

APPENDIX A
TECHNICAL WORK GROUP RECOMMENDATIONS

MEMORANDUM

TO: Mark Denbaly

FROM: Nancy Cole

DATE: 1/26/2010
FoodAps-002

SUBJECT: Summary of FoodAps TWG Meeting

This memorandum provides a summary of the first meeting of the Technical Working Group for the *National Household Food Acquisition and Purchase Survey* (FoodAps). TWG participants are listed below. Additional ERS staff observed and asked questions: David Smallwood, Mark Prell, John Kirlin, Betsey Frazao, and Ephraim Leibtag.

Participant List

From Mathematica

Nancy Cole
Laura Kalb
John Hall
Mary Kay Fox

From ERS

Mark Denbaly
Laurian Unnevehr

From academia

Steve Heeringa
Helen Jensen
Suzanne Murphy
Sarah Nusser
Roger Tourangeau
Parke Wilde

From government

Lori Borrud, NCHS
Donna Rhodes, ARS
John Eltinge, BLS
Susan Krebs-Smith, NCI
Kelly Kinnison, FNS

The original meeting agenda was modified to focus on data collection instruments rather than a general discussion of issues. The meeting was organized in five sections:

1. Overview of the Survey Design and General Q&A
2. Statistical Issues
3. Presentation of the Instruments for Collecting Food Data
LUNCH BREAK
4. TWG Member Discussion of the Food Instruments
5. Goals of the Field Test

The meeting began at 8:30 and concluded at about 3:30. Following the meeting, Mathematica and ERS discussed the main points of feedback.

The remainder of this memorandum documents the feedback from the TWG, and concludes with recommendations for next steps.

MEMO TO: Mark Denbaly
FROM: Nancy Cole
DATE: 1/26/2010
PAGE: 2

FEEDBACK FROM THE TECHNICAL WORKING GROUP

The primary points of feedback are listed first. Each point is discussed in detail in the following sections.

1. The instruments for collecting food data are too complicated. The TWG recommends simplifying the instruments. It was recognized that simplification will likely mean that the survey will not be able to collect some data elements
2. A review of research questions and priorities is needed, along with a mapping of data elements to research priorities. The results of this exercise can be used to guide the decisions about data elements that can be dropped or collected in a less detailed manner.
3. Plans for pre-testing and cognitive testing of instruments should be expanded.
4. The field test should be expanded to more than 200 completed interviews, with sequential testing of multiple data collection protocols and different incentives. This point was emphasized by multiple TWG members who considered expansion of the field test (coupled with expanded pre-testing and cognitive testing) to be critical to ensuring the survey's success.
5. Expand the field period for the full survey to at least 6 months. With a 4-month field period the survey will not be nationally representative. Expanding the field period will provide better coverage of the calendar year, mitigating potential concerns about seasonality. A longer field period will also provide more time to make adjustments in sample release to achieve the required number of completed interviews.
6. Three changes were suggested for the sampling plan: (a) consideration of a dual sampling frame for SNAP and non-SNAP households to reduce design effects; (b) construction of different sampling weights for different subsets of the sample defined by the level of response or survey completeness (multi-phase response adjustment); and (c) expansion of the field period (as noted above) to better manage sample release.
7. Some concern was expressed about the relative costs and benefits of Option 2.
8. Some concern was expressed about the ability to identify foods purchased with WIC vouchers.
9. Other comments. This final section of the memorandum summarizes other comments and questions from the TWG which we will address in our data collection plan.

MEMO TO: Mark Denbaly
FROM: Nancy Cole
DATE: 1/26/2010
PAGE: 3

1. Instruments for Collecting Food Data Are Too Complicated

Mathematica presented the scanner, the draft Barcode Book for collecting food at home (FAH), and the draft adult food diary for collecting food-away-from-home (FAFH). Both FAH and FAFH have two methods of reporting, and the TWG agreed with our overall four-part organization of the data collection:

1. Food at home, scannable
2. Food at home, not scannable
3. Food away from home, with receipts
4. Food away from home, without receipts

However, the TWG found the instrument package in its totality to be too complicated, largely due to the level of detail required by the SOW. The TWG also noted that the response burden will vary substantially depending on where households obtain food, the types of food obtained, and the number of people in the household. The TWG offered two suggestions for partially mitigating this burden differential: (a) collect less detail for events that pose the greatest reporting burden, and (b) provide an incentive to each member of the household.

The TWG's main suggestion for simplifying the instrument package is that we should not try to obtain full detail on every event, but accept less detail for the small percentage of events that pose the largest reporting burden – such as FAFH obtained from sources other than restaurants or fast food establishments and FAH that cannot be scanned. (Specific suggestions appear below.)

There also seemed to be a consensus that a combination of diary and telephone interviews should be used to collect information about FAFH because: (a) telephone follow-up allows for greater simplification of the paper instrument (regardless of whether respondents save receipts); and (b) telephone follow-ups will help maintain data quality throughout the data collection week. One TWG member noted that other surveys show a decline in quality of diary entries over time. If ERS concurs with this TWG consensus, then the main methodological test proposed for the field test—assisted versus unassisted diaries—can be replaced with a test of different diaries formats.

Specific Suggestions for Simplifying the Instruments

1. **Scanner and Instructions** – three types of simplification were suggested:

- Simplify the overall instructions. Change the cover page to include a picture of the scanner, and include a “Quick Start” guide at the front of the Barcode Book guided by pictures. Move detailed instructions to the back of the book (but don't sacrifice completeness). Different organizations of this book should be tested in the pre-test and

MEMO TO: Mark Denbaly
FROM: Nancy Cole
DATE: 1/26/2010
PAGE: 4

perhaps also in the field test. TWG members also recommended cognitive testing of the instructions, which we discuss in a later section on expending the pre-test.

- Instruct households to scan all foods and beverages that have a barcode, rather than telling them not to scan store-printed barcodes for variable weight items.
- Program the scanner so that it does not accept store-printed barcodes on variable weight items. (On follow-up, we have determined that this is not possible because store-printed barcodes follow the same 12-digit format as UPCs.¹ We provide information in Attachment 1 about the expected volume of variable weight items and changes in labeling standards.)

2. **Blue form** for collecting information about FAH “events” – two alternatives were suggested for re-configuring and simplifying this form:

- a) Keep the form basically as is, but reorganize it so that the mandatory fields are more prominent. Assume most FAH is purchased. Include “Free” as a response category for the “How did you pay?” question. Change the presentation of the “paid” section so that it does not appear optional. Replace “Who left the house to get food?” with “Who got the food?”²
- b) Reconfigure the blue form as a landscape grid of all FAH events for the week. This may require elimination of some items to get everything on the grid.

After presenting the FAFH diary, TWG members also suggested that FAH events could be incorporated on a “Daily List” page like the one presented in the diary, but modified to collect basic information about all food acquisitions (both FAH and FAFH):

- One suggestion was to include a single Daily List for the household. This list would remain in the home and each member of the household (or the main respondent, using diaries provided by household members) would add entries each day.
- Another suggestion was to include a comprehensive Daily List in each food diary so that each member of the household records all of their own food acquisitions in one place, including FAH or FAFH. Although at least one TWG member thought children are unlikely to acquire FAH and would not recommend this approach.

¹ Store-printed barcodes are used on meat, poultry, seafood, deli, and cheese items that are shipped in bulk and packaged and priced by the retailer. The barcodes on these items are 12-digit U.P.C. numbers. Because of its space constraint of 12-digit UPC, information is limited to the price of the product and a 4-digit commodity code. (“Industry Roadmap: Building the Fresh Foods Supply Chain of the Future,” www.iddba.org/pdfs/roadmap.pdf)

² ERS clarified that they would like information about who purchased groceries but do not need to know the identity of every household member who may have accompanied the shopper.

MEMO TO: Mark Denbaly
FROM: Nancy Cole
DATE: 1/26/2010
PAGE: 5

TWG members generally agreed that some form of a comprehensive Daily List should be used (with individual-level or household-level) to ensure that events are not missed due to confusion about which events should be recorded in which instrument. The details of the events would be recorded on blue forms (FAH) and diary pages (FAFH). With this approach, we will obtain the full picture of food acquisitions even if we don't get full detail on every event. And we can use the telephone interview to prompt respondents for details.

One TWG member asked if the Barcode Book could be replaced with a camera for taking pictures of items that could not be scanned. Another suggested that we review protocols developed by Molly Kretsch (ARS), which included a barcode book and scale to record items and weights. We believe that the addition of any tool (camera or scale) would further add to the complexity because we would give households another separate tool and instructions on when and how to use it. The tool developed by Kretsch is a computerized system with a scale and barcode reader linked to a portable computer.³

3. **Adult Food Diary** for collecting FAFH information – the TWG offered suggestions for organizing this instrument and for simplifying and reducing the amount of data collected:

- Modify the Daily List as discussed above.
- Modify the format so that each household member receives a small-format Food Diary that they can carry with them.
- Collect precise information to characterize each food acquisition event, which includes the location of the acquisition, amount paid, day of the week, and meal.
- Collect precise names of menu items for FAFH items purchased at chain restaurants and generic descriptions of meals for non-chain restaurants (this approach was included in our proposal).
- Accept less precision about meals obtained from non-restaurant sources. Identify meals/snacks involved, but don't collect information about specific foods or quantities.⁴

³ Kretsch, Mary and Fong, Alice. "Validation of a new computerized technique for quantitating individual dietary intake: the Nutrition Evaluation Scale System (NESSy) vs. the weighed food record," American Journal of Clinical Nutrition, 1990. "NESSy consists of a computerized food scale system (an electronic food balance and barcode reader interfaced to a portable computer with a built-in modem) and a bar-coded food identification catalogue (FIC). Menu-driven, user-friendly computer software prompts the research volunteer through a series of simple questions and commands enabling the user to record food weight and food identification automatically."

⁴ During the meeting we asked, "Is it enough to know that that a child got a school meal and not know exactly what was in the meal?" The response was "probably." In later conversations, there seemed to be some interest in

MEMO TO: Mark Denbaly
FROM: Nancy Cole
DATE: 1/26/2010
PAGE: 6

2. Review the Research Questions and Priorities Relative to Data Collection

The TWG suggested that Mathematica and ERS review the research objectives to reevaluate the level of detail needed in the collected data. This was a response to the complexity of the instruments for collecting food data.

This discussion began with the observation that the data collection is burdensome because it is comprehensive, and TWG members questioned the level of detail needed. “For example, if we want to know if people buy fruits and vegetables versus McDonald’s, we don’t need to know what they buy at McDonalds.”

One TWG member asked why information is needed about all household members on all days of the week, for both FAH and FAFH. He suggested that perhaps we should sample people within households, sample days of the week, and/or sample shopping events. He also suggested that we should not collect information about vending machines because these purchases are such small dollar amounts. These suggestions were rejected by other TWG members who asserted that information is needed on a household level for both FAH and FAFH because prior surveys have focused on either FAH or FAFH and there are no data to tell us how food from these two sources fit together for a household. In addition, while vending machine purchases may be a small percentage of food expenditures, they may make important contributions to available amounts of calories, saturated fat, and/or added sugars.

The general consensus among researchers who have worked with comparable data on food expenditures and nutrient values was that we should not sacrifice the comprehensiveness of the data collected for major food acquisition events (i.e., purchased FAH). Rather, we should sacrifice the level of detail obtained for the less common events (free FAH, with the possible exception of food obtained from food pantries and FAFH obtained from non-restaurant sources, with the possible exception of school meals.)⁵

(continued)

obtaining more information about school meals, but a general consensus that the details were not necessary for meals and snacks eaten at friend’s or relative’s homes, in afterschool programs, or other non-restaurant settings.

⁵ The discussion about which food acquisition events would be considered “major” and “not major” was not exhaustive and no group consensus was reached. The above is what Mathematica staff took away from the discussion, but it is not definitive. For example, if we take the approach of collecting less detail on non-restaurant FAFH, we will have to decide if food from a convenience store, coffee truck, or street vendor, would be considered non-restaurant foods.

MEMO TO: Mark Denbaly
FROM: Nancy Cole
DATE: 1/26/2010
PAGE: 7

Research priorities. ERS confirmed that the research questions on page 5 of the SOW are ordered by their research priorities. These are:

1. How do price and income influence food choices and the dietary quality of food purchases of Americans across all income levels?
2. What do SNAP participants buy and how much does it cost?
3. How does food assistance program participation influence food purchases and acquisitions?
4. What is the relationship between food acquisition decisions and levels of food security?
5. How do access and retail outlet choice influence food purchases and the resulting dietary quality of purchases?
6. What is the influence of nutrition knowledge and attitudes on food purchases?

The first research question requires data on prices, quantities, and the dietary quality of acquired foods, plus information on household income. Remaining research questions relate characteristics of food acquisitions or food purchases to other household characteristics. However, food acquisitions may be characterized in a number of ways. For example, “what SNAP participants buy” (research question 2) can be measured by the source of food, food group analysis, broad measures of dietary quality, or specific nutrient characteristics of acquired foods. Food acquisition decisions (question 4) might be characterized by the number and timing of shopping trips, the source of food, or the characteristics of acquired foods.

Mathematica will prepare a matrix for inclusion in the data collection plan to map research questions to the data items specified in the SOW. It might be useful for ERS to independently compile a similar matrix to prioritize the methods by which ERS would like to see these questions address and the data items needed to implement those methods.

3. Expand Plans for Pre-Tests and Cognitive Testing

The current schedule includes a single pre-test of all instruments with nine households over a 7-day period in April-May 2010 after the instruments have been reviewed by ERS and the TWG. The schedule includes the following due dates: Draft instruments-January 27; Revised draft instruments-March 17; Second revised draft instruments (the versions to be used in the full pre-test)-April 23; and Final instruments (revised after the pre-test)- May 21.

MEMO TO: Mark Denbaly
FROM: Nancy Cole
DATE: 1/26/2010
PAGE: 8

The TWG was in agreement that plans for instrument testing must be expanded because the instruments and protocols for collecting food data are new and untested. They strongly recommend that questions about instrument design should be largely resolved at this stage through multiple pre-tests. The TWG suggested iterative sequential cognitive testing of the food instruments prior to pre-testing all instruments with nine households over a 7-day period.⁶

For cognitive testing, the TWG suggested that we test each food instrument in isolation to determine respondents' level of understanding, ability to adhere to protocols, and burden. For example, we can test the FAH protocols in isolation by recruiting a household, instructing them on use of the scanner and barcode book, and observing their ability to correctly scan grocery items. The food diary would be tested on a similar small scale with up to 9 households, but apart from other data collection instruments. This provides opportunities to test, revise, and re-test before a full scale pre-test of all instruments over a 7-day period. TWG members from BLS and NCHS indicated that multiple pre-tests with 9 respondents is permitted by OMB as instruments are revised.

4. Expand the Field Test Sample and Sequentially Test Alternative Designs and Incentives

As discussed above, the TWG offered several suggestions for simplifying the food instruments, but they noted that there is no research to guide the instrument development. Thus they consider it critical to test alternative instrument designs on a fairly large scale prior to fielding the full survey. Similarly, the TWG concurred that there is little data to guide us with respect to the correct incentive amount. Incentives should be tested because they have impacts on response rates, but cannot be so big as to be coercive.

The current contract includes a field test of all survey instruments and procedures during an 8-week period from January-March 2011 to obtain survey completes with 200 households. Mathematica proposed to test alternate methods of collecting FAFH data during the field test: (a) *assisted food diaries* versus (b) *unassisted food diaries*.⁷ The TWG however, was in agreement that the food diaries should be fielded with telephone interviews because this allows greater simplification of the paper diaries and telephone contact will contribute to better data quality throughout the data collection week. Thus, the planned methodological test can and should be replaced. Instead of testing diary-alone vs. diary + telephone methods, the TWG

⁶ Cognitive testing could be done between the March 17 and April 23 deliverables, if resources are available.

⁷ We originally proposed a sample of 400 completed interviews and an additional test of two alternate incentive schemes. ERS instructed all bidders to cost a field test of 200 completes, and this led us to drop the second test of incentive schemes.

MEMO TO: Mark Denbaly
FROM: Nancy Cole
DATE: 1/26/2010
PAGE: 9

suggested that we can test different version of the diary to be used in conjunction with the telephone interview.

The TWG strongly recommended an expansion of the field test to include a larger sample size and multiple tests of alternative designs: to test the food instruments and the incentive. As discussed above, the TWG offered several suggestions for simplifying the food instruments, but they noted that there is no research to guide the instrument development. Thus, it is critical to test alternative instrument designs on a large scale prior to fielding the full survey.

One suggestion is to develop four alternative instrument designs and field test A and B in a first round, and C and D in a second round. Alternatively, we could develop three alternatives and test A versus B, then C versus the better performer among A versus B (this assumes that we obtain information needed to judge A versus B from the telephone interviewers, which can be analyzed in near real time). These alternatives could include different designs for diaries used by all household members, or a single joint household diary that is monitored by the gatekeeper.

Iterative testing of alternative designs in the field test requires complete development of alternatives in advance, for submission to OMB. Sequential testing also requires a change in the schedule for the field test without jeopardizing the start of the full survey in March 2012.⁸

Another suggestion for the field test (also discussed under Sample Design for the full survey), is to use a nested sample similar to NHANES, whereby event data would be collected from a full sample of households, but details about food items would be collected from a subsample (similar to the nesting of the NHANES examination sample within the interview sample). This would provide a validation for the reporting of events. It was noted, however, that roughly equal sample sizes would be needed and the sample would be better used for other tests.

Incentives. There is little evidence to guide the correct incentive amount or incentive design because FoodAps will impose a large burden and the survey requires participation of all household members.⁹ The TWG suggested the following alternatives for consideration:

- Fixed amount

⁸ The single 8-week field test period could be replaced with two 6-week periods, with the schedule compressed elsewhere to meet the deadline for OMB submission for the full survey. However, any schedule adjustments following the field test will reduce the amount of time available to compile and review field test results and use those results to inform the design of the full survey.

⁹ One TWG member suggested that information about incentives may be available from grant-funded surveys that have more flexibility in providing incentives because they do not need to obtain OMB approval.

MEMO TO: Mark Denbaly
FROM: Nancy Cole
DATE: 1/26/2010
PAGE: 10

- Escalating amount tied to telephone reporting of FAFH
- Different incentive amounts according to household size (to reflect differential burden)
- Incentives for each member of the household, to encourage response on the Food Diary and compensate for individual burden
- Set incentive amount plus a bonus for completing everything, but present it to respondents as a loss. For example, they can receive \$100 for the week, but if they fail to complete a telephone interview, they lose \$20.

The TWG noted that incentive schemes that vary from a fixed incentive can be designed to provide \$100 on average, but information from the field test would be needed to set the amounts to achieve this average. Also, varying (or pro-rated) incentives should reflect the value of the data, considering at what point a partial response is worthless.

One constraint in varying from fixed amounts is that IRBs may consider it unethical, unless varying incentives are justified by differential burden. For example, an IRB may object to providing a larger incentive to higher-income households on the basis of the market value of their time, but would not object to larger incentives for larger households who have a larger reporting burden.

TWG members also commented on OMB expectations regarding incentives. A cost-benefit assessment should be used to justify incentive amounts. Thus, incentives should be linked to cost savings in recruiting households, or data loss from nonresponse. Varying incentives should be provided on a pro-rated basis, not all-or-nothing.

Tests of different incentives can be combined with tests of different instrument designs. TWG members noted that we will not lose power by “cutting the sample” for different tests unless we expect interactions between the tests. If the design of the instrument affects respondents independently of the incentive amount, we will not lose power.

5. Expand the Field Period for the Full Survey to 6 Months

As noted above, TWG members recommended a longer field period for two reasons: (a) to provide more time for making adjustments in sample release to achieve the required number of completed interviews, and (b) to provide better coverage of seasonality.¹⁰

¹⁰ The TWG noted that ideally, a survey designed to provide nationally representative estimates should operate for 12 months to represent all seasons. However, they also noted that that it is difficult and costly to run a field operation in the winter months because it is not safe for field interviewers to knock on doors after dark. There was some discussion of whether the November to January holidays should be included because this period is

MEMO TO: Mark Denbaly
FROM: Nancy Cole
DATE: 1/26/2010
PAGE: 11

Mathematica will need to work through our budget assumptions to determine the cost implications of an expanded field period. The field costs are driven by:

- a) The fixed cost of hiring and training a fixed number of field interviewers – which must be large enough to cover the number of SSUs.
- b) The marginal interviewing cost, which depends on the number of completes and is not influenced by the field period.
- c) The cost of managing the field effort, which is determined by the supervisor/interviewer ratio and the period of time.
- d) The cost of obtaining Nielsen price data, which is priced per PSU and month, and will increase with a longer field period.

During initial negotiations, we suggested that savings could be obtained for Option 1 by extending the field period. With a four-month field period, we need to hire additional interviewers to cover the full added sample of Option 1. With a longer field period, we can hire additional interviewers to cover the additional SSUs added for Option 1, and assign remaining Option 1 sample to the Base contract field staff. (If only the Base is fielded, a longer field period does not provide savings because we cannot reduce the number of field interviewers below the number needed to cover all SSUs.)

6. Sample Design

The TWG offered several comments and questions about the sample design. TWG members concurred with three suggestions.

Comments and questions:

- **Number of PSUs** – There was some concern that 50 PSUs seemed thin for a national household survey. That concern, however, was alleviated after John Hall explained that the design for the base contract focuses on two separate samples of 1,500 SNAP households and 1,500 non-SNAP households, rather than on a combined sample of 3,000. It was suggested that perhaps the pilot can provide more information to optimize the sample design (including the numbers of PSUs).

(continued)

characterized by increased food acquisitions. ERS staff reported that they decided against including this period before issuing the SOW.

MEMO TO: Mark Denbaly
FROM: Nancy Cole
DATE: 1/26/2010
PAGE: 12

- **Will there be certainty PSUs?** There could be, but it depends on how we define a PSU. For example, if New York City is defined as one PSU, it will have a good chance of being included with certainty; but if each borough of NYC is its own PSU none are probably large enough to include with certainty.
- **Will there be too much clustering with four SSUs per PSU?** It is not possible to know this now. The sampling design provided reasonable assumptions of intracluster correlation. We will obtain estimates of intracluster correlation from the field test to use in re-evaluating the optimal design at the SSU level. The TWG agreed that, because of the relatively small sample size, and the relative cost of PSUs and SSUs, the current design should not be changed at this time. It was also suggested that we look at the design for the National Election Study because it has a similar sample size.¹¹
- **Did we consider subsampling?** A TWG member asked if we considered structuring the data collection similar to NHANES, whereby a large sample would provide information about all food acquisition events, and a smaller subsample would provide full details about events and food items. This reduces the complexity of data collection for part of the sample, but requires a larger overall sample. Others noted the limitations of this approach:
 - Subsampling was suggested in part because of the opinion that “some households will only be able to do the Daily List, some will save receipts, and some will also scan.” The subsampling, however, cannot be voluntary because that would create bias. Therefore, this approach does not address the concern that some households will not be capable of completing the data collection.
 - ERS explained that this approach would not provide adequate data for analyzing price and income elasticities of demand.
- **Did we consider a replicate week for validation?** A replicate week of data collection was proposed during which data would be collected only about events. This replicate week would provide a check on the validity of the primary data collection week by

¹¹ The American National Election Study (ANES) is not a field effort comparable to FoodAps. In 2008-2009, the ANES was a series of surveys of a representative sample of the US population of households with a landline telephone. The study included 12,000 individuals recruited by telephone to participate in monthly surveys over the internet. In earlier years the survey was conducted by telephone (2004) and/or in person: the 2000 sample was split with half by telephone and half in person for a total sample of 1,807 pre-election and 1,555 post-election.

MEMO TO: Mark Denbaly
FROM: Nancy Cole
DATE: 1/26/2010
PAGE: 13

providing data to compare the average number of events reported in the primary and replicate weeks for the full sample and relevant subsamples. One TWG member noted that this approach may be limited depending on the relative variability of week-to-week food acquisitions relative to household-to-household variability.¹² In addition, this would divert resources from other activities because the replicate and primary samples would require roughly equal sample sizes.

Suggested changes to consider:

- A. **Dual Sample Frame at the SSU level:** Mathematica proposed a single sample frame at the first and second stage with a composite measure of size (MOS). The composite MOS would be based on the counts of SNAP households and counts of low-income households (based on Census estimates net of SNAP counts). The composite MOS will result in samples of SNAP households and of non-SNAP households that have approximately equal probabilities of selection (within sample) but may result in an uneven distribution of the samples across SSUs. One TWG member suggested that equal sample sizes across SSUs was desirable for some types of analysis.

A TWG member suggested a dual sample frame, such that: (a) a composite MOS is used for the first stage selection of PSUs, and (b) two separate MOS's are used for selection of SSUs at the second stage. Therefore, in each PSU, we use a separate SNAP and non-SNAP MOS to select two SNAP SSUs and two non-SNAP SSUs. This should result in more equal sample sizes across SSUs, but may result in higher design effects due to clustering since each sample would be concentrated in fewer SSUs.¹³

- B. **Release of sample:** Mathematica plans to release sample within each SSU on a phased basis. Adjustments to sample release (if the screening, recruitment, or response rates vary from our assumptions) would be made globally across all SSUs. Our example showed 70 percent of the sample released in the first wave. This percentage is high because of the short field period. We could, however, be at risk if we release too much sample at once and do not have enough time for correction. The TWG was in agreement that sample release should be more gradual over a longer field period to allow more room for correction. "The current field

¹² ERS suggested that they can examine week-to-week versus household-to-household variability using Homescan data.

¹³ One TWG member suggested a Chris Skinner reference on the design effects associated with a composite MOS. We identified this to be Skinner, CJ, Holmes, DJ, and Holt, D. "Multiple Frame Sampling for Multivariate Stratification," *International Statistical Review*, 1994: 62.

MEMO TO: Mark Denbaly
FROM: Nancy Cole
DATE: 1/26/2010
PAGE: 14

period is so short that the risk of failure is high and the probability of successful correction is low.”

The TWG also noted the disadvantages of a field period longer than 6 months:

- Avoid winter because heating costs compete with food for the household food budget
- Avoid winter because field costs rise when there is less daylight

ERS noted that one advantage of a short field period is that we can look cross-sectionally across households without a lot of variation. The TWG noted that with a longer field period, we will have smaller samples in any one month or season for univariate analyses, but the seasonality can be used in regression analyses.

C. **Nonresponse analysis:** Mathematica discussed our approach for analyzing nonresponse bias and asked for suggestions for integrating the data items needed for nonresponse analysis in our survey. This elicited the following suggestions:

- Look at nonresponse in other diary surveys to learn about the characteristics of responders and nonresponders.
- Look at the data items collected for analysis of nonresponse in the *National Survey of Family Growth* and the *Time Use Survey*.
- Include a short set of questions for households that refuse to participate. These should be general questions about household characteristics and food acquisition patterns; and field interviewers should be instructed to obtain basic household composition from neighbors if the sampled household does not respond.¹⁴
- Ask about “types of jobs” not just employment per se, because the type of job may affect the ability to complete the record keeping tasks for this survey.
- Evaluate item nonresponse by household composition. The reporting by the household head may be different from response of other household members.
- Examine whether characteristics of field interviewers are related to nonresponse.
- Have the field interviewers collect environmental data that may be related to response, such as weather during the data collection week.

¹⁴ In response to this suggestion, we have included a “Short Form for Refusals” in our draft Data Collection Instruments.

MEMO TO: Mark Denbaly
FROM: Nancy Cole
DATE: 1/26/2010
PAGE: 15

- Make sure the data items needed for analyzing nonresponse are in the household screener (for analyzing refusals vs. participants) and first household interview (for analyzing attrition).

The TWG suggested multi-phase nonresponse adjustment to retain all partial completes in the final files and assign different sampling weights to different subsets of the sample according to where they dropped out of the data collection. This differs from the less costly proposed approach which retains completes that meet a single definition of survey completion. The proposed approach is less costly because the weights are produced as a by-product of the sampling procedures, whereas a multi-phase nonresponse adjustment requires additional post-stratification and trimming.

7. Option 2: Stratify the Low-Income Sample

Mathematica asked for TWG input about the costs and benefits of conducting Option 2, relative to reallocating those resources for additional testing prior to the full survey. Option 2 will add 500 completed interviews to the non-SNAP low-income sample and stratify the sample to obtain 800 completes with households under 100% of poverty and 1200 completes with households between 100 and 185% of poverty.

We received the following comments and questions:

- **What is the income distribution of the SNAP population?** Option 2 includes the very poor (<100% of poverty) in addition to low-income households up to 185% of poverty. For purposes of comparing SNAP households with income-eligible nonparticipants, we will have two parts of the income distribution: <100% and 100-130%. The low-end of the distribution will be obtained at very high cost, but it is not clear if Option 2 will yield enough sample to compare income-eligible nonparticipants at the high end of the income-eligible distribution (100-130%).
- **How will the data be used?** Researchers will not be able to properly evaluate differences in food acquisitions of SNAP and non-SNAP households with the planned data collection because it does not provide sufficient information for modeling the participation decision. Thus all analyses are subject to selection bias.
- **Feasibility** – It was suggested that we consider the size of the sample of households with income between 100-130% of poverty expected from the base contract. These households may be significantly easier to find (less costly) and would provide a feasible comparison group for SNAP households.

MEMO TO: Mark Denbaly
FROM: Nancy Cole
DATE: 1/26/2010
PAGE: 16

8. WIC Foods

TWG members asked about how WIC food acquisitions will be identified. The instruments presented at the TWG meeting included identification of events when WIC vouchers are used, the total amount of the WIC purchase, and food categories for which vouchers were used. The instruments do not, however, provide a way for respondents to indicate the exact items purchased with WIC vouchers because we considered that task too burdensome.

In our Data Collection Plan, we will discuss the methods by which Mathematica will identify precise WIC foods in the collected FAH data. The methods that we propose are similar to methods used for the *WIC Cost Containment Study* that Abt Associates completed for ERS.¹⁵ In that study, scanner data were collected from a sample of supermarkets for all transactions that included WIC items. Transactions involving multiple tender types did not include precise identification of WIC items. Abt Associates identified WIC items as the combination of WIC-eligible items that summed to the total WIC tender amount. For FoodAps, we will similarly identify WIC-eligible items using Federal regulations and State WIC food lists, and identify the purchased WIC items as those items that sum to the total WIC amount tendered.

9. Other Comments and Questions

These additional comments and questions concern the data collection plan, rather than the instruments and overall study design. These TWG comment will be addressed in our Data Collection Plan.

How will SNAP addresses be located within Census block groups? The TWG would like to see the specification of geocoding methods and handling of non-geocodable addresses such as PO Boxes. The TWG member noted that there is research showing evidence of geocoding errors (incorrect latitude and longitude) from geocoding services which result in valid addresses associated with the wrong Census block.

What methods will be used to mask household addresses for the public use files? The TWG asked whether we will be able to ensure confidentiality if we release household ZIP codes, names of places where they obtain food, and distances to places where they obtain food?

¹⁵ Kirlin, J., Cole, N, and Logan, C. "Assessment of WIC Cost Containment Practices: Final Report," USDA, Economic Research Service, February 2003 (E-FAN-03-005).

MEMO TO: Mark Denbaly
FROM: Nancy Cole
DATE: 1/26/2010
PAGE: 17

Ongoing communication with the Bureau of Labor Statistics – John Eltinge from BLS indicated that his agency is examining redesign options for the Consumer Expenditure Survey and would maintain communications with ERS about respondent burden, response rates, and completion rates.

Food that goes out of the home – a question was asked about how we will capture food that “leaves the household” because it is consumed by guests. During the final household interview after the data collection week, we proposed to ask whether the household had dinner guests during the week, on what day, and how many guests. A follow-up question addressed large FAH purchases during the week for events that occur after the data collection period: for example, a large shopping trip on Thursday for a birthday party on Saturday (after the data collection week). We currently have no plans to ask about events outside of the data collection week because it increases burden on all responders to capture few events. We will revisit this decision after assessing the burden for the final household visit.

Higher-income households – Lori Borrud from NCHS said that experience from NHANES showed that higher-income households welcome materials that inform them about how the survey data will be used. NHANES also provides households with a report on how their data compares to others. In addition she recommends development of a website to provide detailed information about the survey to households who want to seek it out. This website gives the study credibility among low-income and higher-income households.

Publicizing the study –The survey should be publicized in the PSUs through advertisements to make people aware of the survey and lend it credibility before Mathematica staff knock on doors. In addition, we should notify the local police, churches, and other community groups. NHANES provides a link on their website “if you are a survey participant” to add credibility.

ALERT data – Mathematica proposed to use ALERT data to analyze nonresponse amount SNAP households. The TWG also suggested that the ALERT data could provide measures of shopping patterns for the neighborhood.

Certificate of Confidentiality – This certificate goes beyond the standard consent form to inform SNAP households that we will obtain administrative data but the data will not be used to investigate fraud.

MEMO TO: Mark Denbaly
FROM: Nancy Cole
DATE: 1/26/2010
PAGE: 18

PROPOSED NEXT STEPS

The TWG offered three recommendations that will require a significant change in the study design:

1. Increase resources for pre-tests and cognitive testing.
2. Increase the size of the field test. Include tests of multiple food instrument designs and multiple incentive schemes. Do not test an unassisted diary, because the telephone interviews are considered critical to maintaining data quality throughout the data collection week.
3. Extend the field period to at least six months, to provide better control over sample release and better coverage of seasonality.

As noted during the TWG meeting, these changes require additional funding or a reallocation of the existing budget. ERS input is needed before we proceed to incorporate these changes in our data collection plan, or to evaluate budgetary implications.

Our current plan is to deliver the Data Collection Instruments on January 27, in accordance with the current schedule. All instruments will be delivered except the food instruments. The food instruments were provided to ERS in draft form at the TWG meeting, and will be revised after we reach agreement with ERS on which revisions will be made. Before beginning the revisions, we think it is advisable to have a meeting or conference call with ERS staff in which we review the summary of TWG recommendations presented in this memo and the reaction of ERS staff. For example, ERS staff may have “heard” some comments differently than we did and/or may place more weight on some comments than others. We believe that the level of TWG feedback and the potential implications for instrumentation merits a careful and coordinated review by ERS and Mathematica staff. After we reach joint decisions about the specific directions to take in revising the draft instruments, we will be happy to develop multiple versions (for use in expanded pre-testing), should funds become available. ,

Mathematica will deliver, separately, a budget estimate for an extended field period of 6 months, after receiving a request from the ERS contracting officer. Please let us know whether to prepare the budget for the Base contract or the Base plus Option 1. This budget must be prepared by Laura Kalb, our survey director, who is conducting a training for another study through the end of the month. Thus mid-February is the earliest that we can deliver a budget estimate.

MEMO TO: Mark Denbaly
FROM: Nancy Cole
DATE: 1/26/2010
PAGE: 19

Attachment 1 – Sales of Variable Weight Items

Mathematica proposed to obtain the weights of variable weight items from receipts, or by imputation. We continue to develop plans for collecting item weights without burdening respondents and to evaluate the percentage of variable weight food acquisitions for which we can collect data on weight.

Supermarket sales comprise over 80 percent of SNAP benefit redemption. Supermarket sales in departments that include variable weight items represent the following percentages of supermarket food sales:¹⁶

- Produce – 12.3%
- Service deli – 3.9%
- Self-serve deli – 1.5% (includes pre-packaged items with UPCs)
- In-store bakery – 2.3%
- Meat, fish, and poultry – 15.6% (includes pre-packaged items with UPCs)

Produce should not pose a challenge for this study when purchased at supermarkets because item weights are printed on the receipt when the item is weighted at the point of sale (POS) and the scale is integrated with the POS system. However, we can expect that produce weights will not be printed on receipts at small retailers who lack an integrated system or at farmer's markets/produce stands.

Weights for other variable weight items are not likely to be printed on the receipt because these items are weighed before barcodes are attached to the item (they are not weighted at the POS). Service deli and in-store bakery items represent small percentages of food sales, but meat, fish, and poultry sales are substantial. The weights of these items are not printed on receipts. We will inquire with Nielsen about whether price per pound information is available in the data that they collect from retailers.

Emerging barcode symbology: We will monitor the implementation of the new Databar barcode symbology.¹⁷ Information about the Databar was sent to the ERS separately. This new industry standard is being rolled out in 2010 for fresh produce, with other food categories (deli, meat) to follow, and full rollout is expected by 2014. Databar (formerly known as RSS, or

¹⁶ Food Marketing Institute, Supermarket Sales by Department, 2007. Adjusted for non-food sales. ERS may be able to obtain better estimates of food expenditures, not limited to supermarkets, using Homescan data.

¹⁷ GS1, Industry Roadmap: Building the Fresh Foods Supply Chain of the Future," available at www.iddba.org/pdfs/roadmap.pdf.

MEMO TO: Mark Denbaly
FROM: Nancy Cole
DATE: 1/26/2010
PAGE: 20

Reduced Space Symbology) was developed to provide standardization and traceability for two types of fresh foods:

- Loose variable measure items – product that is sold loose and can be bagged by the consumer
- Pre-packed variable measure items – product that is packaged by the supplier, vendor or retailer

Because of its small size, Databar provides a scannable barcode that can be printed on PLU stickers for loose fresh produce. This implementation of Databar provides a standard scannable code for identifying produce items. The scanner that Mathematica proposes for this study, and demonstrated at the TWG meeting, is compatible with Databar. For pre-packed items, Databar will be implemented in its expanded form, to include information about product name, price, and item weight.

We can expect that Databar barcodes will be used in some stores and some markets during every phase of our data collection, and we will continue to monitor its rollout.