

JUL 02 1997

0074-A-01

A98.2:W 84/7

HSR

Health Systems Research, Inc.
2021 L Street NW Suite 400
Washington DC 20036
(202) 828.5100
Fax: (202) 728.9469

The Feasibility of Analyzing WIC Food Purchasing Patterns Using Combined Issuance and Transaction Data:

Feasibility Volume

Final Report

Prepared for:

U.S. Department of Agriculture
Food and Consumer Service
Alexandria, VA

Prepared by:

Loren Bell, Lawrence Bartlett, Renee Schwalberg,
Katherine Clayton, Leanne Belongie, Hilary Bellamy and Renee Simon
Health Systems Research, Inc.
Washington, DC

April 1997

97-027600

e



Health Systems Research, Inc.
2021 L Street NW Suite 400
Washington DC 20036
(202) 828.5100
Fax: (202) 728.9469

FINAL REPORT

SUBMITTED BY:

Health Systems Research, Inc.
2021 L Street, N.W., Suite 400
Washington, DC 20036

CONTRACT TITLE:

WIC Food Purchasing Study
Contract No. FCS: 53-3198-4-024

FCS PROJECT OFFICER:

U.S. Department of Agriculture
Food and Consumer Service
Office of Analysis and Evaluation, Room 208
Attn: Marie Mitchell, Contracting Officer's
Representative
3101 Park Center Drive, 2nd Floor
Alexandria, VA 22302

CONTRACT OFFICER:

U.S. Department of Agriculture
Food and Consumer Service
Contract Management Branch, Room 914
Attn: Karen Scott, Contract Specialist
3101 Park Center Drive
Alexandria, VA 22302

HSR PROJECT DIRECTOR:

Loren Bell
Senior Associate
Health Systems Research, Inc.
2021 L Street, N.W., Suite 400
Washington, DC 20036

Table of Contents

Volume I. Feasibility Study

Chapter I.	Introduction and Overview	1
	A. Background	1
	B. Purpose of the Study	6
	C. Potential Uses	7
	D. Overview of the Methodology	11
	E. Feasibility Questions	12
Chapter II.	Feasibility of Using Scanner Technology to Examine WIC Participant Food Purchasing Patterns	15
	A. Introduction and Overview	15
	B. Major Findings Related to the Feasibility of Matching WIC State Agency Data With Store Transaction Data	16
Appendix A.	Description of Merge Data Set	

CHAPTER I

Introduction and Overview

This chapter provides a brief overview of the Special Supplemental Nutrition Program for Women, Infants and Children (WIC), highlighting the need for more data about the food purchasing patterns of WIC participants that led to the implementation of the feasibility study described in this report. Potential uses for data on the food purchasing patterns of WIC participants are described. A brief overview of the methodology of linking supermarket point-of-sale data with information from the WIC agency database is presented. Feasibility questions encountered in implementing this methodology are outlined.

A. Background

1. The WIC Program

The WIC Program was established in 1972 through an amendment to the Child Nutrition Act. Its purpose is to provide low-income pregnant, breastfeeding, and postpartum women, infants, and children up to age five with supplemental foods, nutrition education, and health care referrals to counteract the adverse affects of poverty on their nutrition and health status.

There are four methods used by States to deliver WIC Program benefits to participants. The majority of WIC State and local agencies have relied primarily on the distribution of a "food instrument (FI)," which can be in the form of a check or a voucher. Participants are given the FI at their local WIC clinic for the purpose of obtaining WIC-approved foods. In a State using checks as food instruments, the check is accepted by a store in exchange for WIC-authorized foods, and then deposited in the store's bank. The check clears through the Federal Reserve system, and like other checks, is paid from the bank account managed by the State WIC

agency. If a State uses a voucher, the store must send the voucher to the State or its fiscal agent for payment. In either case, the FI is exchanged for food at a retail grocery store.

However, some States operate different types of systems. For example, six counties in Wyoming are testing replacement of their check system with an electronic benefits transfer (EBT) system that requires participants to use a "smart card," a card containing a computer chip, to purchase WIC foods. Mississippi distributes food directly to participants through State-operated warehouses. Vermont uses a home delivery service to provide WIC participants with foods, while Ohio uses both a retail check system and a home delivery system (in 32 of 88 counties). Alaska uses a direct delivery system in areas of the State where retail grocers are not available.

States that operate systems using a food instrument are known as "retail delivery States." These States issue WIC FIs to participants who qualify for program benefits. WIC participants use their FIs to purchase specific foods at authorized grocery stores. Relying on guidelines and regulations established by the United States Department of Agriculture's (USDA) Food and Consumer Service (FCS), each State develops a list of approved WIC foods from which participants may choose. Each of the general food categories and the amount of food authorized to be purchased is clearly displayed on the WIC FI. The exact type and quantity of food a participant may purchase is based on one of five WIC participant eligibility categories: pregnant women, postpartum women, breastfeeding women, infants, and children.

Table I-1 displays the types of foods that participants in each of the five eligibility categories are authorized to purchase. Because the foods are chosen to meet the specific nutritional needs of persons in the different categories, the FIs are regarded by Federal regulation and program staff as "food prescriptions."

Table I-1. Authorized WIC Foods by Eligibility Category					
Food Type	Pregnant Women	Postpartum Women	Breastfeeding Women	Infants	Children
Milk	X	X	X		X
Cheese	X	X	X		X
Eggs	X	X	X		X
Cereal	X	X	X		X
Juice	X	X	X		X
Beans or Peanut Butter	X	X	X		X
Tuna			X		
Carrots			X		
Infant Formula				X	
Infant Juice				X	
Infant Cereal				X	
Special Formula*	X	X	X	X	X
* Under special conditions only					

Participants must redeem their WIC FIs at retail stores that are authorized by the program. Some States operate a vendor-specific retail system, in which participants are required to select a single authorized store and to purchase all of their WIC foods at that store. Other States operate open systems, in which participants may use their FIs at any WIC-approved store. Among geographic WIC State agencies with retail delivery systems, there are 11 that are vendor-specific and 35 which operate open systems. Figure I-1 displays the food delivery systems of all 50 geographic State agencies.

2. Procedures for Purchasing WIC Foods

Each month the WIC participant receives a series of FIs that may combine a variety of food categories. For example, a participant may be issued certain FIs that authorize the purchase of milk, eggs, cereal, and juice, and other FIs that allow the purchase of peanut butter and cheese. Within each food category, participants are given a choice of food products. For example, participants can select different types of cheese, such as Colby, Cheddar, or Swiss; or various brands of WIC-approved cereals.

Legend:

- Open System States
- Vendor - Specific States
- Home Delivery State
- Direct Delivery State

- Page 4

Each WIC FI has a number, which is recorded in the participant's issuance file at the WIC agency where the FI is issued. A participant takes the FI to an authorized store, makes selections, and pays for the foods by presenting the WIC FI to the cashier. The WIC agency is required to use the FI number and the total dollar amount it was redeemed for to create a redemption database.

3. The Need for Food Purchasing Data

In any State that uses a FI for delivery of WIC benefits, officials administering the program have no means of conducting an examination of the purchasing patterns of WIC participants. The only purchasing information available to program officials is the total expenditure amount for each WIC transaction, which is reflected by the dollar amount entered on each FI by the store cashier. Although WIC nutritionists encourage participants to purchase all their prescribed foods, WIC staff have no way of tracking whether these foods are actually purchased. In addition, even though officials know the range of foods available for purchase, they do not know which brands and forms of food participants are purchasing. They have no information about how many trips participants make to the store, what foods they do or do not purchase, or the costs of individual food categories.

Information about WIC participants' food purchasing patterns could enhance the ability of FCS and WIC State and local agencies to administer the program. Data indicating how often participants shop, which foods they purchase or do not purchase, and the cost of each WIC food category could better prepare WIC officials to make decisions about policies related to food authorization, the content of nutrition education materials, and selection of foods to include in their WIC food package. States could also project monthly food costs more accurately and examine reasonable alternatives when considering which food products to include in the program's authorized food list.

B. Purpose of the Study

With the introduction in recent years of bar-code scanning technology and Universal Product Codes (UPCs) and Local Product Codes (LPCs), supermarket checkout systems can generate a wealth of data about individual food purchases. UPCs and LPCs are unique codes assigned to individual food products. Stores use these codes to create an individual product description in their data systems that identifies the food product by such characteristics as product brand name, type of food, container size, and price. When the product is scanned at the checkout stand, the computer system recognizes the code and records the purchase of the food. These data can be aggregated and analyzed in various ways to study food purchasing patterns of individuals and groups, including WIC participants.

The purpose of this study, which was implemented in October 1994, was to test the feasibility of combining data from a WIC agency's participant records, which include food prescription and client demographic data, with information from local grocery stores' UPC and LPC data files created by point-of-purchase scanning. By linking these databases, information about the shopping patterns of WIC participants can be obtained.

In examining food purchasing patterns of WIC participants, the primary goal of FCS is to ensure that participants obtain maximal nutritional benefit from the food package issued to them. Data about food purchases will allow FCS to evaluate the food packages and whether adjusting the types and amounts of food would improve the nutrition of participants. Thus, FCS has two main research objectives: to determine the feasibility of using scanning technology to describe food purchasing patterns of participants and to examine differences among participants in such patterns. The differences in purchasing patterns examined in this study were between rural and urban participants, between seasons, between WIC eligibility categories, and among four ethnic groups: Blacks, Hispanics, Asians or Pacific Islanders, and Whites. To explore feasibility under a variety of conditions and with different populations, two States were selected to participate in the study—Wisconsin and Pennsylvania. To examine

seasonal purchasing patterns, data were collected over two time periods, one in the spring/summer and one in the fall/winter of 1995.

It should be emphasized that the purpose of this project was to test the feasibility of linking the two databases, as well as the feasibility of analyzing the resulting combined data set. This study was not designed to describe the purchasing patterns of a scientifically selected representative sample of WIC participants or the stores in which they shop. The data reported here are applicable only to the population included in the study. In addition, the chain stores selected may not necessarily represent the types of stores where the majority of WIC participants shop. Moreover, no attempt was made to determine reasons for shopping patterns of WIC participants. Some of the data described in this report are highly suggestive of particular interpretations; however, because no valid comparison groups were available, no attempt was made to statistically test the significance of the findings or to examine the validity of any conclusions that might be made from these data.

C. Potential Uses

If this process appears feasible and/or cost effective, the data could be used for several purposes. This section will describe the potential uses of the data by Federal and State WIC Program officials. As indicated earlier, data collected and analyzed using the technology described in this feasibility study can be useful to Federal and State officials in evaluating the food purchasing patterns of WIC participants. Four key areas in which these data may be useful are described below.

1. Potential Uses: Food Selection and Benefit Utilization

a. Developing and Assessing the Impact of Food Selection Restrictions

Under Federal WIC regulations, WIC agencies or State directors are authorized to develop criteria for selecting specific brands and forms of food to be included in their food package. In addition, State WIC directors may need to make decisions to restrict participants' food choices to lower priced foods when necessary to remain within the

State's budget limitations. Within each food category, high-cost items can be eliminated as long as other food choices within that category are available. For example, during the months data were being collected for this study, the issuance of peanut butter, a relatively high-cost item, was restricted in Wisconsin to odd-numbered months. Participants were authorized to purchase dried beans, a nutritionally similar food, during even-numbered months. States can also set stricter nutritional requirements for authorized foods, for example, by further limiting the sugar content of cereals offered.

In current WIC State retail delivery systems, the exact quantity of individual food items purchased cannot be determined. Therefore, when WIC State directors make the decision to restrict or remove a food from the program, they do not know what the overall impact of that decision will be—the number of participants affected by the change, or the total amount of savings that may be created or the impact on demand at WIC vendors.

Data about the types of food and the forms of food selected by participants would allow State officials to make more informed decisions about food restrictions and to better assess the impact of their decisions on participants and stores. For example, data concerning cereals and juices popular with certain ethnic groups would allow States to evaluate the impact of eliminating these choices from the WIC food categories and substituting other lower-cost items when necessary.

b. Developing Food Package Tailoring Policies

WIC local agency nutritionists have the option of tailoring participants' food packages to better meet their nutritional needs. Under the current system, the nutritionist must rely on information provided by individual participants to determine if they are purchasing prescribed foods. By having information about the foods not purchased by participants, the nutritionist could develop policies for tailoring food packages to maximize the benefit to the participants. For example, if data reveal that certain ethnic

groups do not purchase all of the prescribed milk, but purchase all of the prescribed cheese, the nutritionist will have additional information to develop guidelines and policies for reducing the amount of milk issued and increasing the cheese issuance.

c. *Developing Nutrition Education Activities and Materials*

Most WIC nutrition education is focused on the importance of eating the appropriate foods. If State and local nutritionists know that participants do not purchase foods authorized, they can direct their efforts to encourage participants to purchase and use those foods. For example, adolescents may not purchase dried beans because they do not know how to cook them. Adults may not understand the importance of limiting fat intake by drinking low-fat and skim milk rather than whole milk. With data about food purchases, WIC nutritionists can better target nutrition education.

d. *Improving Product Stocking Requirements for Vendors*

Food purchasing data would also help WIC State and local agencies ensure that newly authorized WIC vendors stock enough of the more popular food items to provide an adequate supply for WIC participants. When a store applies to become an authorized WIC vendor, States could better direct stores in stocking WIC foods. In addition, States could use this information to develop indicators that a store is not stocking sufficient foods. If data show that certain foods are not purchased in a particular store, this may indicate that the store is carrying insufficient stock of the food product.

2. *Potential Uses: Number of Shopping Trips/Checks Used*

States can also use data regarding the number of shopping trips made by WIC participants. In particular, States may use such data to assist them in developing policies related to the number of FIs the food package is distributed over. States have an interest in reducing the number of FIs issued to participants. Each FI issued by a program and processed by a bank or fiscal intermediary incurs a cost to the State in a processing fee. The total amount of FI charges is significant considering that millions of FIs are issued. Some participants may make frequent

shopping trips due to transportation difficulties, lack of refrigeration, or the large number of infant formula cans they must purchase. In addition, States may wish to explore the underlying causes of a low number of trips to a specific store. While a low number of trips is not necessarily a problem, States may wish to examine the reasons for participants making a small number of shopping trips, including the convenience of the store location.

Data about the number of trips participants make would allow States to assess whether they are issuing an appropriate number of FIs. If participants are making only a few trips for reasons unrelated to grocer convenience, States may decide to reduce the number of FIs participants receive by combining more foods on one FI. This would not only reduce the cost of FI processing fees, but may also expedite the processing of transactions at the checkout counter.

3. Potential Uses: WIC Food Expenditures

An additional use of food purchasing data would be to address the cost of the WIC food package and the amount spent or not spent. While there is limited use for expenditure data from a single store or a group of stores, expansion of the methodology may have specific benefits to WIC State agency fiscal staff.

a. Enhancing Food Package Expenditure Forecasting

States are required by FCS to project monthly WIC food expenditures and report actual expenditures against projected expenditures. States usually project expenditures based upon monthly redemption data that provide officials with an average cost per food package. One problem with this method is that it relies on data captured several months after the FIs have been redeemed. It is difficult to use these data to project future expenditures if there are changes in the allowable foods during the fiscal year. Having accurate and timely food product redemption rates would provide State officials with an additional tool to estimate food costs based upon actual purchasing patterns for individual food categories, thus allowing adjustments for changes in policy over the course of the fiscal year.

b. *Calculating Food Expenditures Associated With Caseload Expansion*

Having expenditure data for each food item would assist State officials to project the fiscal impact of caseload expansion. Knowing the differences in food package costs for urban or rural stores and for participants with various demographics would allow States to project the increased costs of adding participants in urban and rural areas or increasing services in areas serving a particular group.

D. Overview of the Methodology

The study examined the feasibility of linking data from two databases: food purchasing data from a supermarket point-of-sale transaction and data from the WIC State agency database. As noted above, under the current system, WIC agencies can determine only that a FI was used or not used, and the total dollar amount for which the FI was cashed. Detailed information about the type and form of foods purchased are unavailable. With the introduction of supermarket scanning and point-of-sale databases, it became possible for the WIC Program to tap into the wealth of data about food purchases contained in supermarket databases.

The means used to link the supermarket data and the WIC participant data was the WIC FI number. Each WIC FI has a unique number, and the WIC agency database contains a record of the numbered FIs issued to each participant. At the supermarket point of sale, after scanning the WIC-authorized foods at the checkout counter, the cashier entered the WIC FI number into the transaction record. The WIC FI number became the single common data element of both databases and a key step of the methodology described in this report is matching the supermarket transaction record to the participant data in the WIC database using this number. References below and in Chapter II to the "match rate" indicate the extent to which each WIC point-of-sale transaction was able to be matched by its check number to participant food issuance and demographic data in the WIC agency database.

Processing the food purchasing data supplied by the supermarket and matching it to WIC participant data required the building of a combined database. Periodically, the supermarket

supplied the data processing agency with a file containing raw UPC/LPC data for all the WIC purchases made at the study stores during the data collection period. The first task of the data processor was to convert the raw data into usable information. For this task, a conversion table was needed. A conversion table is a list of all the foods authorized for purchase by WIC participants and their respective UPC/LPC codes, including different types of a food (e.g., Colby, Swiss, Cheddar, etc.) and the different forms of a food (sliced cheese, bulk, etc.), the container size, and the price. Armed with these data and with the WIC food instrument (FI) numbers for each of the supermarket transactions, the data processor was able to match an individual shopping transaction with WIC participant data.

Appendix A presents a description of the merged data set, containing detailed information about the foods purchased in the transaction and WIC participant issuance and demographic data. These data can be analyzed in numerous ways, some of which have been described above in the section on potential uses for the data.

The next section lists specific feasibility questions that were answered by implementing the methodology described above.

E. Feasibility Questions

Because this was a feasibility study, the findings of most interest and value are related to the lessons learned in selecting States and stores to participate in the study, in working with WIC and chain store representatives in Pennsylvania and Wisconsin to link their databases, in ensuring that the merged data set that resulted was flexible and user-friendly, and in learning to examine the data in meaningful ways.

The feasibility findings are addressed in detail in Chapter II. These findings are based upon a series of feasibility study questions detailed below.

1. **Is technology available that will allow data from WIC State agency files to be matched with point-of-sale transaction data at WIC-authorized stores?**
 - 1.1 What elements are necessary in a WIC State agency data system to make the technology feasible?
 - 1.2 What are the required characteristics of a store's point-of-sale data system to allow transaction data to be matched with WIC State agency data?
 - 1.3 What changes would stores need to make in their data systems to participate in a project to match State WIC data with transaction data? Are these changes feasible?
 - 1.4 What is the cost to the stores of making changes in their data systems to accommodate the data collection requirements?
 - 1.5 By what methods is it feasible to link WIC State agency data with store transaction data?
2. **Is using UPC/LPC data sufficient to identify the foods purchased by WIC participants?**
 - 2.1 How can WIC-authorized foods be identified in the store's UPC/LPC database?
 - 2.2 Once identified, are data in the UPC/LPC database complete enough to determine the type and form of foods purchased?
 - 2.3 What is the best method of converting UPC/LPC codes into useful transaction data?
 - 2.4 Can WIC transactions be separated from non-WIC purchases in the transaction record?
3. **How will the completeness of the data be affected by cashier point-of-sale data entry?**
 - 3.1 How will cashier errors and omissions affect the completeness of transaction data?
 - 3.2 How will cashier errors and omissions affect the match rate between State data and transaction data?
 - 3.3 Does training of cashiers improve the data entry error rates?
 - 3.4 Does cashier familiarity and experience over time improve the data entry error rates?

4. **Are algorithms available to compensate for data entry errors and improve the match rate?**
 - 4.1 What sources of data are available to use in the development of the algorithms?
 - 4.2 How effective are the algorithms in improving the match rate?
5. **Are there methods available to improve point-of-sale data entry?**
6. **How do State food delivery systems affect the completeness of the data?**
 - 6.1 How many States use a vendor-specific food delivery system? An open system?
 - 6.2 Are there differences between the completeness of data in vendor-specific States and open system States?
 - 6.3 What are the advantages and disadvantages of each system in collecting complete matched data?

CHAPTER II

Feasibility of Using Scanner Technology to Examine WIC Participant Food Purchasing Patterns

A. Introduction and Overview

This study examined the feasibility of merging data from WIC State agency files with grocery store transaction data tapes built from scanned UPC/LPC codes, which contain a detailed record of information about WIC participant purchases, to analyze the food purchasing patterns of WIC participants. The merged database provides information about the types, forms, cost, and amounts of food purchased by each WIC participant with each WIC FI, and, conversely, what authorized foods participants did not purchase.

The purpose of a feasibility study is to separate what is realistically possible from what is theoretically possible—and to account for the difference. The main focus of this chapter is on the findings of the feasibility study. The findings are based on a study carried out over approximately 18 months (October 1994–April 1996). At the beginning of the study, several States and chain stores were considered as possible candidates to participate in the study. The chain stores selected to participate, and selected local stores in the chain, met the requirements for having the appropriate scanning technology, serving rural and urban WIC shoppers of diverse ethnic backgrounds, and having a high and a low volume of WIC sales. WIC State agencies and the selected store chains were briefed on the study's purpose and methods, and staff were prepared for data collection.

B. Major Findings Related to the Feasibility of Matching WIC State Agency Data With Store Transaction Data

This section presents the major findings related to the feasibility of collecting and matching data from WIC State agency participant files with store transaction data. However, there are two important considerations to take into account when examining the findings.

First, it is important to note that the success of the entire study is based on the ability and willingness of WIC-authorized vendors to cooperate with the objectives of the study and implement new procedures within their stores. Not all WIC-authorized stores will have the necessary technology to participate in a study of this nature. As will be discussed in these findings, the primary technology that must be available in a store is the capacity to scan UPC/LPC codes to record purchase transactions into a database. In attempting to determine how many stores serving WIC participants have this technology, researchers contacted various trade associations, scanner companies, and the FCS Supplemental Food Program Division to ascertain if data were available on the number of stores with scanning systems. However, no data currently exist that provide information about the number of WIC authorized stores with scan technology. Researchers were informed by FCS that such information will be collected from WIC State agencies in 1998. Once these data are collected, WIC State agencies will be able to determine the extent to which a study similar to this one can be conducted.

Second, the feasibility study was designed to assess the capacity of *geographic* State agencies to collect and analyze food purchasing data. No attempt was made to determine the feasibility of Indian Tribal Organizations (ITO) conducting a study of this nature. While some of the findings are relevant to ITOs, differences in data system capability and types of stores used by clients affect the generalizability of the findings. In addition, the overall cost of conducting this type of research may render it infeasible for these agencies.

Finding #1. Data from WIC State agency participant files can be matched with store transaction data to provide information about WIC participant food purchasing patterns.

By developing a method by which food purchasing information collected at the point-of-sale can be linked to WIC State agency files, a database can be developed and analyzed to provide a description of the food purchasing patterns of WIC participants. Information contained in the matched database provides information about WIC participant food purchasing patterns, including the number of shopping trips, the type and form of the foods purchased, the foods not purchased, and food cost information. Information can be analyzed by both participant and store demographic characteristics. A complete description of the structure of the combined database is included as Appendix A.

Once data are received from a WIC State agency and the selected stores, building a complete database requires both computer programming and data processing time. To build the complete database for Wisconsin, 455 programming and data processing hours were required for the first data collection period and 154 hours were required for the second. To develop the Pennsylvania database, 372 programming and data processing hours were required for the first database and 184 hours for the second.

There are key elements that must be present in both the WIC State agency and the grocery store to make this methodology successful. The WIC State agency must have an automated data system that captures both client demographic data and FI issuance information for each client. The agency must have redemption data that can link a paid WIC FI with the participant using it and the store where it was cashed.

Stores selected for conducting this type of study must have several elements in place for the methodology to be successful.

- ***Stores must have a product scanning system that captures transaction data at the point of sale and links those sales to specific food products through the use of Universal Product Codes (UPCs) or Local Product Codes (LPCs).***
In order to create the automated database, a store must have a method by which

food purchasing information is captured electronically. Use of a product scanning system that identifies food products by their UPC/LPC identifiers is critical to capturing transaction data. Without the ability to identify food products by individual type, form, and cost, the methodology will not be successful.

- ***The store's system must be able to distinguish WIC transactions from other routine transactions.*** Stores with electronic scanning systems have methods by which they record the type of payment being used to make a purchase. This identification of a payment is known as a "tender." The most common tenders are cash, check, or Food Stamp. In order to identify the transaction as being paid for by WIC FIs, the store must have a method to include a "WIC tender" in their system. This will usually require the store to reprogram their data system to include a WIC tender. Chain stores participating in this study reported that there was little additional cost to conduct the reprogramming. In both States' chain stores, the cost of reprogramming their computer systems to meet this requirement was less than \$500 for each chain.

If a purchase cannot be identified as a WIC transaction in the store's database, it will be impossible to separate it from any other transaction in the files. Because WIC foods are common foods purchased by most people, and because WIC foods supplement a participant's monthly food purchases, the lack of a method by which to identify the transaction as a WIC transaction makes it difficult to determine if a WIC participant purchased her foods with a WIC FI or with another form of payment.

- ***The store's data system must have a method by which the WIC FI number or client identification number can be entered into the transaction record.*** Simply identifying a transaction as a WIC tender is not sufficient to link the transaction to a particular participant. The store must be able to enter a WIC FI number or a participant identification number into the database in order to link the transaction to the participant's WIC food prescription. A WIC FI number is the preferred method for linking the transaction, as this creates a direct link between the purchase and the FI. When a participant identification number is used, the method by which the transaction is linked to the issuance record requires several additional steps.

Scanning systems have the capacity for entering numbers, such as a personal FI number or a driver's license number, into their system. These number entry systems can be used to enter a WIC identification number, whether it is the FI number or the participant identification number. While this creates a small additional burden to the cashier, stores participating in the study reported little disruption to the checkout process. The total cost of reprogramming the stores' data systems to capture the WIC FI numbers was less than \$1000 per chain store in each State.

However, having all of the above systems in place does not guarantee that complete food purchasing records will be available for all WIC participants using the store. As is seen in the remaining findings, there are a number of variables that affect the completeness of the merged database.

Finding # 2 Developing a transaction file from raw UPC/LPC data requires refinement on the part of the stores and the State data processing agency.

Building a transaction file from raw UPC/LPC data contained on the supermarket data tapes requires significant refinement on the part of the stores and data processors. The data processing system used by the stores collects data on all purchases made by any person shopping at the store. WIC participants represent only a portion of those transactions. In order to match food purchases to a WIC participant, a file must be created that separates the WIC purchases contained in the database from those made by other customers. To do this, the store must sort the files by the code developed to identify the purchase as a "WIC tender." All records that are identified as a WIC purchase are then placed in a separate file for transmission to the State data processing agency for matching with individual food issuance data. Two specific factors influence how the transaction database was developed.

Finding 2.1 Conversion of UPC/LPC codes requires cooperation between the WIC State agency, the data processing agency, and the selected stores.

The data provided to the WIC State agency contain the UPC/LPC coded information about the type, form, date of purchase, and cost of each authorized WIC food purchased by the participant during a single transaction. In order for the information to be useful, the code must be converted to the actual product name, type of food, size of container, date of purchase, and price. However, to provide a State data processing agency with the information to convert the codes into useful information, the store must first have information about which food products are authorized by the WIC State agency. Without this information from the State agency, the stores cannot identify all of the WIC-eligible food products by their UPC/LPC codes.

To develop UPC/LPC conversion tables listing WIC food items and their corresponding codes for this study, the data processing agency had to initially sort out all of the UPC/LPC codes contained in the transaction data and provide the chain store officials with a list of those codes. The stores were then able to identify each product and send the list back to the data processing agency with the necessary information for each food product listed in the conversion table.

States and stores will need to work out how information needed to create conversion tables will be transmitted to the stores before data collection begins. States will need to pay close attention to the data contained in the conversion tables to ensure that the food purchasing data collected are as comprehensive and detailed as possible. Some stores do not identify food products in as much detail as WIC State agencies may wish. There were three specific food groups for which the lack of specific product information limited the ability to identify a food product in detail.

- ***Cheese Purchases.*** Data on cheese purchases collected from the chain store in Pennsylvania could not identify the specific type of cheese if it was purchased as a deli item. Cheese purchases were broken down by type of cheese (Colby, Swiss, American, etc.) and form of cheese (sliced, block, shredded, etc.). Each type and form was identified by a separate code. However, from 30 percent to 50 percent of the cheese sold in Pennsylvania was not able to be identified by type or form because it was identified in the store's LPC system only as "deli cheese." Deli cheese is typically, but not exclusively, block cheese, but can be of any type.
- ***Juice Purchases.*** Similarly, in both Pennsylvania and Wisconsin, only single-flavored juices (orange, grape, apple, etc.) could be identified by their separate codes. All mixed juices were identified by a single code "juice." In this case, however, while the specific flavor of the juice could not be identified, the form of the juice (canned, bottled, or frozen) was able to be identified. In both States, these mixed juices were the type most frequently purchased by WIC shoppers, and they could not be identified further by flavor.
- ***Bean/Legume Purchases.*** In some cases, beans that were purchased in bulk from bins in the stores were identified as "beans/legumes," rather than by the specific type of bean.

Because the study used the initial UPC/LPC codes from actual WIC purchases during the first round of data collection as the basis for building the conversion tables, a problem was identified in the second round of data collection. During the second round of data collection, new UPC/LPC codes appeared in the data sets. This was a result of WIC participants purchasing some approved foods in the second round that had not been purchased in the first round. Therefore, new UPC/LPC conversion tables were built using the same process described above to add the new products. Had a complete listing of all possible authorized WIC food products been developed during the first data collection period, this problem would not have occurred.

Finding 2.2 Including non-WIC transactions in the database may hinder accurate reporting of foods purchased.

A second issue that required significant attention by the data processing agency was the fact that more than 50 percent of the stores' raw transaction files contained information about non-WIC purchases as well as WIC purchases. This situation occurred because WIC foods are normally scanned separately from other foods the participant is purchasing, but only subtotaled into a single transaction rather than counted as two separate transactions. Having these non-WIC data on the record made accurate identification and analysis of the WIC purchases more difficult. Foods that were not WIC-eligible were easily identified and were sorted from the files. However, if clients purchased additional WIC-eligible foods with cash, the record became more difficult to analyze. For example, if the issuance record authorized the purchase of two gallons of milk, and the client purchased two gallons of milk with the WIC FI and one additional gallon with cash, the transaction record did not distinguish between the transactions, and it appeared that the client over-purchased milk. For purposes of this study, food products in excess of the maximum allowed in the issuance record were removed from the transaction record.

The only method to determine if the milk was purchased by using a WIC FI was to compare the total prices of the issued food on the FI with the total dollar amount for which the FI was deposited. In the example above, if the total dollar amount of the deposited FI was the same as

the total of the prices for foods issued on that FI, then the client did not use the WIC FI to purchase the additional milk. If the total amount of the deposited FI included three gallons of milk, then the client would have exceeded the authorized limit.

While the identification of UPC/LPCs created a problem for completing the transaction record, the problem was solved through communication between the State, the store, and the data processing agency. Approximately 20 hours of chain store and data processing staff time were required to build a complete UPC/LPC conversion table. However, other problems affecting the accuracy of the data occurred during the study, some of which could not be completely resolved. These problems are discussed in Finding #3.

Finding #3. Data entry errors and omissions made at the point of sale have a negative impact on the completeness of the matched data set and the subsequent analyses.

Complete, accurate, and matchable transaction data are critical to the success of the methodology. Perhaps the single most important finding of the feasibility study was that the accuracy of the cashier in entering the WIC FI number was the most critical element in creating a successful merging of the WIC participant data and the store transaction data. Cashier error played a large role when store transaction data could not be matched to WIC issuance data. The extent to which these errors affected the match rate are discussed in later findings and displayed in Table II-1. The major errors identified include:

- ***Cashiers Did Not Identify the Transaction as a WIC Transaction.*** In Wisconsin, the cashier was required to enter the WIC tender before they could complete the transactions. However, in Pennsylvania the system used by the chain store allowed a cashier to bypass the process of identifying the WIC transaction and still complete the sale. As a result, an unknown number of the WIC transactions did not get identified as such. Thus, the WIC transaction appeared as a normal cash or FI transaction, and was not included in the matched database.
- ***Cashiers Did Not Enter the WIC FI Number or Entered Incomplete FI Numbers.*** In both States, even when a purchase was identified as a WIC transaction, it was possible for the cashier to bypass entering the WIC FI number into the system and still complete the sale. In addition, some records

had fewer than the required number of digits entered in the FI number field, indicating that a complete FI number was not entered.

- **Cashiers Entered Incorrect WIC FI Numbers.** The majority of errors in Wisconsin occurred in the entry of the FI number into the system. This was due in part to the similarities between the participant identification number and the WIC FI numbers. The two numbers are the same length, and both are printed on the face of the FI. In 909 cases (14 percent) during the first data collection period and 1045 cases (11 percent) in the second, the cashiers entered the client's WIC identification number in place of the FI number.

Even with the problems associated with cashier error, there are methods by which some of the transaction record errors can be resolved. These are discussed in the next two findings.

Finding #4. Training and experience seem to improve cashier error rates, thus improving the ability to match transaction data with WIC participant data.

Senior representatives of the chain store in Pennsylvania preferred that cashier training be conducted by customer service managers in the individual participating stores, rather than by outside trainers, as this approach would be less disruptive to store operations. The purpose of training is to inform store managers and cashiers about the study and review the keying sequence for processing WIC transactions. Training materials were developed and delivered to the chain headquarters in advance of the training sessions, which included:

- ***A one-page summary of the study describing the data collection methodology and the steps cashiers would need to take when handling WIC transactions.*** The summary included information on the reason for the study, an overview of the study design, and an explanation of the contractor's role.
- ***Two substantially enlarged facsimile WIC FIs, one handwritten and one computer-generated.*** The purpose of these samples was to illustrate where FI numbers are located on WIC FIs.
- ***Five hundred cards that contained an explanation of the study.*** These cards were sent to the store headquarters for distribution to store managers. They were to be given to WIC customers who inquired about the study.

However, the chain store in Pennsylvania failed to carry out training before the first data collection period began. As a result, cashiers failed to identify transactions using the WIC tender, or failed to enter the FI number into the transaction record. The final match rate for the first data collection period was 5,379 matching records from 17,531 eligible issuance records, or 36 percent, much lower than expected.

Several steps were taken to develop a reliable training plan before the second data collection period. In a memorandum to the customer service managers at the study stores, senior management stated that they must inform cashiers to enter WIC FI numbers. The memo also required each customer service manager to have all cashiers sign a form stating that they had been informed of the study and understood what they were being asked to do.

In addition, the study protocol was reviewed individually with each store manager, and any questions were answered. Each step in the keying sequence was reviewed, and it was explained that if the cashiers did not complete the entire keying sequence, the transaction would not be identified as a WIC sale. As a result, the final match rate during the second round of data collection improved to 13,908 matching records from 21,829 eligible issuances records, or 47 percent.

The contract for this study did not call for training of cashiers in Wisconsin. The chain store selected in Wisconsin was already using a WIC tender system. Cashiers working in the selected chain store simply were instructed to enter the WIC FI number into the system once the tender was complete. The initial match rate for Wisconsin—based on matching only the WIC FI numbers—was 60 percent of all FIs issued to the study stores.

However, as in Pennsylvania, the match rate between transaction and issuance records in Wisconsin improved between the first round of data collection and the second round. During the intervening months, even though data were not being gathered, cashiers continued to enter WIC FI numbers into the system. As a result, by October, the cashiers had been entering the WIC FI numbers for five months and the match rate improved to 80 percent. Store officials

attributed the higher match rate to the fact that checkers became more skilled at entering the WIC FI number. Early training of cashiers prior to the data collection period increases the likelihood of high-quality data entry at the point of sale, and accuracy of data entry skills improves over time.

Finding #5. *Use of an algorithm based upon information contained in the WIC State agency issuance files and bank records can also improve the match rate between issuance and transaction data.*

Even with training and experience, it is likely that errors in the data entry process will still occur. To create the initial merged database, data from the WIC State agency files were matched with data from the stores' transaction records by means of the WIC FI number. However, largely due to the data entry problems mentioned in findings #3 and #4, the match on the FI number produced an unacceptably low rate of successful matches. Other methods of matching the available data records were needed, and a matching algorithm was developed to match issuance and transaction records using data elements other than the FI number.

The matching algorithm (i.e., the series of steps described below to improve the match rate) was initially developed and tested on unmatched items from the first round of data from Wisconsin and then used in all subsequent rounds of data collection. The steps in the algorithm are as follows:

1. **Matching on Participant's Identification Number.** In some cases, a cashier entered the participant's WIC identification number instead of the FI number. The number in the "FI number" field in the transaction record was then matched to the participant identification in the issuance record.
2. **Matching on Bank Record.** If the FI number from the transaction record did not match either the FI number or the participant identification number in the issuance record, the two records could sometimes be matched using information provided when the store redeemed the FI at a bank and the canceled FI was returned to the WIC agency. The bank record that is appended to the issuance record contains the date the FI was cashed by the store, the identification number of the store redeeming the FI, and the amount paid to the store. In the feasibility study, to match an issuance record to a transaction record using this information, for each unmatched issuance record, all transaction records that met three criteria were identified: (1) the transaction took place at the store that

cashed the FI, (2) the total WIC amount matched the amount paid by the bank, and (3) the date of the transaction preceded the date the FI was cashed by the store. If only one transaction record matched a given issuance under these criteria, the records were considered matching.

When these criteria are used, it is common for more than one transaction record to emerge as a potential match for a given issuance. To narrow the field of potential matches, one more step was added to the algorithm. This step limited the match criteria to those records in which the date of the transaction was no more than five days before the date the FI was cashed. Five days was selected as a limit to take into account situations when a transaction took place on a Friday evening of a holiday weekend and the FI could not be deposited until the following week.

3. ***Manipulating Incomplete FI Numbers.*** In the feasibility study, many transaction records had a number in the "FI number" field that had more or fewer digits than an actual WIC FI number. If such records were not matched in either of the steps described above, two additional efforts were made to match them. First, to attempt to match transaction records in which the FI number entered was shorter than a true FI number, FI numbers on all unmatched issuances were searched for a substring of digits that matched those in the transaction record's FI number. If such a record was found, it was considered to be a match if the store number and transaction amount were the same and if the transaction date preceded the date on which the FI was cashed by the bank.

Second, transaction records in which the FI number was too long were searched for cases in which a double zero appeared. If it did, one of the zeroes was eliminated and the record was compared with the remaining issuance records. (This was intended to account for cases when a cashier hit the "00" key instead of "0.") If a match was found, it was considered a true match only if the store number and amount matched, and the transaction date preceded the date the FI was cashed.

4. ***Hand Matches.*** Finally, a printout of all remaining unmatched issuance and transaction records was examined to find records in which the FI number was inaccurately keyed by the cashier. By examining each unmatched issuance against a list of candidate transaction records (ones that took place at the same store, for the same amount, with the transaction date preceding the date the FI was cashed), records in which the FI number was entered incorrectly were identified and matched to the correct issuance record.

Once the matched database was created, the various steps in the matching algorithm were tested by searching for records in which none of the items purchased (as described in the detail

records) were included in the WIC issuance. Such instances would be considered "bad matches" and would be deleted from the final data set. Matches were considered invalid only in cases in which the issuance and transaction records contained no items in common. No such records were found in any of the data sets.

The final match rate can be calculated in either of two ways: (1) as a percentage of the total number of eligible issuances to participants shopping at the study stores for which matching WIC purchasing transactions were found, or (2) as a percentage of the total number of identified WIC purchasing transactions for which matching issuances were found. The first method reflects the feasibility of the study design from the perspective of the WIC State agency's issuance records, the second presents a measure of the technological feasibility of the project to identify those transactions identified as valid WIC transactions in a particular retail chain. If all WIC transactions are correctly identified, the two percentages will be the same; however, since some WIC transactions are not likely to be identified because of cashier errors and omissions, and are therefore irretrievable, the number of identifiable WIC transaction records in the store files represents the maximum number of records that may be matched.

In Wisconsin, the match rate was calculated for each round of data collection in both ways. In Pennsylvania, only the percentage of cashed issuances that were matched was calculated. Because the study stores' data included all WIC transactions occurring during the month, the Pennsylvania transaction data sets included transactions that used issuances from months other than the study months. In Wisconsin, issuances from prior months were eliminated because the data processing agency had all of Wisconsin's issuance information for the months around the study period. These issuance data were not available to the data processing agency for Pennsylvania, and could not be eliminated from the data set. Therefore, using the total WIC transaction number as the denominator would produce an artificially low match rate.

Table II-1 displays the results of the matching algorithm for both rounds of data collection in Wisconsin and Pennsylvania. The number of matches found in each step was higher in the second round of data collection. The number of transaction records available for matching was

higher as well, indicating that cashiers were more thorough in identifying WIC transactions in the second round.

As Table II-1 shows, the match rate increased in the second round of data collection in both study States. In Wisconsin, the fact that the total number of WIC transaction records in the first round of data collection is lower than the number of WIC issuances indicates that the stores did not consistently code WIC transactions as such. In the second round, the two numbers were similar, indicating improvement in this regard. This was also reflected in the increase in the percentage of issuances that were matched.

Finding #6. A system that requires the cashier to enter the FI number into the transaction at the point of sale would likely produce higher match rates.

As was indicated above, a major problem with regard to data entry was the fact that a cashier could bypass either a WIC tender or enter an incomplete FI number into a transaction and still complete the WIC sale. Because there was no built-in method requiring the WIC transaction to be identified and a FI number to be entered into the data system, cashiers could circumvent the store's policy.

The feasibility study initially planned to test the use of a computer prompt to remind the cashier to enter the WIC FI number into the store's data system. The Pennsylvania chain store management indicated a willingness to program their computer system to require the entry of the WIC FI number into the data system. However, because the programming would have affected all of the chain's outlets, doing so to accommodate data collection in the 10 study stores was considered too disruptive. Had the cashiers been required to identify the WIC transaction and enter the FI number before they could complete a WIC sale, the problem of cashiers not entering the WIC FI number would have been eliminated.

Table II-1.
Results of Algorithm Used in the Feasibility Study to Increase the Match Rate
Between Issuance and Transaction Records

	<i>Wisconsin</i>		<i>Pennsylvania</i>	
	<i>Round I</i>	<i>Round II</i>	<i>Round I</i>	<i>Round II</i>
Total Eligible Issuance Records	8,668*	9,435*	15,081 [†]	29,651 [†]
Total Eligible Transaction Records	6,742**	9,486**	17,531 [‡]	21,829 [‡]
Step A: FI Number Match	3,762	5,803	3,009	12,163
Step B: Participant Identification Match	909	1,045	5	3
Step C: Bank Record Match (No Date Limit)	378	493	1,870	1,108
Step D: Bank Record Match (Date Within 5 Days)	102	107	471	601
Step E: Matches of Incomplete FI Numbers	0	0	24	33
Step F: Matches made by Hand	52	128	0	0
Total Matches	5,203	7,576	5,379	13,908
Final Match Rate	60% of issuances 77% of transactions	80% of issuances 80% of transactions	36% of issuances	47% of issuances

* FIs issued in the study month and cashed at one of the 12 study stores

** Identified WIC transactions taking place in the study months that did *not* use previous or later months' issuances

[†] FIs issued and cashed in the study month(s) for any of the 10 study stores

[‡] Identified WIC transactions taking place in the study months; includes those using previous months' issuances

Source: Health Systems Research, Inc.

Stores with scanning systems are already likely to be programmed to require the entry of personal check numbers or check cashing card numbers into their system. Expanding this technology to include the entry of WIC FI numbers would not necessarily constitute a major system change. However, if a chain store operates in more than one State, it would need to customize its systems to match different FI numbering systems in each State, increasing the computer programming costs.

Finding #7. *A State's ability to collect comprehensive data for a large number of participants, as well as the completeness of the data regarding food purchasing patterns for any individual, is affected by the nature of the State's food delivery system.*

The feasibility study demonstrated that data regarding food purchasing patterns of WIC participants can be collected, matched, and analyzed in one selected chain store within a State. However, the ability of a State to expand the data collection beyond a very limited number of stores that may agree to "volunteer" to participate in this type of study is limited by the approach States take to selling the use of the technology. The completeness of the food purchasing data for an individual participant is affected by the number of authorized WIC vendors that have the capacity to provide WIC transaction data to the WIC State agency. For example, small neighborhood markets may not have the scanning technology required to provide the necessary WIC transaction data to a WIC State Agency.

These issues are discussed below.

Finding 7.1 **States using a vendor-specific system are more likely to collect complete participant transaction data than States using an open system.**

One of the major findings of this feasibility study concerns the type of WIC FI system operated in the State. To a great extent, the comprehensiveness of the data collected depends on whether the State agency operates a vendor-specific FI system or an open system. As discussed in Chapter I, vendor-specific systems require WIC participants to identify a single store in which they will cash their FIs. In a vendor-specific State, WIC participants select one store in which to do all of their WIC shopping. Thus, when collecting and analyzing data from a single chain store, complete information on the participants' food purchase is available in a vendor-specific State. When collecting data from a single chain store in an open system State, complete participant food purchasing data is only available for participants choosing to shop at one of the study stores. The only way to capture more complete food purchasing information in an open system State would be to collect data from a number of different stores in a geographic area.

There are currently 12 vendor-specific States operating retail delivery systems. However, based upon data contained in the July 1996 State Agency Participation Report compiled by FCS, these States represent 35 percent of the total WIC population being served in all retail delivery States. This represents over 1.9 million participants shopping in States where complete participant food purchasing information could be made available through the selection of a single chain or group of stores.

Finding 7.2 Increased use of electronic benefits transfer (EBT) will likely increase the feasibility of utilizing the study technology.

Future methods of providing WIC benefits will lend themselves to this methodology. The increased use of EBT systems will cause States to reevaluate their grocer authorization practices. EBT will require stores to have in place the technology to read a participant's benefit card. Because EBT systems automatically record a client identification number at the point of sale, a clear link is established between the WIC transaction and the client, and cashier error is bypassed. Testing of the methodology developed through this study in an EBT environment would likely produce complete purchasing profiles of participants, regardless of whether they live in a vendor-specific State.

Appendix A: Description of Merged Data Set

Description Of The Merged Data Set

The following is a description of the merged data set created from matching the WIC State agency participant data with the store point-of-purchase transaction data.

1. The Header Record and the Detail Record

Table A-1 presents the elements included in the merged data set, which is the end product of the methodology described in this report. Deciding on the structure was a complex task, because each WIC transaction can contain multiple food items and each WIC participant can have many transactions. The final data set has to be flexible enough to facilitate analysis on the levels of item, transaction, person, and family.

As shown in Table A-1, each WIC transaction record contains a header record and at least one detail record. The header record includes demographic information about the participant to whom the WIC FI was issued and information about the participant's WIC category, status, and priority (column 1). The header record also shows the types and amounts of foods the participant is authorized to purchase (column 2). These data are obtained from the demographic issuance records maintained by the WIC State agency. Information about the transaction, such as the FI number, store number, and dollar amount, is also contained in the header record (column 3). This information comes from the records submitted by the retail store.

The detail records are files containing the specific food purchasing information obtained from the stores. These detail records show the type and amount of a particular food actually purchased in that transaction (column 4). It contains comprehensive information on the item purchased, including brand, type, form, volume, quantity, and price, and the number of the FI used to purchase the item. Each item purchased in a given WIC transaction will have its own detail record describing the product and linking the purchase to a FI number. After the sets of header and detail records have been created, the two data sets can be combined by linking a header to its corresponding detail(s) using the FI number. Detail records can then be aggregated to produce a

summary record of all items purchased with a given FI. Similarly, an individual WIC participant's records may be combined to produce a single record containing that client's total issuance and purchasing information.

Table A-1. Data Elements From WIC Agency Files and Supermarket Transaction Records Included in the Merged Data Set			
<i>Header (Transaction) Record</i>			<i>Detail (Item) Record</i>
<i>Demographic Information</i>	<i>Issuance Information</i>	<i>Transaction Information</i>	
Encrypted participant identifier	Gallons of milk	FI number	UPC code
Date of birth	Pounds of cheese	Transaction date	Description
Education level	Dozens of eggs	Store number	Brand
Ethnic group	Pounds of carrots	Total purchase amount	Food category
Migrant/refugee code	Cans of tuna	FI paid amount	Type
Homeless code	Pounds of beans/peas	Draft type	Form
Household smoking code	Ounces of peanut butter	Match step (for algorithm)	Item price
Health care source	Cans of juice and size of cans (in ounces)	Number of FIs used	Item quantity
Sex	Ounces of cereal		Item volume
Family income	Cans of formula		FI number
Income period	Cans of infant juice		
Family size	Ounces of infant cereal		
Breastfeeding status (for infants)			
Breastfeeding duration			
Immunizations			
WIC status			
WIC category			
Priority			
Source: Health Systems Research, Inc.			

The merged record was originally designed to represent one WIC FI, and to include all detail records of items purchased with that FI. However, individuals occasionally use more than one FI in a single transaction; in these cases, it was unclear which header record the associated detail records should be linked to. Therefore, the header record was redesigned to represent a single *transaction*, rather than a single *check*. When only one check is used in a WIC transaction, the two will be the same. When more than one check is used in a transaction, the header record represents the *total* amount of food in each category issued on all checks used in the transaction. The check number used to link this record to the associated detail records is that of the last check used in the transaction. A final data element in the header record represents the number of checks used in the transaction. This allows identification of records in which more than one check was used and entered into a single transaction record. The records could then be separated by participant and transaction when a single transaction record reflected a purchase by multiple participants, such as a pregnant woman who shopped for herself and an a child participant.

2. Presentation of the Data in the Merged Data Set

a. Identifiers

The data set was designed to exclude information with which to determine the actual identity of WIC participants. Instead, the data included two encrypted identifiers: a five-digit family identification code, created by numbering sequentially all families in the data set from an undisclosed starting point, and an individual identifier, created by adding one digit to the family identification. In this field, a "0" represents a pregnant or postpartum woman; other numbers represent children, numbered in order of age, youngest to oldest.

b. Demographic Characteristics

A number of demographic data elements, including information relating to WIC enrollment (WIC eligibility category, status, and priority) and the participant's ethnic background and socioeconomic status (racial/ethnic code, income, and education level) are included in the header records. Based on the participant's date of birth, the participant's age was calculated as of the middle of the study period. Using the

participant's reported family income, the family's poverty status was calculated as defined by the 1995 Federal Poverty Income Guidelines.

c. *Foods Issued*

All WIC FIs issued in Wisconsin and Pennsylvania fall into more than 300 draft type codes in each State, each of which describes a specific package of WIC foods. The original issuance records contained a field for this "draft type" code rather than a list of all foods issued. To create usable variables describing the amount of food issued in each food category, each draft type was converted into 12 separate variables, each containing the amount of one of the 12 WIC food categories issued on that type of check. Table A-2 presents the units of measure used for each category.

d. *Other Information in the Issuance Record*

In addition to demographic data and information about food issued, WIC issuance records contain information about the ultimate disposition of the WIC FI. If the FI is not cashed within its valid time period, the record indicates that the FI had been voided. If the FI is redeemed by the store, the record gives the date the FI was deposited at the bank and the amount the store was paid. In most States, the record also contains the identification number of the store depositing the check.

e. *Foods Purchased*

The detail records contain information about each food item purchased with one or more WIC FIs. This information is based on the UPC and LPC codes received from the stores, each of which contains the brand, type, and volume of the item purchased. For the feasibility study, these codes were converted into three variables, each of which has different possible values for each food category. The values for these variables are presented in Table A-2.

To compare the amount of food issued with the amount purchased, the amount of each item purchased is presented in the data set in the units used for the food issued. In addition to the volume of the item purchased, the detail record contains the number of items purchased and their total price.

Table A-2.
Values for Descriptive Variables by Food Category

<i>Food Category</i>	<i>Units</i>	<i>Type</i>	<i>Form</i>	<i>Brand</i>
Milk	gallons	1=skim 2=lowfat 3=whole 4=lactose reduced	1=liquid 2=powdered 3=evaporated	alpha
Eggs	dozens	n/a	n/a	n/a
Cheese	pounds	1=American 2=Colby 3=Swiss 4=Cheddar 5=mozzarella 6=Monterey Jack 7=Muenster 8=unknown	1=block 2=shredded 3=sliced 4=string cheese 5=cheese balls (mozzarella only)	not collected
Cereal	ounces	n/a	n/a	alpha
Juice/ Infant Juice	cans	1=orange 5=grapefruit 2=grape 6=lemonade 3=apple 7=pineapple 4=tomato 8=berry 9=mixed/unknown	1=frozen concentrate 2=canned single- strength	alpha
Peanut Butter	ounces	1=crunchy 2=creamy	n/a	alpha
Beans	pounds	1=navy 4=lima 2=pinto 5=lentil 3=black 6= unknown	1=dry 2=canned	alpha
Infant Cereal	ounces	1=rice 4=mixed 2=barley 3=oatmeal	n/a	alpha
Infant Formula ¹	cans	1=soy-based 2=milk-based	1=powdered 2=concentrate 3=ready-to-feed	alpha
Tuna ¹	cans	[white/light/dark packed in water or oil]	[solid, chopped, flaked, grated]	alpha
Carrots	pounds	n/a	[canned, fresh]	alpha
1 Foods issued in cans were identified by can size, in ounces.				