue devoted to PBS, provider based sampling, you would think you need to supplement that sample with coverage for the cases that you've missed, either through non-coverage or non-response.

And the right place to go there is to birthing centers. So going to hospitals is clearly a sensible way to go there. Now if you think of that as a stratum, in other words think of that as a part of your sample design rather than as a separate venture unrelated to the other sample design, then you have a unified, essentially what we would call a stratified approach to the sample design, in

which you have no problem in accumulating the data across the two.

So the big problem I have with the cohorts is not with the concept of different ways of collecting data but that if you do it in an unorganized way - and I mean this in a very positive sense. So if you don't think in advance about how you're going to put these pieces together then you're not going to have a good entrance at the end of the process.

And therefore the argument that I would put forward is that you don't think of this as allocating a proportion to two different methodologies whose characteristics you're not sure about, each of them has benefits but think about it as using the appropriate methodology for an appropriate part of the population. When you have a single sample design you don't have any of the problem of dual estimation or multiple frame estimation or trying to figure out what the joint probabilities were with births which came from one cohort rather than the other because it really is only one design with multiple components.

And it seems to me that this ought to be something on which we could agree. We can subsequently debate what proportion of the sample should be in each of these strata but now thinking of the strata and not of

different approaches to the same population because they are really tackling different parts of the population.

I did want to mention briefly the point Dr. Duan made earlier that in comparing the costs there's been a regrettable concentration on recruitment costs rather than the costs of the study. And most of us I think would agree if you want to think about relative cost, you think about the whole cost of the venture which is the cost of the 21 years of data collection. And you don't disregard something because it has higher costs at one particular point which may well become negligible if you cost it out over the full length of the NCS.

Happily, there are many problems with whatever solution we might come up with because otherwise we wouldn't have an opportunity to meet in these pleasant surroundings over a number of additional decades in addition to the decade in which we've already been meeting to talk about these problems but let me say I look forward to many more decades of discussion of the approach. Ideally, in the presence of actual data as time goes on, so it would be nice to think that we're actually recruiting births as time goes on because it would give us a little more substance to talk about.

There are serious problems with waiting and I mention this only because it's part of the remit we had for

this group that you have to think carefully about the probabilities that are already in place for the counties, for the PSUs, for the study centers already there. And you have to take those in to account in terms of thinking how you might sample either from birthing locations or from providers because you've already given a very high probability to locations like Cook County for example which is a very high probability of selection as a county and therefore you have to use a much smaller sampling fraction within Cook County if you're to have anything like reasonably balanced probabilities of selection for different mothers.

In conclusion, but happily we can talk, we can figure out all of these things at relatively low cost and only a small number of years. I would like to say that it seems to me obvious, for which I have no data, this is an alternative way of expressing, I don't know why but I believe it, that siblings should be included and it seems to me there are obvious advantages to including later births to women who were recruited into the sample. And again that has nothing to do with the original point of recruitment.

It's not dependent on whether you do it through hospitals or through providers or for that matter through households. So I think that's a separate argument if

people should wish to have it but on balance it seems to me most people feel that recruiting subsequent births to a mother recruited into the NCS has strong advantages both in terms of pre-pregnancy, prenatal and other measurements that seem well worth having.

DR. COHEN: Thank you, Colm. We are going to have another discussion by Rick Valliant who is a professor at the University of Maryland and also Professor at the Joint Program in Statistical Methodology tied to the University of Michigan.

DR. VALLIANT: Thanks Steve. Unlike Graham and Colm, I'm a person who knows almost nothing about NCS. The good thing about recruiting somebody like me is you may get some new ideas and the bad thing is you may get some ideas that don't make any sense. Because of that I'm tempted to say I agree with everything that Graham and Colm said and then sit down.

So this unified design that Graham talked about seems like it really has some advantages to me. And my study of the NCS started a couple of days ago so I'm not up on all the previous discussions but it does clarify the thinking very well. So I have three sets of things I'd like to talk about which I think mesh pretty well with what's been said here.

One consideration is optimal design, which even

if you think in this unified design approach, not how many do we allocate to a birth cohort, how many do we allocate to a prenatal cohort, you still have allocation issues in the sense of how many PSUs should you have, geographic PSUs, how many providers per PSU and then how many women per provider should be sampled.

And the usual way that we try to do that if we had data would be to estimate variance components associated with each of those steps and that would require identifying one or more statistics that you think are important.

They could be descriptive statistics. How many women had underweight babies? How many women were exposed prenatally to something? What the relative risk is for certain condition? The more complicated the estimator is all you do is linearize it, get the variance of the linear combination, write it in such a way that you have variance components for PSUs, providers and women.

Now the big trouble with that is, as I understand it, there's probably insufficient data that's directly related to the variables that the Children's Study is going to collect. So you have to cobble together whatever you can in order to at least make somewhat informed decisions about these allocations.

There is the Vanguard data which is only about

4000 or 5000 women at this point I think. There are somewhat related datasets. The NHANES data at least is health related and it's got a lot of physical measurements on different people. The American Hospital Association publishes hospital data.

The problem in deciding, particularly how many providers do you want to sample per PSU, boils down to thinking about how much alike are the women within a provider, the women who visit a particular provider. But another way to think about it is how much do the providers themselves differ in terms of size and the way the math works out in this, the variance between providers is highly dependent on how many women they service. And there's no control in this population over that.

There are providers that are hospitals. They can serve hundreds, thousands of women in a year's time. There are individual doctor's offices who are much smaller. So there's this built in disparity in size and mathematically what that's going to push you toward is sampling more and more providers rather than more women per provider. Which, you know, there's cost implications of that so you have to think about that even setting up a problem to get a good allocation.

In fact, if you had enough data you could do something like Professor Duan was talking about earlier.

It turns in to a mathematical programming problem. If you have a bunch of different statistics you're interested in you can weight them, their variances according to their importance to the survey and then you constrain the problem.

You have a fixed budget as a constraint. You may want to lay on constraints about I'll select at least some minimum number of providers and women per PSU and so forth and to do that you need a lot of data. So many times in designing a sample where you don't know much, you really have to make some rough approximations in where the data would come from to flush this out. I don't understand enough about NCS and what's available. So that's one issue.

Another issue is this idea of missing prenatal covariates for women that you only pick up at the point of birth. So even in this unified design there will be some women that you get only at the hospital assuming you can recruit them there. You won't have prenatal covariates except to the extent that you can get them by recall or consulting medical records for those women.

So one option there would be to use the sample women for whom you do have prenatal covariates, exposures of different kinds say and use those as donors to impute for what has been called the birth cohort but in this

unified design you'd think of it as just women who are missing the prenatal covariates.

So one way to do this is with multiple imputations and to ever evaluate whether you're doing a good job that way, you need an imputation model that is correct and can be quoted since these models are never exactly right. You need a model that's a good enough so that if you do the imputing you put these two types of women together, the ones with and without actual prenatal covariates. You can estimate population values for the inferential population that Colm was talking about.

So what one possible way to try to evaluate that would be through simulation I think. Because this is not a thing that's easily subject to analysis, mathematical analysis to assure you're doing the right thing.

So if it were possible to put together a pseudo-population maybe based on the Vanguard data, maybe based on NHANES or something else and then divide that population into women with and without the prenatal covariates. And that's an experimental parameter.

If you're worried about the fact that there's sort of an imbalance between those two groups that will occur in actuality you try to incorporate that into the simulation design and then just do this many, many times, see if you get unbiased estimates of that pseudo-

population.

Another statistic that could be used that we normally use in one of the EMI multiple imputation operations is the fraction of missing data. So there's a question of if you have to impute so much data, are you doing more harm than good. That's one of the questions we were asked to address.

This fraction of missing information which is essentially - there's a between and within component for a multiple imputation variant so the between is a measure of what variability is being injected by the fact that you're doing the imputing. If that's a big proportion of the total variance then you feel uncomfortable with that. So that could be measured in a simulation study, in addition to mean square error or accounting for bias.

The third issue we were asked to look at is estimation in general and this idea of oversampling has come up. Do you need weights? I'd say yes if you're actually going to do oversampling. Usually the oversampling is done because you're trying to pinpoint groups that you think are different in some way, low socioeconomic status or a certain race ethnicity group or something like that.

And even if you like to think in terms of models, which I do, the fact that you've oversampled those

different groups is sort of prima facie evidence that you think they're different and they probably need a different model that's going to lead you, at least in the prediction sense, to requiring weights that are different for the oversampled groups. So I think weights are important.

More importantly, there's the issue of whether the sample that you start with and that changes over time is a decent representation of the population and to make that happen you're going to have weights. Graham alluded to this fact that there are vital statistics records and birth certificate data that NCHS publishes. Those seem like the obvious sources of control totals for different things.

So there are two questions there. Where do the control totals come from? It seems like Vital Statistics is place there. Which ones do you use to create calibrated weights? That's kind of a modeling problem I think.

There are at least two reasons that you want to use these calibrated estimators. One is if you undercover or mis-cover different parts of the population, you know, if you do a poorer job of recruiting lower socioeconomic status women for instance, and in a household survey in the US this is kind of typical.

There are certain groups where if we draw household sample we make inverse probability estimates of

total number of people, like young, black males. We get about 75 percent of what the last census showed. That's sort of standard and it would be no surprise if something like that happened in the NCS so this calibrating can attempt to - can rebalance your sample versus the population.

Of course, if you do that the crutch you're leaning on is that the sample women you got are a good representation of the non-sample women. So you can weight up the one's you've got but if there sort of a skewed representation of the lower socioeconomic women then calibrating is not going to save you from that.

The other reason is reduced variances. But in any case you look for covariates that are related to coverage and to the substantive things that you're trying to estimate and there are a lot of things on birth certificates that I assume are available - birth weight, APGAR score, whether the infant required assisted ventilation or they were admitted to the NICU.

There's a whole bunch of potential things that you can tabulate in addition to mother's characteristics. And those are all fair game for control totals. So that's a research project to figure out which of those to use.

DR. COHEN: Thank you, Rick. So in term of some of the comments by the panelists, I'd like to just go a

little deeper on the issue of how estimation again connects back to the sample design. And Graham put forward say an integrated unified design as an alternative to the split that we have in front of us in terms of whether 50/50 or some sort of balancing. And Colm was in the direction of what Graham proposed but he moved a little further in terms of defining say the prenatal group and the birth group from the hospitals as different strata, perhaps overlapping, and also good alternatives to consider.

But if we go back to what's still on the table and we think of situations where ostensibly, and we know both ways of collecting samples will not get the entire population. But for that subset that is covered by say the birth cohort and the prenatal cohort is a very large representation of the children that you want to have in your sample.

There are situations where you have the same target population. It's often done where you have multiple data collection organizations and you find that over and above the fact that you should get comparable estimates, there is a data collection organization effect. So, if in fact there actually was an effect of the vehicle of data collection actually having more representation might be an advantage.

And even if that wasn't the case, are there

strategies if the NCS went in the direction of these separate cohorts to have composite estimation that actually has greater precision than - well, you're going to pool the estimates but in terms of the best way of pooling the estimates over if you had the full sample in one of the designs.

I think in small area estimation it's been demonstrated that ways of weighting the different composites could actually sandwich the precision. So just to get back to those issues, if the panel could react to that.

DR. O'MUIRCHEARTAIGH: So the first part is to clarify, and thank you Steve, to clarify strata do not overlap so the strata point design is a non-overlapping design. So the principal is simplicity where possible.

So my default option would be, were it possible, to have equal probabilities of selection for each birth in the defined inferential population but an efficient design so that it's not simple random sampling but it's a design that takes advantage of structured hierarchalistics and much of that is already in the design.

So I see the combination, I don't see this as a conflict between the idea of cohorts and the idea of a unified design. It's simply seeing the cohorts slightly differently as strata rather than as possibly overlapping

units with very difficult to disentangle joint probabilities.

So by thinking of it as a single design, you identify for each birth which stratum it's in. Some of this will be empirical. So for example providers who refuse to participate, their patients will arrive in a hospital, can be identified as being from this provider and they are therefore eligible for selection in the hospital.

Births where there is no provider will also be identified as not in the provider frame and therefore as eligible for selection in the hospital so the combination gives you essentially in principal a probability sample of all births.

And all you're doing in the cohorts is identifying the most effective way to get them. You take a hit on those who are included only in the hospital and that you don't get pregnancy data, you don't get pre-birth data for those. You couldn't have got that anyway so this is not something you're giving up, this is something that you wouldn't have. And it seems to me you'd want quite a strong argument to go against the notion of uniformity in the probabilities of selection.

In other words, the size of these cohorts if we want to think of them as cohorts or the cohort/strata will be determined by the empirical reality of data collection.

Where there are a lot of births for whom there were no providers prenatally than that sample would be larger because there would be more eligible births at the point of delivery from that stratum or that cohort. If there were a large number of providers who refuse than equally there will be a larger size to that cohort because they would present in the hospital as non-respondent providers and therefore will be eligible for selection. And if they are providers who are missed on the frame, they will also present in the hospital as not having been covered in the other strata.

So that it's not necessary to decide to now how big the cohorts are, the cohorts will define themselves because they're strata in the population rather than a predetermination that we need to make about what their relative sizes will be. So it seems to me there isn't conflict. It's simply by separating the cohorts, by making them straighter rather than allowing them to overlap, you get a coherent design where it's possible to make decisions on cost optimization or whatever, but you know what you're covering and the analysis and estimation is much more efficient and also more valid than it would be otherwise.

DR. COHEN: Good clarification. Any other reaction to the original design as specified and these alternatives.

DR. VALLIANT: Let me just say one thing that Graham told me earlier. This idea of the hospital being the last resort selection is important in the sample selection in this unified design because if you sample things, probably proportional to number of births that they deal with, you'll tend to get these really big hospitals in for sure. But the right way to do would be if the hospital is the last resort for picking up women who didn't get prenatal care, you'd want some adjusted measure of size that reflected the number of births they handle who were to women that did not get prenatal care.

So this is kind of a sticky technical detail but it also avoids this big issue that was brought up earlier in the day, I think that if you sample hospitals and a few big ones refuse then you've lost a huge number of births that way.

DR. KALTON: You are right, Rick, that one of the issues is how to determine a measure of size for the PBS selection of hospitals. Because it isn't the total number of births in that hospital it is the total number of births that would not have been picked up by the provider, whether it's because the provider wasn't listed, whether it's because there's no prenatal care and that's only I think about two percent of women don't have any prenatal care so that's a pretty small number. So it depends on that

compilation and there's an issue there and you're not going to get it right but you want to get it in the right ballpark because you're going to apply sampling fractions to it and that'll clear it up. It's just you'll have a variable sample so you don't want to cap that too wildly out of line.

DR. COHEN: Before we move it to the floor, in addition to estimation issues and design issues, there potentially is going to be this set aside special population sample of about 10,000 and one of the speakers this morning talked about using that as a vehicle, or at least a portion of that, where you get some retrospective data and then you have more detailed data whether it's on prenatal care and you could use this for modeling. That's one use of the data.

There's also going to be attrition over time. Graham talked about weighting back to the original sample but after something like 15 years with all the different levels of non-response, you might be fairly low in terms of your overall representation and the idea of potentially using this for a replenishment is a possibility. So are there insights in terms of recommendations for that set aside sample to inform this design.

DR. KALTON: Let me just comment on question of the attrition and whether you try to replenish.

Essentially, I see the NCS as being a longitudinal study. So if you've missed the first 5 years of the data what are you doing at that point? So replenishment has its limitation. I think many panels don't try to replenish in quite that way because they are focusing on the longitudinal aspect. Replenishment would also be - I'm trying to think how you'd do it is another question.

DR. COHEN: You could have a spare cohort that's going simultaneously - there are different way but -

DR. KALTON: But there are real difficulties of dealing with that. I think my own sense about the special populations is that that could be reserved for such things as births that came about from assisted technologies and so on. And you could identify clinics that do that and if you could factor them in, that kind of thing.

Now you could design a study of that sort. It's not going to contribute to the national estimates in any way. You're oversampling at a very high rate from miniscule population so it isn't going to help there but if you have special interests in that then you could use this methodology or the NCS techniques and apply them to that and then give you a benchmark comparison.

DR. O'MUIRCHEARTAIGH: On the issue of replenishment, I think it is a difficult issue. I agree with Graham that this is a longitudinal study and its

strength comes from the fact that you've measured people very early on and then continued to measure them over time.

You could of course conceptualize a replenishment as in the model that Graham had earlier, you would then make inferences backwards from the sample that was boosted by the replenishment, so you could think of this as a sample of adolescents or whatever, you're following them forward and you can track some of their characteristics in relation to the earlier panel. It seems to me that's a question that would arise in 10 or 15 years, 5, 7, 10, 15 years down the road, if you decide that there's enough interest in specific characteristics of the population that you want to get some within group information at that age that would be beneficial because the precision at that age has become less -- but of course you've lost all of the prior data.

I expect Rod Little eventually will stand up and say everything is a missing data problem so I'll stay it on his behalf. All of life is a missing data problem.

So the less data you have, the less good your inferences will be. But you could argue that in 15 years time there might be a benefit to attaching a parallel cohort at that point, assuming that you would cover some of the cases that had attrited over a period of time in the National Children's Study. But I think that's an argument

you would want to make then on the basis that there was enough information to be gained from it that it was worth doing.

On the special panel, I'm agnostic. It doesn't contribute to the overall National Children's Study. Whether it's desirable essentially to set some money aside from within this budget to tackle a specific problem that can't be tackled within the framework of the NCS, that's a decision people can make on the basis of the value of that study on its own. You can do some linkage because it will be contemporaneous with the National Children's Study, will have some characteristics that it will have in common but if it's not linked to the design then it doesn't really give you much strength in terms of inferences you make about the national sample.

DR. VALLIANT: I have one little comment about the attrition. The University of Michigan does several of these longitudinal surveys. The Health and Retirement Study, the Panel Survey of Income Dynamics and what happens in those surveys - they have a number of cohorts going. Every 5 years or so recruit another bunch and if the attrition compounded over time, if you lost 5 percent every year, eventually you'd be down to nothing.

But what happens in practice, at least in those surveys, is your big losses occur immediately, total non-

response, people who don't want to cooperate and then among those who do cooperate, they tend to stick with you. The Health and Retirement Survey is older people and I think they like to have somebody to talk to periodically but attrition is very low after say the first couple of interviews. So in the National Children's Study, I'm sure you can convince people how important it is and once they sign up for it, they'll stay on.

DR. O'MUIRCHEARTAIGH: Perhaps I could just add one thing to what Rick said. It's particularly difficult for NCS. But almost all longitudinal surveys have almost all of their non-response in the first wave. This is when most of the non-response occurs and the conditional response rates are really quite high, often over 95, 98, 99 percent, beyond that point.

This argues for minimal intrusion at the earliest stage which is exactly the opposite of the intention of the NCS which is maximal intrusion as early as possible to get as much data as possible but I would advise minimizing that maximum intrusion to the extent possible.

In other words, if you want to have these people in the sample for a long time then it might pay to be a little conservative in the early stages rather than feel this is the only time you will ever get data from them and you've got to get it now or all would be lost.

Some of the data you'd be looking for is not necessary at the earliest points, some is and I would simply argue that you want to select at the earliest point only the data that are actually necessary in order to maximize initial response because that's what's going to determine the long run response rate for the study.

If I remember correctly from PSID, and I'm sure Greg will be able to give a better number than I, after 20 years the overall attrition from PSID was something around 50 percent. In other words, the unconditional response rate, the proportion of initial cases that were still there was about 50 percent. And half of these were lost in the first wave. In other words, the initial response we have 75 percent was responsible for half of the lack of representation in the sample after 20 years.

So you really do want to think carefully about how you maximize that initial response so that you don't lose people from whom you could get very valuable information over a longer period of time.

And I do think that in some of the early stages of NCS, some of the early versions of the data collection had really quite extraordinary burden on the respondents at an early stage and all that's going to do is get you a lot of data about a very small number of people. So I would argue against it - at enormous cost. Other than that it's

a great idea.

DR. COHEN: I guess let me make one more point for the panelists to be covering because any talk on analysis and estimation has to focus on the variance estimation. And what Graham put forward and the difficulty in getting the right probabilities of selection that Colm pointed to under the split that's under consideration seems to be partially mitigated by the design that you put forward because it's going to be a complex design.

There's going to be the standard complex design procedures and then you might have additional complexities so is there anything the panel wants to point out in terms of doing the variance estimation for the estimates under the alternative models under consideration.

DR. KALTON: I'll just correct you on one thing, you said partially mitigated. It is totally mitigated because all this is now a standard probability sample. It'll be a clustered sample.

Just as an aside there, if you were to go the hospital route initially followed by providers, there is actually no need to do any geographical clustering initially. You don't have to, you can or you needn't. Because there is a list of hospitals, AHA has a list of whatever 6000 odd hospitals. They have the number of births in nearly all of them and so you could just sample

straightforwardly from that.

But if you did that it still would be, if you wanted to go that route, it would still be a clustered sample design. The hospital would become the cluster as distinct from the geographical unit but once you've got the geography, it's a multistage stratified design which we are used to dealing with so I'm not sure that I see any particular problems there.

DR. O'MUIRCHEARTAIGH: Unfortunately, I agree with Graham again. He makes it harder to speak at length, although not impossible, as you will find out.

So if you have a stratified design based either on initially starting with hospitals or providers and using hospitals as a component to the design, then the variance estimation is a trivial problem. All it will cost is statisticians and computing time and so on. It's not that it's easy but we know how do it. It's simply a straightforward problem.

If you have a separate cohort design then it's essentially impossible, unless you do a lot of work on trying to unify post hoc the probability selection of each of the births across these two cohorts. And I would argue that you don't get any benefit from the separation.

And that therefore really all Graham and I are suggesting is a slight rationalization of the structure

where you accept that there are these two components in the data collection, that there are two different ways of getting to your objective which is births and to cover them appropriately across the population and then it becomes a standard estimation problem.

It will be more complicated if the design is a little messy. It'll be less complicated if the design is cleaner. A lot of these things will be dictated by practical considerations but it provides you with the possibility of a valid variance estimate. Two separate, kind of overlapping but not entirely clearly how they do it components or cohorts will make that impossible.

You can kind of guess at it but there's no need to would be my argument, not that we couldn't tackle it if we had to but I don't see any strong reason why that would be a good idea. It would still be better to have strata regardless of the proportion you put in these strata. So my argument would be the default, make it as close as possible to equal probability for each birth. That's a good plan in the absence of overwhelming information elsewhere. But should you decide not to do that keeping it as a unified design still makes it possible to make appropriate inferences that would satisfy not just the population-based people but model-based people and all other based people because it would be a coherent design.

DR. COHEN: The floor is now open.

DR. GARFINKEL: I have just one question. If you don't have data on the proportion of births that are served by each of the prenatal clinics, how is it, so you don't have the denominator where you clearly have that in the hospital, how are you going to weight back up?

DR. KALTON: I'll give you two answers to that. The first one is the way in which the PBS is currently going forward is they have determined estimates of the number of births from each provider and their sample would be probability proportion to size.

But even if we didn't have that, we could still get probabilities. We'd say okay, we're going to take in this particular place we're going to take one week in four. So the weight is four times whatever the cluster probability was. So it's not difficult to get those probabilities. You just want to get efficient probabilities and that's a little more difficult.

All we need to know is the probability of someone being in the sample. The probability of being in the sample is the probability of the clusters picked times the probability that the providers picked times the probability that they're picked. And we can devise methods for all of those and just multiply them out.

DR. O'MUIRCHEARTAIGH: Perhaps to add to that.

We can fix the probabilities or we can fix the sample size. We can only fix both if we have a lot of information. So the absence of information simply means you have to accept more possible fluctuation in the outcome in terms of the sample size you finish up with. So fixing the probabilities is not difficult. Fixing the probabilities while simultaneously controlling the sample size requires information.

If you have good prior information, so should it turn out that these provider data are accurate, reasonably precise then you finish up with exactly more or less the sample size you had planned. Should they turn out to be completely wrong then it will allow for fluctuation in the outcomes.

DR. GARFINKEL: Both Colm and Graham are proposing to use the hospital only as the last resort. So my question, just in terms of this question isn't it simpler and don't you get better data and need less assumptions if you have the information on number of hospitals and number of births, isn't that more reliable?

DR. KALTON: Yes, but that is not we want. And it is possible to - you're right, what we want to do is to know the number of women who come to that hospital who didn't have prenatal care or who had prenatal care that was from a different provider. Now you have to guestimate it

or if you're lucky you can probably base it off some birth certificate data from the past year or something.

But the point being, I agree this is a problem. I'm not trying to minimize it but I think it's a problem we can deal with. The point being that what Colm just said was if we determine a probability, we misestimate that and we underestimate it, then we're going to have a sampling fraction which will mean that we take too many births when we get there or more than we planned, but it doesn't matter, we live with that.

DR. GARFINKEL: How would you know if you underestimated or overestimated?

DR. KALTON: We will draw a sample and we'll find out. We will say, go this hospital - the way we have been working with the providers and indeed with the plans for the hospitals and the provider based is mostly a time based thing.

So if we think if this particular place doesn't seem like it's going to have very many births then we would take a rather high number of time intervals there and we know how many we're taking. Let's say we're doing it over a year, there are 52 weeks. And let's suppose we go there for 5 weeks. We know the probability is 5 out of 52. Then we take all the births in those particular periods and the numbers of those, the total will vary because we've got it

wrong. But so be it.

DR. O'MUIRCHEARTAIGH: I think it's unfortunate to use the term last resort in this. It's the appropriate probability within a hospital. So it's simply saying where the birth has not been covered by our sample of providers, the hospital will be the stratum that generates this birth.

Now one part of it we guess which is the number of births with no provider, clearly these will have to be in the hospital. The second we learn in the field, which is the proportion of providers that turn us down, because the larger that proportion is, again the higher the proportion of the births that will be generated in the hospital rather than by the provider.

And sadly this whole operation has uncertainty. It would have uncertainty no matter how we did it. No matter how you sample you won't know exactly how many births you finish up with but you can determine the probability and you simply apply that probability and take the number of births that it generates.

So it's not a - it's something that's empirically determined by the population and not by some presupposition we have to estimate in advance approximately what it is, but the facts will determine what happens rather than our presuppositions.

DR. VALLIANT: Not having complete control over

the sample size in the survey is pretty standard. The other unknown besides provider cooperation is cooperation of the women. How many of them are willing to do this and it's kind of a traumatic time in life and one more job is something that they may not want to take on. So if you have to make advance estimates of what your cooperation rates at all these stages are going to be and even given pretty intelligent advanced estimates, there's still going to be some slack in what you end up with.

Probably in this case you could do what household surveys typically do which is create these replicates of sample units and if you go for 6 months and you can see you're coming up short then you release another replicate of the provider sample and go out and try to recruit them and try to control it that way.

DR. COHEN: Thank you. Michael.

DR. BRACKEN: Michael Bracken. So I would like to go back to imputation. Dr. Valliant very nicely made the point about the difficulty of imputing when you are imputing for a lot of data. And of course in the plan in front of us we'll be imputing, if we're lucky, half of the samples pre-pregnancy data would need to be imputed. And if we're unlucky, particularly if we're talking about rare disease, it's quite possible that in fact the entire rare disease group would be in the group where environmental

exposures need to be imputed. Is that not a problem? I would hesitate to do some epidemiology where all my exposure data was imputed and try and get that published in the Journal of Pediatrics.

DR. O'MUIRCHEARTAIGH: I didn't quite follow why half of it would be imputed but -

DR. BRACKEN: Well, because the proposal is half of the babies being sampled at birth so you would be -

DR. O'MUIRCHEARTAIGH: That may be a proposal. It certainly wasn't the proposal that we were -

DR. BRACKEN: No, no, not yours, but the one that the NCS has put forward to us.

DR. O'MUIRCHEARTAIGH: I think it is a misinterpretation to think of that as a proposal. I think it's more a starting point from which to have a discussion.

DR. KALTON: I think we hesitate to discuss this with Rod here.

DR. LITTLE: I'm going to use a little bit of notation and I'm sorry about this but if you're doing a regression, you have a Y, you have some Z variables that are observed and you have an X variable which is the early pregnancy variable that's missing for some cases. If you're imputing that value of the Xs just purely based on Y and Z that gives you no information really about the association between X and Y which is the thing you're

really interested in.

The only way you're going to get additional information is by having auxiliary data that you can then use to help with the imputation. That auxiliary data could be recall data or it could be data from other sources or whatever or time lag data, whatever it is you choose.

But it's important to realize that multiple imputation only helps you if there's some additional information to be recovered in the data you're imputing. And if you're really interested in the relationship between Y and X, you have to have some other variables.

DR. KALTON: I was going to make the same point in regard to the birth certificates. Because it would be very valuable to have the birth certificate data for all the sample because that would enable you to have a lot of data that may be useful for this purpose and may not be either in the regression as such but could be very helpful for that purpose.

DR. PANETH: Nigel Paneth. I really thank the panel for clarifying that really one cohort perhaps with different strata is the sensible approach and clearly separating the distinct cohorts would just cause more difficulties than it solves.

I'd like to though really raise a more basic question. And it has to do with all of these statistical

prioritization questions which are very deep and it's kind of obvious and it's come up so much in the side bars that these all depend upon what questions you're asking.

And right now the Children's Study is a study about every childhood outcome and every potential exposure. And with that as a framework to decide whether the prenatal is more important than the delivery, that the placenta is more important than postnatal is impossible to make.

And I think the struggle over sampling strategy and design reflects the fact the absence of any previous struggle over prioritization of public health relevant outcomes, key exposures that need to be investigated and their relationships. Some of those were subsumed, some of them, by the hypotheses we once had. Now we don't have hypotheses, and I think this vacuum you'll have to struggle against until such time as the Children's Study says really what it is about, what its priorities are and having not heard them, I don't see how you can come to any conclusions about which fraction of any sample should be oversampled, undersampled or not sampled.

DR. COHEN: Graham.

DR. KALTON: Just a quick reaction on that. The integrated design gives you 100,000 births to follow from birth forward. If you split them up, you've got this mixture, you're not sure where you are.

Now prior to birth we don't know how many you're going to get and how early you're going to get them. And that is a critical issue. I think you raised the question earlier about will there be subgroups of women who will, the socially disadvantaged, will not come in until late if they come in.

And so there are issues about the effectiveness of this strategy that need to be examined but it seems that if you're going that route, this is the best you're going to be able to do.

DR. O'MUIRCHEARTAIGH: I would also argue that if you define the problem as one of obtaining a representative sample of 100,000 births, that this unified design with equal probabilities of selection would be the best design.

This means that you are getting a representative sample of these births with as much information as possible. In other words you're getting prenatal data in as many cases as you can. You will get pre-pregnancy data for later siblings in as many cases as you can. So it'll maximize the amount of information contained in a representative of sample at births. And that seems a noble ambition and a fine achievement were it the outcome.

And then you can argue either before the event in terms of saying you wish to over-represent urban areas, inner urban areas, poor rural areas, and these are possible

within the structured design should there be a particular reason that you want to do that. But it gives you an opportunity to represent not only the things you know about but also the things you don't know about.

And one of the things I think that was mentioned this morning is that there are aspects of our environment that we don't yet realize or the one's that are actually killing us and not the one's that we're concerned about are the ones that are making us healthy that we're not concerned about.

So by taking this population representation approach where you maximize the information on as representative of a sample as possible, it seems to me that this creates a platform on which many studies of different kinds can be based including studies that we don't know about yet because we don't know that these are things we should be looking at.

And any departure from that where we make deliberate decisions to exclude parts of the population or to take only certain kinds of information from some people where we could have had more, seems to me to militate against that and therefore I would argue for the basic simple approach that Graham and I are advocating.

DR. COHEN: Greg.

DR. DUNCAN: Greg Duncan. So I heard Colm talk

about subsequent births. I didn't hear Graham talk about subsequent births. And I guess if we think about there's sort of a disconnect between the first session this morning and this session. Because we heard in the first session about the importance of exposures very early in pregnancy, potential importance of exposures preconception and unless there are subsequent births, in your design as I understand it, you would have - it's not that we want a representative sample of births. We want a representative sample of births for whom we have very early pregnancy information, exposure information.

We'd like to have a sample of births for whom we have preconception information, at least to some extent. And it sounds like from what I heard Graham talk about there's no way in which you would be able to collect preconception information and very early in pregnancy exposure information. Maybe I misunderstood that.

DR. KALTON: Let me respond to that. I was talking about what I view as the sort of basic design. The question of siblings comes up. I put two things on the table that Dave Hubble and I were talking about just a couple of days ago. And that is within the say it is a 2-year enrollment period, there may be some good grounds for saying, well, if there's a second birth in the family during that 2-year period, then it comes in as with a

certainty. And then to make sure the probabilities are right, you exclude from sampling those who have second births coming from the providers. And the advantage of that is - and I can see some real potential advantages of having sibling data.

People often want to make comparisons and so on and it also has some statistical efficiencies. There's a little wrinkle on the design. But putting that aside, I see the sibling sample as an adjunct in some fashion. I think it needs very careful examination to see how it can be applied to provide the data that you think you can get from it.

It's very easy to be facile and say, oh, we're following these women but we need to get these data and we need to get data from women at these particular points. Now what data we need to collect, I'm not sure what those data are. But if it's going to be blood draws and things and so on, how are you going operationalize that to make it effective? But it has attractions if it can be worked out but I think it could turn out to be very expensive.

DR. O'MUIRCHEARTAIGH: If I could follow up. First if I could answer as well, Greg, an earlier design which was widely advocated was household based probability sampled women in which you would interview women regardless of their pregnancy status if they were in the childbearing

age range defined. So clearly it's possible - that would be the way were there no costs or practical consideration - that would be the way to recruit the sample and wisely that was attempted. So I think I was advocate of it.

I did allow that practicality might hold sway. There was some evidence that it was impractical. The evidence may not be quite as strong as sometimes described but clearly that's no longer a part of the design but that would be the design that you would use for doing that.

Any design that doesn't involve recruitment of women in those age ranges regardless of pregnancy status is not going to collect those data so that's a policy decision, that's a science decision to say that is not something that is being attempted as a representative population based sample. If it were then it wouldn't be difficult to produce a design that would obtain it.

DR. DUNCAN: Just a follow on for Graham I guess. So I appreciate if you set this 2-year interval, you'll get some second births. It would be a rather strange sample, I mean short birth intervals.

The longer that interval is, right, if it were 5 years rather than 2 years, you get more births, there are more representative births. So I guess I'm just thinking about amending your proposal to include a longer interval but then instead of considering these siblings as just kind

of annoyances that might provide some interesting information about some things, we all have annoying siblings right.

Why not think about the births over a 5-year period, right, with oversampling births early in the period and you're getting some subsequent births, you're getting some first births later on. It's a headache for samplers I'm sure but it's another way of potentially providing a single integrated sample over a 5-year period that would include both the initial and subsequent births.

DR. KALTON: I'm trying to think that through. It seems to me a very expensive design by doing that. If I'm hearing you right, you're saying instead of taking a 2-year enrollment period extend it to 5 and then follow on the model that I just put forward and you get more of these other births.

I would argue in the other direction for a 1-year enrollment period for a variety of reasons of efficiency, of data collections, of avoiding problems of the field workers going with interview number five with this household and number two with this, number one with that and all these mixtures, that really makes a mess of things. But you've also got the providers changing over time and all of these problems.

You could do that but I still need to be

convinced that you can tie that design into the basic children data collections that will be going on. And get the data close enough to the point of time that you want it. And if you're looking for preconception, it means all these women have to be followed and go through questionnaires or whatever it is that you collect and only some of them are going to become pregnant -- was it 20 percent was the figure you mentioned - so you'd have to follow these women through for that period of time.

I'm trying to make a distinction between collecting child data which is a schedule of every 3 months initially and then every 6 months and how do you match that in to wanting to know about this woman having become pregnant almost immediately. You don't have a method. I think someone suggested we should have them send us - we should do pregnancy tests for them by mail or something, I don't know.

But you've got to get a method of data collection for the women to fit in with this otherwise it's going to cost you a lot and I'm not clear what you need to collect or how you would do that. That's my problem.

DR. COHEN: So it's a good question to be answered.

DR. O'MUIRCHEARTAIGH: I agree that it's a complex and difficult question and therefore I won't answer

it. But I'll answer a simpler question which is this doesn't affect the overall design that you would offer as the base design. In other words, it might reduce the number of initial recruitments if you had decided that you wanted to supplement it with siblings over a 2-year or 5-year period but the principal of the design would not be affected and that I think was -

DR. DUAN: So I'd like to follow up Greg's point about the duration of the recruitment window. I think Greg pointed out having a longer duration will enhance the representativeness of the sibling cohort. I think that's an important consideration.

I would think that in addition having a longer duration might have its own merits. If we focus the sample entirely within one year, we are bound by the idiosyncrasies that are happening in that year and having a longer duration gives us a better representation over time.

We're not really just interested in the population of the children born in 2014. We are interested in the universe of children who will be coming and going so having a longer duration has the advantage that it will help us capture variations in other economy and environmental events and in the weather.

So I just appreciate Graham's point that it will be more costly for the same sample size but potentially it

might yield a more useful scientific question.

Another comment I'd like to make is to follow on Graham's very eloquent proposal to look at the likely missing or early pregnancy data in this unified approach which I think is a very good approach. Namely many of the women who we can recruit through the providers will not be in the sampling frame until they pass the first trimester.

So I think Graham made a point where our statistical methods like weighting and imputation has been routinely used very successfully for missing data that occurs in longitudinal studies and we can apply the same method backwards to look to see what happens in the past.

But I think there is indeed a difference between time forward and time backward. Because looking at time forward, as I think both Colm and Richard commented, a good study with a good field operation usually has a very good way to maintain the sample over time. So I guess the conditional response rate after the recruitment is usually very good.

Going backward we're trying to impute missing data that is not in our control, this is missing data that occurred before we got our hands on the participants. So the missing data rate going backward will be much higher than going forward. And I think that this missing data methodology that can be sensitive toward assumptions

underlying. And I think because of that I would think that to supplement the data with either a sibling cohort or with alternative ways to get to the early pregnancy data will be helpful with this exercise.

DR. KALTON: Yes, I don't disagree. I think one of the key issues is what proportion of the women can you get during pregnancy and in the first trimester. That's an important consideration. But you're right, let me just quote what Colm said, the response rate was 75 percent in the first wave PSID and if we were in a position that the first trimester we could pick up the 70 odd percent that way, what's the difference?

DR. COHEN: Irwin.

DR. GARFINKEL: So Colm you made a point which I think is worth emphasizing that a lot of attrition is likely to occur early on. So if you get prenatal data that's very expensive and if there's a lot of attrition that, whatever the attrition is, very expensive data has been wasted.

You made another point which I disagree with which I think is really important. You said it doesn't matter when we spend the money. But if you believe siblings are important then it matters greatly when you spend the money because if it costs - take an example, it costs $18,000 to collect the data on the prenatal births

and it costs $2000 to enroll them and it only costs $2000 to enroll the mothers at the hospital. You can enroll 10 times more in a birth cohort than you can in your prenatal cohort and you're going to lose a bunch of those prenatals so it matters when you spend the money. That's just one example of why it matters.

DR. O'MUIRCHEARTAIGH: I think I should clarify what I meant at least - I have no idea what I said.

So were it to cost $18,000 to recruit one way and $2000 the other, that does not mean you can have nine times as many one way than you have the other. Because you have to think about maintaining these people in the National Children's Study throughout the 21 years.

So that's why you don't use only the short-term cost in determining what the optimum allocation is. If you were to think that for each child it would cost $100,000 over the 21 years or $200,000 over the 21 years, then the comparison is between $218,000 and $202,000 in terms of the cost of a case in the NCS.

It's only if you're thinking that the decision has to be made and how much money you're going to have this year that you would make that decision but that's entirely the wrong decision. And it's critical - I think it's really important to remember that these short-term recruitment costs are only a small fraction of the costs of

the case in the NCS. And the comparison should be made on the total cost of a case, soon to be discounted for a future expenditure, the total cost of the case under each of the scenarios. And they're going to converge, obviously they are going to converge, there's no reason to believe that the later costs are any different depending on the method of recruitment and therefore that imbalance is not 9 to 1 but perhaps 1.05 to 1.

DR. KALTON: I would like to agree with Colm and I think Naihua made the same point this morning that you should be looking - the cost of investing in a good sample is you pay benefits of that over the time so the investment is worth it. So that's one point.

The second point was in your costing - I'm not sure I fully understand it. It's like, well, I'm going to get all these prenatal data and forget about them because you're saying they're in that cost but you're not saying they have any value and there's value to them so it isn't a very fair comparison if I understood what you were saying correctly.

DR. COHEN: We have time for one more question. Jennifer.

DR. MADANS: Jennifer Madans, National Center for Health Statistics. I think I missed something about how the stratification worked. You made a point, I think it

was an important point, that if you use this unifying design you will get a good birth sample which would be equivalent to just taking a birth sample from the hospital.

It's clear to me the women who agree, you follow them, you get to the birth and they're counted. Then you have a group that the provider didn't agree to be in the study so all those are gone. Then you have a couple of - you have women who have no provider and then you have the women who the provider said okay but they didn't.

So then you're going to go to the hospital to fill in the cohort. How are you identifying at the hospital the women who have no care or were at a provider who refused? That's what I missed.

DR. KALTON: The way in which that's currently operating is the data collectors are given a list of the providers from which the sample was drawn and they are told if it came from that provider then that women is not eligible and that can be determined either prior to data collection or it's part of the screening interview.

DR. MADANS: During your week the sample then becomes everyone who does not have the provider that was in the sample, is that right?

DR. KALTON: No, you have to look at the frame. Did they have a chance of appearing from the frame?

DR. MADANS: But where are you getting that

information? How are you determining whether -

DR. KALTON: You make a list of all the providers, you make a list of - and so a woman has different routes of getting into the sample and she's only on one route. The one route is that they come in for their first visit to any provider. So when you interview them at the provider, you ask them have they had any other prenatal care visits and you then establish whether it was to one of the providers on your frame or not. If they have, not many of them will so it isn't a big deal. Mostly they would have been to this particular provider anyway. But it's the first visit, so that uniquely defines them.

For the hospital cases, exactly the same criterion, have they had any other provider visit at a provider that was on the sampling frame?

DR. MADANS: Where is that information coming from?

DR. KALTON: You check it. You've got a list of the providers.

DR. MADANS: That was the answer - that you have to look at everyone to determine they're not in scope.

DR. KALTON: It's a little variant on that actually in practice at the moment but that's for a different reason.

But yes, you have to have an eligibility

screener. And the eligibility screener would include age, depending on how you do it, I won't get into this in detail but whether they live in the county or not and have they been to a provider and there's a list of providers that's given to the data collectors.

So even if they pass - so they can be prescreened by the hospital as being not eligible because you can just look at the records. If they're not prescreened out then they go to a screener to make sure that they are indeed eligible.

DR. COHEN: So now that our session is 5 minutes over the period allotted, let's thank our panelists for very insightful comments.

Break 2:47

Resume 3:06

Agenda Item: Factors, Issues, and Values to Balance and Consider in Reaching Decisions about the NCS Design

DR. MCLANAHAN: So this is the last session and we're going to identify and synthesize some of the issues here and Greg Duncan is going to moderate the discussion.

DR. DUNCAN: Thank you, Sara. So this is the big think session for the day. The instructions that we heard from Steve today are to identify and synthesize tradeoffs. We have a terrific panel of big thinkers up here. So let

me go in the order that they are listed in the program. Ed Sondik first.

DR. SONDIK: So at 2:39, I said things really are looking pretty good. At 2:40, I said, now I'm not so sure.

In making my list of points, I wrote down areas of agreement, points of agreement, where there doesn't seem to be agreement, tradeoffs and then I have a couple of suggestions. I tried to look at this from the standpoint of the study and the study moving forward and what kind of information it needs to have to move forward at this point.

So in terms of agreement, and this is where I came up a little bit - I had written down, looking at this from the standpoint of the population being the kids born over a 2-year period and then all of a sudden we had 5 years. We weren't quite so sure about that but the idea of looking at it that way over a fixed period, it sounds to me like we really do have agreement over that and it's from that that you can then look at how you want to divide up that sample.

It seems as if a bit surprise to me that we have agreement - and when I say it was agreement it means that we haven't taken a vote but I didn't hear great objection to this either from the panelists or from the floor - about the preconception sample. We're saying that in the design we're evaluating here is that it's a relatively small

sample and I didn't hear people saying this is unbelievable, this is where all the action is and this would be a huge error to do that. So I haven't heard that.

The point about costs, I think there was real agreement that the way to look at cost is to add them together. And Graham I think said it well, somebody else I thought said it well, that you really want to look at the total costs, the total costs here. And it may be that recruitment costs would be high but you need to look at that in the context of what we actually get from the study.

There seemed to be agreement on mobility and loss as an issue but there wasn't really a lot of discussion of that. And it seemed to fall back on prior studies and that probably I shouldn't have put under my agreement area but I think there's agreement with that but it seems as if there's work to be done with that in terms of characterizing the design.

There was a line said kind of in response to the question I asked earlier that more PSUs are definitely better than fewer PSUs. And I suppose that's true. But there is a point in the description that we all had that talked about the possibility of going from 100 down to some smaller number, 40 whatever and I think that's a complex issue that hasn't been discussed and I would think in going forward that's a very important decision that needs to be

made.

But allied with that is something that has always troubled me about this - which is how to handle the geographic environmental variables. And whether or not these are clustered, whether these are - clearly if they are uniform then presumably we'll handle that with the right measurements but if they're not uniform then how do we bring that into the sample, the sampling design, the PSUs and so forth. But I put that under my agreement list because I have a sense that people would agree that in general more is better but I think the issue here really has to do with the operation of the study.

Now where there's no agreement or there isn't agreement, we really had very little discussion today over what we hope to learn from the study. And it really wasn't directly on the agenda but it was interesting to me that a prioritization of the questions to be answered really wasn't up for discussion today. Well, it really wasn't on the agenda per se but it's also a very complicated complex issue.

And that's something I think that again the program needs to be able to, I think, articulate pretty clearly in terms of the importance of this and the ability of the design to produce information that's important to knowledge and important to public health.

And I was struck by the three points that were made on the right hand side down here of the first panel about asthma, endocrine disruptors and neurological problems. And what struck me was that's all we really heard today about that and what it raised for me is okay how does the design, will the design enable us to understand the impact of some subset of environmental variables on those or the fact that there is no reasonable impact. In other words, what's the power of the study there?

Now in terms of tradeoffs, one of the variables that is open it seemed to me after the last discussion is the extent to which we want the design should have the prenatal measures. Graham said that his estimate is that 70 percent of the women see a doctor at that early point, what was it, 3 months, I think he said. But then someone brought up, yes but the action is at 6 weeks. And how important is that and are there ways of getting that. And I think that's a tradeoff in looking at the design, that's a tradeoff that needs to be considered -- how to get that.

The agenda started off posing that perhaps we could think of this - actually it said that this was the design, was 50/50 and then the 10/10 so to speak or 45/45 and then or was it 40/40, whatever, the point is - 45/45 and 10. And then it was raised perhaps the ratio could be

rather than 45/45 80/20 and that seems to me to get back to the science which I think is a crucial issue and we can't obviously solve that here. There are probably as many opinions here or at least half as many opinions as there are people here but that I think needs to be in the design.

The point that had occurred to me on occasion was the first born but I must say I never really dwelled on it being first born. I figured that was best, I guess, who knows, but my sister's a lot smarter, let me tell you, and accomplished. It strikes me that this is really an important issue and so in terms of thinking about a variety of other demographic variables or perhaps strata, that's something that should be considered. It's come up and I think when people hear about the design, I think that, it strikes me anyway, as that resonates. So I would consider that a tradeoff.

And Graham's design, I like the kind of the elegance of that. But I was also struck earlier today by the point that interviewing women in labor or just post labor raises a number of issues.

And I mean I can just imagine saying we'd like to enroll you and your child in this study so we could find out what's wrong with them later on. I mean this is not exactly what people want to hear and I don't mean to make light of it. I thought the points that came up earlier

about it really gave me pause and so I think in the design that's something that really needs to be considered. If somehow this could be done a pure prenatal sample, that sounds great, because it would eliminate that. I don't think that's possible so the question is how would one go about that.

And finally let me make a couple of suggestions that go back to the science and to the design. I think it's important, whenever I come to this, and I should say I'm an ex-officio member of the Advisory Board and I represent CDC on that, and whenever I go to that or really think about this, I think about what the power of the study is to determine relationships. And I understand I think the reluctance to say this is the specific set of hypotheses.

But it strikes me that a way to evaluate this is in terms of what could be called an exemplar set of hypotheses. And look at the power that exists within the design to evaluate these and the main information that we have on these are the two power tables which really don't get at relationships, the ability to ferret out relationships.

So I'm thinking about the firstborn and asthma. First born, the poverty level of the family, income level of the family, race and asthma - what power do we have to

determine relationships in that? And I, having the discussion with someone here and I raised this and the point was brought up well you really need data to look at that. But I don't think you do. Someone else mentioned simulation. I think you can look at this in terms of what the potential relationships are and does the study have the ability to identify that relationship.

So I thought given that there's relatively little time available, it occurred to me that a panel, there could be a panel that would look at a set of science questions, reasonable science questions. And prioritize those questions in terms of their importance and given those priorities, perhaps a second panel or this could be done internally, look at this set of priorities in terms of the design and the ability of the design, the capability, the power of the design to indentify relationships.

And it strikes me that that would be a very powerful argument in putting this before the decision makers, up through the chain of decision makers in the department and beyond and including Congress and saying this is what we're focused on. And we think this is really the best bet to do this but also to get lots of other information.

I just have one more comment, that in the past Framingham was raised as something in which there were not

a set of hypotheses to begin with. Well, I really wasn't there at the times to know whether there were or there weren't hypotheses but certainly with a broad database, there will be possibilities for exploring relationships.

But it's important I think to be able to say here's something we know we have enough power to look at this. We know we have enough power to look at this but we don't have enough power to look at this relationship and it strikes me, keeping track of those estimates early on as the study progresses and as the sample develops, would be a very important management tool. Thanks.

DR. DUNCAN: Thanks Ed. Next up is Rod Little from the University of Michigan.

DR. LITTLE: So since we've talking about firstborn siblings I was going to start with a joke. It doesn't quite work because I have an older sister but - I'm a twin. My mother was age 40; she didn't know she was having twins so my other twin tells a story. So he says that when my mother had Janet my older sister, he said it's a girl and when they had me they said it's a boy and when they had Chris they said, it's another.

Okay so I have a few random comments. I was on the Federal Advisory Committee early on in this study so I've seen it evolve over the years. It's been an interesting phenomenon. One question that - and some of

this echoes what Ed said I think. So what is the study about in some sense - I think there needs to be some articulation of this. In particular, what's it adding over existing studies?

I think that in the last few years there have been a lot of new studies that have come out. There's a tendency in the States to think that the whole world is the United States and there's nothing else going on anywhere else but I think it's worth paying attention to what people are doing in other parts of the world.

When I was on the Advisory Committee earlier on, there was this huge effort to develop hundreds of hypotheses so they had lots and lots of committees generating hundreds and hundreds, literally hundreds of hypotheses in every conceivable area. And I was actually a little bit critical of the scope of that effort although it was very laudable in some ways I think. But it seems like we've lurched completely to the other end now so now we have no hypotheses, it's just a data platform that's somehow going to address lots of different things.

I think there's a happy medium somewhere between having hundreds and hundreds of hypotheses and having only a few and I don't see any obvious way to make decisions about optimal design without having some specific objectives articulated through hypotheses.

So I would suggest that NICHD comes out with a set of relatively small number of sentinel hypotheses that they view as being sort of burning issues in the area right now and then show me some power calculations for those sentinel hypotheses. If you do a $5 million study these days and you go to a study section, you're expected to produce a reasonable power calculation to show that you're doing what you're doing. So if you have to do that for a $5 million study, then I think you should have to do it for a study of this scale.

The fact that it's not a university-based study seems neither here nor there to me. So I would really like to see some detailed power calculations. You have Vanguard data. Maybe there are other data sources you can use. So really spending some time to try and develop a detailed power calculation seems to be very, very important and this fits in with what I had said I think.

One comment in terms of the subject matter - I think this workshop has focused a lot on the role of prenatal exposures, particularly environmental exposures. But I think it's important to bear in mind, particularly if you haven't been in the game that my understanding is that that's one component of the NCS but it's not the only component. There's a lot of interesting work that happens after birth and so I mean you get a distorted view if you

think the only thing that matters is what happens with the prenatal exposures, although clearly they are very important.

One thing that I like about the draft that was put out, as I mentioned earlier, is the supplemental sample of 10,000. I agree with Colm and Graham that in general, for most of the sample, at least an equal probability sample design makes sense, particularly given the fact we don't have very clearly articulated hypotheses.

On the other hand, I think that getting a good variability in some exposures has a lot to be said for it and I would personally be interested in seeing some index of environmental risk or something and oversampling areas that have high areas of that risk. I could see some benefit in doing that since that might increase the power for looking at some of these associations and might be a worthwhile way of spending that additional sample.

Then on sampling designs - I must say I think there's been a lot of progress. So I'm really very heartened, as someone who spoke up for probability sampling right from the beginning here, I'm really extremely pleased that it looks like we've now evolved to debate about which particular kind of probability sample we are going to be doing rather than doing some kind of another kind of a sample that's less scientific from my point of view.

So I think there's been a lot of progress and I think I can see coming to a reasonable conclusion based on the workshop. By the way, I should say I really appreciated all the presenters that gave. I thought it was really a great workshop and I feel like I learned a lot from the presenters.

So I really like probability sampling and I'm much more willing to accept essentially a probability sample platform or something that's as close as possible to a probability sample with the possibility that there's going to be some missing data. So some things are harder to collect than others so early trimester information may be very difficult so we may not get that for everybody. We may have to live with that. But if we're living with partial information but then still a probability sample, I think this is going to be a still a very useful study because there are lots of things that you can analyze that don't necessarily use that information.

In terms of the specific choices of a design, I think there are three overarching issues and actually the last panel knows more about this than I do so I sort of defer to the expertise to some degree. There's the choice of the frame, whether you use a provider frame or a hospital frame, the point of contact and the timing of the initial visit. Those seem to me to be the three key

issues.

In terms of the choices, I think a couple of people have said that devil is in the details and I think that's an important point. I think having very detailed specifications for these alternatives is important since arguing from 30,000 feet may not be all that useful.

The birth cohort versus siblings versus provider cohort argument should be based I think on the cost and that's been kind of - I got the view that there was quite a bit of divergence and sort of confusion about what the relative costs were in these different things.

Also the utility of the information for the hypotheses I tend to leave in actually trying to get the direct information for at least as many people as possible in the early pregnancy, since as I mentioned earlier, I don't think that multiple imputation is necessarily going to recover that information very well and the representativeness of the sample.

So there's been conflicting information about cost and practicality. I'm a little bit more inclined, based on what I heard today, to like the provider approach, provided it can be sort of operationalized properly, satisfactorily. And I really defer to people who have actually been in the field and doing this work so the folks who are current investigators I pay attention to the fact

that they are doing the real work here. So I tend to defer to that information.

So I think the provider sampling approach looks to me to be promising and I really liked the sort of unified way of thinking about the design that the last group was talking about. I'm not a big fan of a hybrid design and I think a unified design really works - we should be thinking about it the way that the last group was thinking about it, Colm and Graham.

So imputation of early pregnancy data, I said this earlier but it's important to bear in mind the fact that you're not making up information by imputation. You're using imputation to make use of the available information you have for the cases you are imputing.

And the value of imputation depends on whether that information is adding anything. So the only value I think for imputing early pregnancy data if you're interested in the relationship between those variables and the outcomes is if you have good auxiliary data available, either from proxy interviews or from some other source.

My final comment is the question about the original 110 PSUs versus a smaller number of PSUs. I would need quite a lot of persuading that the added variance from going to a more highly clustered design is really worth the savings and costs when you amortize it over the whole study

because the recruitment costs in terms of the overall cost of the study is clearly going to be a very small component. So I would need pretty strong argument as to why you'd want to go to a more clustered design.

DR. DUNCAN: Thank you, Rod. Our third speaker Ana Diez Roux from the University of Michigan.

DR. ROUX: Thank you. I also enjoyed the workshop a lot and learned a lot from all the speakers. Many of my comments are going to echo some of the things that the previous commentors said.

First, I think we need to acknowledge, obviously this study is trying to address a very complex and broad ranging issue. It's trying to do many different things and it includes many different disciplines and so it's normal that there's going to be discussion and debate. That being said, the study needs to move and I think you're all aware of that.

So I'm going to raise - we were asked to step back and think about big picture things so that's what I'm going to do. However, I'm not implying by this that addressing or thinking about these things should take 5 years. I think it's something that has been percolating and that can be done relatively quickly and should be done relatively quickly.

So the first point I want to comment a little bit

on is well, what criteria should be used to make design decisions. And then I'm going to talk a little bit - a couple comments on process and then few specific comments on things that came up during the day that I just wanted to point out.

So in terms of criteria that should be used to make design decisions, I think the study will need to grapple with prioritizing various study objectives. And be explicit about these priorities and recognize that there are tradeoffs, that there will be certain things that the main study will not be able to properly address and that is totally acceptable and fine and it just needs to be acknowledged.

I think that it's important for some of the design decisions and I think frankly it's also important for the morale of the study. Because I think when people are collecting data and having a sense that there are specific objectives that we are going after I think helps, at least in my experience working in groups, it helps push groups forward around a common idea.

Now if we think about the objectives that this study could have - again, we were asked to step back so that's what I did. There are a couple of - first of all there are two big sets of objectives. One objective which we haven't talked about much today but that could be

important is estimating incidents and prevalence of different conditions among US children.

The reason I've been thinking about this is because I was fortunate to be part of a panel, an IOM report that just was launched yesterday - and actually there was editorial in the New York Times today - about the US health disadvantage compared to other high income countries. And regrettably, health under age 50 and specifically among children and adolescents features prominently as one of the areas in which the US does substantially worse than other high income countries.

And one of the things that the panel found was that we don't have a lot of good data on the prevalence and incidents of many conditions among children in the US or even that we can compare to other high income countries. So this may be an objective that the study wants to think about as something that it could contribute that would be valuable.

The other big kind of study objective, which is the stuff I think that we've been talking about mostly today, has to do with etiological investigation. And within that one can kind of think about etiological investigation in two ways. It can be driven by very specific questions, very specific research questions. Now of course the disadvantage of this, and I think the study

experienced this a little bit, is that you can get bogged out in many, many, many hypotheses and it becomes completely unmanageable and overwhelming.

The other approach, the other extreme is to be completely agnostic and say okay, we're just going to collect data and then we're going to figure out what we're going to do with it. Now, I believe that a purely agnostic approach is virtually impossible because you have to make decisions and because in making those decisions there are implicit questions that you want to answer because that's what you're using to prioritize those decisions.

It is true that there are some aspects of the design, for example perhaps some aspects of the probability sampling as we heard from the prior panel, may be applicable to many, many different kinds of questions and that's great. But there will be a number of other decisions that have to be made that may require thinking about well, what are the priority objectives.

So is there a middle ground in these two extremes? And I agree with Rod, I think there is a middle ground and my sense is that that would be the most productive avenue for the study to take so what could a middle ground look like?

Well, one option is not to get bogged down in hypotheses, because hypotheses by definition have to be

very specific. Maybe we can think about well what's a typology of the priority questions that the study might answer. For example, are there one set of important questions about prenatal exposures. Then that would tell us, okay, so collecting prenatal information is really important and we need to maximize the design of the study and the instruments to do that. Are there certain kinds of environmental factors that we're especially interested in and it can't be everything? It can't be everything.

Some ancillary studies may do other stuff but what's the priority for the study. Is it environmental factors that are common? Is it environmental factors that we think could have very adverse impacts? And deciding what kinds of environmental factors are the ones that we're interested in will also help us decide some things. Are they environmental factors that vary geographically a lot? That would indicate that a geographically distributed sample is more appropriate.

Are there certain outcomes that - is it our typology of questions about certain kinds of outcomes. And again I'm not arguing for a list of very specific outcomes but a typology. Is it rare outcomes? Or is it common outcomes that are causing us to have much worse health than other high income countries but we don't really know why. I don't have an answer to that but I think that kind of

thinking might also help - is important for some of the design decisions. So is it outcomes that have public health impact? Is it outcomes that are rare but we want to learn about? Is it outcomes that contribute to our disadvantage with respect to other nations?

A third kind of question has to do with well, is the investigation of disparities key to this study? And that also has implications for how we sample and what data we collect. So I think having this kind of typology has implications for the core design, sample size and other issues and also for the core measures and for insuring that we have variability in the key exposures that we're interested in.

And this will require prioritizing and again, I don't think this is an exercise that should take a long time. I think the group will have to come to some consensus and not everybody will agree because that's the nature of humans. But some prioritization that can help guide some of these decisions because I think a lot of the discussions around the design and the measurement reflect underlying differences about what people feel the study should be addressing and so at least making these things explicit. If this is what we're going to address, this is what we have to do.

So, having a setting of core typology questions

that prioritize the study without being overly specific or detailed, and then in addition, of course the study needs to collect as much additional data as possible because we don't know. There will be many new questions that will emerge that we will want to answer and this should include exposure and outcome of these pre-disease markers, things that will allow us to look at epigenetics, all this new stuff. How should we prioritize that because we can't collect everything on everybody?

Well, some criteria have to do with the expected utility based on what we know. This is incomplete criteria because there are many things we don't know but that is one starting point. Ease of collection, some things are really easy to collect and so I say collect them. Getting GPS locations on the houses is very easy to do and that can allow linkage to a wealth of stuff down the line, as some of the panelists today indicated.

Storage - can we store the stuff? If we can store it and it looks like it might be interesting, I say try to get it, of course within logistics. So collect as much as possible of course within budget and logistical constraints and recognize that there will be ancillary studies that will do a lot of other stuff.

I think another thing that the study has to balance in making design data collection decisions is

simplicity versus complexity. And I would certainly weigh towards simplicity.

Simplicity has lots of advantages in terms of running the study on site, of analyzing the data later because remember that the more complicated things get, then the more difficult it's going to be for people to use these data. And even though we may be very sophisticated and able to do a lot of complicated stuff, there will be many people who will want to use this data who will not be able to do that - so simple but not so simple that it defeats the purpose.

Of course, I realize there are some things that need to be complex but if we can make things simpler, I say make them simpler. And this unified design approach that was talked about in the panel I think is very appealing because it is a simpler approach than having multiple cohorts that have to be weighed differently and combined. So I think those were the comments I had on general criteria for making decisions about the design of the study.

In terms of kind of the process of running the study and so forth, we haven't talked about this. But based on my experience in multisite studies I think it's important for the study to find the right balance between centralized and decentralized activities and decide well

which things really have to be centralized and there are many things that do have to be centralized and which things are better decentralized.

But I think being explicit about that and particularly capitalizing and learning from other multisite studies how they have done that and I'm sure you've done that to a certain extent but sometimes - I was on the Advisory Committee for a couple of years. Sometimes I got the sense that there was a wealth of information out there on how these large studies can work with a mixture of centralized and decentralized activity that perhaps wasn't being capitalized on as much. And certainly capitalizing on the experience of the Vanguard Centers, the investigators who were involved in the study as well as other investigators who have experience with these types of cohorts.

So those are my general comments. I'll just make two quick comments on some of the stuff that came up today. One has to do with this first birth issue and this may be moot now because the design that we heard in the prior panel doesn't necessarily - would bring in first births together with others.

Given what we know about differences in the biology of first pregnancies versus subsequent pregnancies and also birth order effects on a number of social and

health outcomes, I think it's very plausible that prenatal factors interact with birth order. So I would be hesitant to - I think we really need to look at that so I'd be worried if we didn't collect that information.

And the other I think interesting topic that we touched on but I'm not sure has been - well, we certainly haven't decided on - is this issue of siblings. I think the study needs to think about the advantages and disadvantages of including siblings.

Certainly the advantages would be the ability to get this preconception information potentially because there are logistic issues involved as well and the ease of recording sibs and perhaps some cost benefits and also the kinds of within family sib comparisons which can be very informative.

However, depending on what the priority questions are, enriching the sample with sibs may reduce variability in some exposures that are invariant, for example, within families and if there's clustering of outcomes within families that could have some power implications as well.

So I think the study needs to think about the tradeoffs of including sibs. And also I think the study needs to think about if the sample becomes more weighted with sibs, is it now no longer representative of a family structure and could that have implications of the US and

could that have implications for some of the inferences that can be drawn. Maybe for some it doesn't but maybe for some it does. That's it.

DR. DUNCAN: Thank you very much. So I detect agreement on some issues across the three panelists. I think everyone endorsed probability samples very heartily. The two of three that talked about the more PSUs versus fewer PSUs endorsed more PSUs.

I think the meatiest discussion was about hypotheses. Everyone thought there ought to be some version of hypotheses or Ana's conception these more general objectives, maybe not very explicit hypotheses.

I guess I would push that a little further to first thinking about the kind of objectives that have important bearing on the design. And if it's true that - Ed said he detected a consensus that preconception exposures didn't seem to be valued very highly by the group today. If that's the case and if it's also the case that exposures very early in pregnancy before we can really pick them up in a prenatal sample, if those really aren't that important then I start thinking very differently about a sibling sample and maybe even thinking that its value isn't worth it. But we need to get some kind of judgement about whether preconception and very early in pregnancy exposures are important questions that we just have to be able to

address with this study.

And the second element of this that Ana pointed out also that related to what Rod said is about geographically varying environmental exposures. If after thinking about what's potentially important, we really don't prioritize geographically variable environmental exposures to the point that we'd really want to sacrifice some sample efficiency to do the kind of oversampling scheme that Rod talked about, that's a fine decision but would follow from not prioritizing the geographically variable environmental exposures. It seems to me that needs to be a very conscious decision because it has very direct implications for what design looks like.

So we've got a little bit of time first for the panelists to react to what the other panelists said and then we'll open it up for questions.

DR. SONDIK: Let me react to the two points you made I completely agree with that I don't feel though that I'm in a position to prioritize the preconception. I don't know what the literature is. I don't know what the models might be, et cetera. But it strikes me that has been part of the study.

And I think it comes back to what hypotheses are associated with that, could be associated with that. And to what degree does the study have of shedding light on

those hypotheses.

I think that's a significant decision. I didn't mean to and I don't think anybody took what I said to say that that really was the consensus. I'm just saying it did not, no one turned red and had smoke coming out their ears from when it was discussed.

And I think the same on the geography. I think it's an important part of the study but I think having an expert look, I don't mean an expert per se, but I mean a good solid look at how a set of environmental factors distribute across the country, I think is important to do. And the question of whether or not the design has an ability to pick up that is a significant variable.

DR. DUNCAN: Okay, the floor is open. Michael, you're the first to the microphone.

DR. BRACKEN: Thank you. Well, I think it is extraordinarily encouraging that all four of you have immediately focused down on the need for hypotheses, not for dozens of them but I think what Rod called sentinel hypotheses. I mean picking out some that really reflect what this study could to, it's what the investigators have called in the past the need to show a bang for the buck. Communities need to be able to focus in on health effects to support this project. It's not enough just to be a data platform. So Ana was urging some speed in this.

Well there is a wealth of information that came out several years back now from, as Rod mentioned, numerous committees, hundreds of people were actually working on this in working groups trying to develop a hypotheses. And they are there, they are archived somewhere in the NCS.

And certainly I would think would be the first port of call to be revisited. Are these still actually the ones we're interested in? Do they need to be updated? And so on. But I would urge the NCS to now go back and look at that documentation because people spend hundreds, thousands of hours on it and they were the experts in the field in their various disciplines.

So that is there, it's a place where it could be done relatively speedily and I think it would be a real encouragement to people who did invest in that work and were very discouraged when it seemed to be abandoned later on.

DR. DUNCAN: Thank you. Nigel.

DR. PANETH: I'll just make three points I think. First of all, again I echo Michael in thanking the panel for emphasizing the need for prioritization, for systematically developing some kind of schema that would allow us to get out of the bind of not knowing what to prioritize. And I also agree with you that there's somewhere a sweetspot between enormous numbers of

hypotheses on the other hand and the current state of absolutely no hypotheses whatsoever on the other that would allow us to get at what the study is really about.

The second point is that within that world of hypotheses, because, and this is the wrong way to go about it, but that's what we're confronted with, because we have already said its 100,000. There are so many hypotheses that have no business being in a study of 100,000 per se. They couldn't possibly motivate it. I cannot imagine a hypothesis on obesity that needs 100,000 people. And then there are other things such as individual cancers which NCI has weighed on many years ago, simply even 100,000 would not get you there.

So there's the sweetspot of what are the hypotheses that truly motivate a study of this size and shape, both in terms of the prevalence of the outcome, the importance of the exposure, the importance of the relationship and so forth. So I think that kind of hard work that Ana has called for, it doesn't have to be large and resources have been pointed out by Michael and others, has to be done if a design that makes sense is to emerge.

The third point I'd like to make is you have spoken, others have spoken. Does anyone listen? Thank you.

DR. DUAN: Naihua Duan from Columbia University.

I would like to share a thought that was partly triggered by the discussions - a little quick note that I've never been affiliated with a study and for me this is a wonderful, marvelous educational experience. I appreciate the panelists and the discussions. I've really learned a lot. As I was sitting in my seat reflecting on what I learned, I kind of begin to wonder about the plan not to go for the household screening sample.

So I thought I might bring up some thoughts for the purpose of brainstorming. I guess one I think pretty strong message we learned from the first panel this morning is that the early pregnancy or maybe even preconception is a high priority and maybe even during the early part of the first trimester.

So some of that could be captured in the unified sampling approach, some of that might not be captured. So there is some question in my mind as where the relevant merits of this prenatal and/or household will accomplish relative to what potentially could be accomplished with a household sample.

I understand that the household sample has been found to be expensive but I think also we have developed some consensus during the discussions today that we should not just be looking at the recruitment cost. We should be looking at the total lifetime costs for the study and

several panelists who are probably more knowledgeable than I am commented that the difference in recruitment costs might not be that large when combined with follow up costs.

So that brings up some questions in my mind whether maybe some household sample is still useful to be retained to answer the important scientific questions. And I think there was some discussion previously about maybe taking a household sample in hotspots or what I will call the warm spots.

At the same time, another angle like what was mentioned is that the high cost of recruitment for the household sample is partly due to many women who could be recruited and followed for a long time without yielding a child. There is probably the potential to think about limiting the house sample to the women who are actively seeking pregnancy and so the relative costs might be somewhat lower.

So we had a lot of discussions about the sibling sample and also some questions about the sibling sample and I do agree with Irv and the other panelists that there are a lot of merit to consider for the sibling sample but there are also some limitations. This question about first born is not a trivial question that I think needs to be really considered carefully. And one advantage of the household sample if that can be retained in some affordable way,

would be to fulfill that gap.

Also the discussions made me realize, the question about the lack of what is not covered in a sibling sample about the earlier pregnancy exposure data, is not just the first born in a family that we will have later siblings, maybe to impute what the first born look like. There are more than a few families that have only one child so that is it so there are no subsequent second born to proxy for the first born.

So for the purpose of brainstorming this is a large study with a lot at stake. I will hope that we don't take the household sample entirely away from consideration. Maybe keep it - I understand this is probably a dead horse - but maybe keep it as a possible option and really evaluate whether there is some residual role that might still be a useful strategy to supplement the other strategies.

DR. DUNCAN: Any other comments from panelists? Our NICHD contingent, would you like to make any comments?

DR. HIRSCHFELD: I want to thank everyone for a very informative and stimulating discussion. We will continue the evolution of the design of the National Children's Study. I think everyone appreciates the potential and our goal is to have that potential not only met by the expectations that we can conceive of here but to

have a platform that would exceed, not necessarily because of anything we can predict but just because of its inherent nature, exceed our expectations and continue to surprise us in the future so it becomes an ongoing resource for informing us about the health and development and growth of children. Thank you.

DR. DUNCAN: Thank you, Steve.

DR. SONDIK: I would make a quick comment in response to the last comment over here about household. For me personally it would be very nice to see a comparison of the characteristics of the household sampling versus the provider in terms of data that can be collected through the provider - or data items, categories of data. And then for those data items which can be collected through which approach and that would provide a very nice basis for saying, well, you know, we really don't need to consider it or here's the potential for this, maybe in some circumstances.

The other thing with the household was we were looking at kids - we're saying that the population was not only the kids born in a particular time period but we were able to circumscribe the sample areas or the PSUs by the geographic area and so that always made me very comfortable if you know what I mean.

In other words, these were the kids that lived in

that particular area whereas here, and of course Graham didn't have time to go into all the details of this but I would expect that would be handled. So it would be very clear that we are getting a representative sample of the US kids when we look at the providers and hospitals on a geographic basis. It was always this complicated issue with the kids being born elsewhere for example and how that was handled in the household situation. So a nice comparison I think could be helpful. Thank you.

DR. KALTON: I would like to make an overview somewhat final comment. In putting the National Children's Study into the context of our society, and everyone in this room understands the value or the potential value, if not the need for the National Children's Study and what it's going to bring to us. And when we start looking at the elephant in the room that we've had through the day of cost, we've got to keep in mind that when we talk about costs we're talking about political will.

And if you step back with society and you look at the political will, how much money is NASA going to spend in the next 20 years to put a man on mars for intellectual curiosity and yet the future of our children of our population is to me exceedingly more important than that. So I just urge caution when we start talking about cost because political will will help us with that and my sense

is that both sides of the aisle understand this study fairly well enough that they support it. So I say, let's go forward and make it a good strong study and don't get caught up in nickel and diming. Make sure that it goes well and right.

DR. DUNCAN: Thank you. You have the official last words.

DR. MCLANAHAN: Thank you all.

(Whereupon, the meeting adjourned at 4:05 pm.)