27300

BY THE U.S. GENERAL ACCOUNTING OFFICE

Report To The Committee On Agriculture, Nutrition, And Forestry United States Senate

WIC Evaluations Provide Some Favorable But No Conclusive Evidence On The Effects Expected For The Special Supplemental Program For Women, Infants, And Children

GAO analyzed the evaluations of the Special Supplemental Program for Women, Infants, and Children to determine whether they support assertions that have been made about the program's positive effects for five outcomes GAO found insufficient evidence for making any general or conclusive judgments about WIC's effectiveness overall. In a limited way, however, the information indicates the likelihood that WIC has modestly positive effects in some areas.

GAO assessed the problems in the state of evaluation efforts. That the program evaluations do not reveal whether WIC is having the effect intended by the legislation underscores the need to design and implement better studies. If the lessons learned from past efforts are heeded, it should be possible to produce information on the overall effectiveness of the WIC program.

GAO/PEMD-84-4
JANUARY 30, 1984

UNITED STATES GENERAL ACCOUNTING OFFICE
WASHINGTON, D C 20548

PROGRAM EVALUATION
AND
METHODOLOGY DIVISION

B-176994

The Honorable Jesse Helms
Chairman, Committee on
Agriculture, Nutrition,
and Forestry
United States Senate

Dear Mr. Chairman:

In your letter of June 30, 1983, you asked that we analyze the technical and methodological soundness of the evaluations of the Special Supplemental Program for Women, Infants, and Children (WIC) and that we assess the credibility of the assertions that have been based on them about the program's effects on certain aspects of the nutrition and health of mothers and their children. This report summarizes our review of the information and discusses our observations regarding what is known about WIC's effectiveness for the outcomes in which you expressed an interest.

As we arranged with your office, we are sending copies of this report to other interested congressional committees, to the U.S. Secretary of Agriculture, and to health-care professionals. Copies will be made available to others who request them.

Sincerely,

Eleanor Chelimsky
Director

GENERAL ACCOUNTING OFFICE
REPORT TO THE COMMITTEE ON
AGRICULTURE, NUTRITION, AND
FORESTRY
UNITED STATES SENATE

WIC EVALUATIONS PROVIDE SOME FAVORABLE BUT NO CONCLUSIVE EVIDENCE ON THE EFFECTS EXPECTED FOR THE SPECIAL SUPPLEMENTAL PROGRAM FOR WOMEN, INFANTS, AND CHILDREN

D I G E S T

The Special Supplemental Program for Women, Infants, and Children (WIC), sponsored by the Food and Nutrition Service of the U.S. Department of Agriculture, was established in fiscal year 1972 to provide food supplements and nutrition education in conjunction with health care to pregnant and postpartum women and to infants and children up to age 5 who have health and nutritional risks as well as low incomes. WIC's annual appropriation grew from $20 million in fiscal year 1974 to more than $1,160 million in fiscal year 1983. In fiscal year 1983, WIC served about 3 million participants.

WIC's proponents have cited its local, state, and national evaluations in support of their claims that WIC is unquestionably effective in improving the health of mothers and their children in specific ways. Others have criticized the studies as being so severely flawed methodologically that drawing conclusions from them is unfounded. The Chairman of the Senate Committee on Agriculture, Nutrition, and Forestry asked GAO to analyze WIC's evaluations to determine the strength of their evidence.

Specifically, the Chairman requested that GAO focus on WIC's effects on miscarriages, stillbirths, and neonatal deaths and on maternal nutrition. With regard to positive pregnancy outcomes, he asked GAO to review WIC's effect on "high-risk" mothers and to review the claims that the length of participation in WIC is directly related to positive outcomes. With regard to infants and children, GAO was asked to look at WIC's effect on the birthweights of infants and the claims that the program reduces the chances for anemia and mental retardation in infants and children.

OBJECTIVES, SCOPE, AND METHODOLOGY

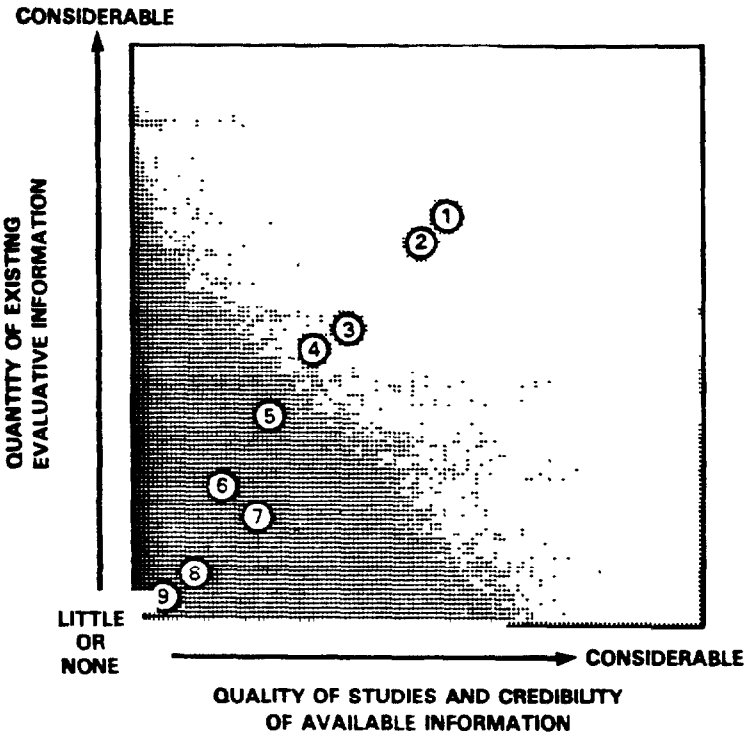
To find out what is known about WIC's effectiveness, GAO formulated specific evaluation questions; identified the evaluation reports that are relevant to those questions; reviewed them for their design, methodology, execution, and findings; rated them on their credibility and soundness; and analyzed their findings. GAO's bibliographic search and consultation with experts identified 61 evaluations relevant to the Committee's interests. (pp. 4-11; app. IV)

WHAT IS KNOWN ABOUT WIC'S EFFECTIVENESS

The accompanying chart displays GAO's assessment of the strength of the evidence in the WIC evaluation reports. To be able to say that supporting evidence is conclusive regarding a specific WIC outcome, GAO looked for evaluative information that was adequate in quantity and high in quality. The absence of topics in the unshaded area of the chart indicates that GAO finds no conclusive evidence of any kind about WIC's success or failure. Data on the birthweight question are substantial, but GAO finds that their quality is moderate. Findings relevant to the remaining questions are pushed toward the "gaps in knowledge" corner of the chart, indicated by the darker shading. In particular, GAO finds little or no information on mental retardation and on the separate effects of WIC's services for food supplements, nutrition education, and adjunct health care. In sum, GAO's critical review of the evaluation designs and their execution leads to the finding that the information is insufficient for making any general or conclusive judgments about whether the WIC program is effective or ineffective overall. However, in a limited way, the information indicates the likelihood that WIC has modestly positive effects in some areas.

Infant birthweights

Six of the WIC studies containing information about infant birthweights are of high or medium quality. They give some support, but [illegible] evidence for the claims that



LEGEND

- □ CONCLUSIVE EVIDENCE
- ◘ SOME OR MODERATE EVIDENCE
- ■ GAPS IN KNOWLEDGE

KEY:

1. INCREASE IN MEAN BIRTHWEIGHTS
2. DECREASE IN PERCENTAGE OF LOW-BIRTHWEIGHT INFANTS
3. EFFECTS, FOR HIGH-RISK GROUPS AND FOR THOSE PARTICIPATING LONGER THAN 6 MONTHS, ON BIRTHWEIGHTS
4. IMPROVEMENT IN MATERNAL NUTRITION
5. DECREASE IN INCIDENCE OF ANEMIA IN INFANTS AND CHILDREN
6. DECREASE IN INCIDENCE OF FETAL AND NEONATAL MORTALITY
7. EFFECTS, BY LENGTH OF PARTICIPATION AND FOR HIGH-RISK GROUPS, ON MATERNAL NUTRITION, FETAL AND NEONATAL MORTALITY, AND ANEMIA IN INFANTS AND CHILDREN
8. DECREASE IN INCIDENCE OF MENTAL RETARDATION IN INFANTS AND CHILDREN
9. EFFECTS OF THE THREE SEPARATE WIC COMPONENTS

studies, about 7.9 percent of the mothers in WIC had infants who were less than 2,500 grams at birth, compared to about 9.5 percent of the mothers who were not in WIC. This translates into the positive finding that, in the six studies, the proportion of infants who are "at risk" at birth because of low weight decreased as much as 20 percent. Average birthweights were between 30 and 50 grams greater for WIC participants, an increase of not more than 2 percent. Both WIC and non-WIC infants weighed about 3,200 grams on average, which is above the 2,500-gram boundary below which neonatal and infant health problems are expected. (pp. 12-24)

Fetal and neonatal mortality

The quality and credibility of the evaluative data on fetal and neonatal mortality are substantially lower than the data on birthweights. GAO rates the reports of WIC's favorable effects low in credibility and insufficient to support claims in either direction about WIC's ability to lessen the number of fetal and neonatal deaths. (pp. 24-25)

Maternal nutrition

On the improvements in maternal nutrition that can be attributed to WIC, the evidence is less strong in quality and quantity than that available for birthweights. There are six studies of moderate quality that differ in several ways, including how they ruled out alternative explanations and what measurements they reported. It is difficult to synthesize their results. Although some evidence does suggest that participating in WIC is associated with a better diet, greater iron levels in the blood and increased weight gain, it is inconclusive. (pp. 28-40)

Anemia in infants and children

GAO finds that the evidence is insufficient to support conclusively the assertion that WIC prevents anemia in infants and children. Limited evidence from two studies of only moderate quality suggests that WIC may be associated with improving the iron levels in their blood. This is also true with regard to children who are classified as anemic when they

Mental retardation in infants and children

There is no evidence on WIC's effect on mental retardation. No WIC evaluation has specifically addressed the incidence of mental retardation. One study focused on the cognitive development of WIC participants, but its favorable conclusions cannot be confidently attributed to the WIC program because of limitations in the study's design and execution. (pp. 48-49)

WIC's effect on different groups

WIC does appear to have greater positive effect on the birthweights of the infants of mothers who are teenagers or blacks or have several health- and nutrition-related risks. (pp. 19-23) However, the information on these differences with respect to WIC's effect on fetal and neonatal mortality, maternal nutrition, and anemia in infants and children is inconsistent and insufficient.

WIC's effect by length of participation

GAO finds some evidence that suggests that participating in WIC for more than 6 months is associated with increases in birthweights and with decreases in the proportion of low-birthweight infants. (pp. 23-24) Longer participation may improve the levels of iron in maternal blood. (p. 38) The greatest reductions in the incidence of anemia in children occurred during the first 6 months of participation. (pp. 46-47) None of this evidence is conclusive, however.

The effects of WIC's three separate components

There is almost no information about the separate effects of WIC's services for food supplements, nutrition counseling, and adjunct health care. Most of the evaluations determined who participated in WIC from unvalidated listings on the WIC roles and give no description of the WIC intervention being studied. The studies that do include data about WIC services do not systematically examine or discuss the separate effects of the three components.

THE CURRENT STATE OF WIC EVALUATION

No one study or group of studies provides the kind of evidence that can either confirm or refute claims that WIC is effective with respect to the outcomes of interest to the Committee. The evidence overall and on particular outcomes falls short of being conclusive.

One of the most important problems in the evidence of WIC's effectiveness is that it is not generalizable to the national WIC program. Funded by the federal government, WIC is administered at the state and local levels. WIC projects can vary considerably in the populations they serve and in the way they provide services. Since most evaluations have focused on only one or on a few projects or on specific geographic areas, their findings are restricted to the specific conditions of those projects or areas.

The shortage of credible evaluative information does not mean that WIC is ineffective; rather, it means that there is not enough clear and indisputable evidence to draw a firm conclusion about it. WIC may have certain positive effects on its participants, but the designs, sample sizes, and measures that have been used in evaluations so far have not always been sensitive enough to detect changes in the well-being of women and children that can be attributed to the WIC intervention. Even the findings that are methodologically the strongest--that is, that are statistically significant--await answers about their clinical meaning.

The studies that do not provide conclusive evaluative information about WIC's overall effectiveness often contain information that is nonetheless useful about WIC's implementation and about operational issues among the local programs.

Many of the documents GAO reviewed do not adequately describe the design, execution, and analyses that were used in the evaluation effort. With this information missing, it is difficult to determine the technical adequacy of the findings or the confidence that can be placed in the findings.

In this synthesis, GAO did not include findings from the clearly poor evaluations. They were so severely flawed that combining them with the findings from studies of high or moderate quality could be misleading.

The following methodological problems are noteworthy in WIC's evaluations:

--they lack research designs that are adequate for establishing a cause and its effect (such as a causal relationship between participating in WIC and a positive outcome);

--the indexes they use to measure nutrition are neither precise nor standardized, and experts do not agree on what the indicators of nutritional inadequacy are;

--the data are of questionable quality because collection and reporting are not sufficiently uniform or consistent;

--the evaluations do not present sufficient, technical details about the WIC interventions that were studied;

--they do not separate the effects of the individual WIC components or of WIC from the effects of other programs, nor do they analyze the relationships between a mother's nutrition, her pregnancy, and the health of her children during the early years of life;

--the evaluations do not build on past research and are not designed to enable subsequent studies to use their results. (pp. 56-57)

Despite these problems, progress can be seen in the improved designs and methodologies of various recent evaluation efforts. The national WIC evaluation that the Food and Nutrition Service has under way has placed considerable emphasis on reviewing past evaluation difficulties in order to guide the design of the new assessment.

Previous reviewers of WIC evaluation studies have offered conclusions ranging between two extremes. Either

--design and methodology problems and program complexity impose such severe constraints

that a meaningful overall assessment of the WIC program is not really possible or

--a substantial body of evidence from WIC evaluations now exists and indicates that the program is having a positive and significant effect on its participants.

GAO's position falls between these two extremes.

GAO finds some sound, but not conclusive, evaluative evidence of favorable program effects on birthweights and little credible evidence on several other measures of effectiveness. That the evaluations do not reveal whether WIC is or is not having the effect intended by the legislation underscores the need to design and implement evaluations that can provide the information that the Congress needs. GAO believes that the lessons learned from past evaluation experience will make it possible to produce this information.

Contents

ABBREVIATIONS

CDC	Centers for Disease Control
FNS	Food and Nutrition Service
GAO	U.S. General Accounting Office
NDDA	Nutrient Dietary Data Analysis
RDA	Recommended daily allowance
WIC	Special Supplemental Food Program for Women, Infants, and Children

CHAPTER 1

INTRODUCTION

The Special Supplemental Food Program for Women, Infants, and Children (usually referred to as the "WIC" program) was authorized by Public Law 92-433, a September 26, 1972, amendment to the Child Nutrition Act of 1966. WIC's purpose is to serve as an adjunct to good health care during critical times of personal growth and development, to prevent health problems, and to improve the health of the low-income citizens who are eligible to participate. According to the Omnibus Budget Reconciliation Act of 1981, supplemental foods, access to health care, and nutrition counseling are authorized and to be provided to eligible women who are pregnant, lactating, or up to 6 months postpartum and to infants and children up to age 5. Eligibility criteria include having inadequate income and special risk with respect to physical and mental health because of inadequate nutrition or health care or both.

The Food and Nutrition Service (FNS) of the U.S. Department of Agriculture provides funding for and gives general oversight and direction to the WIC program, which is administered by state health departments and approved local health clinics. At present, more than 1,500 local WIC projects operate through 84 state agencies and the Indian tribes. Over the program's years, the legislation and regulations have become increasingly specific so that WIC can help the most vulnerable, although the state WIC agencies differ in how they control the local agencies, and the local agencies differ in how they provide supplemental food and nutrition education and coordinate health care services. During those years, WIC's appropriation expanded from the $20 million of fiscal year 1974 to approximtely $1,160 million in fiscal year 1983. In fiscal year 1983, the program's participants included 633,440 women, 852,480 infants, and 1,477,040 children--about 2.96 million persons in all. Authorization for WIC expires at the end of fiscal year 1984.

The Congress has recognized the importance of WIC's evaluation in various hearings and, in November 1978 amendments to the Child Nutrition Act, has allowed for special funding for WIC evaluations. FNS has funded two large, nationally focused evaluations and several smaller evaluations. State and local agencies have also undertaken various assessments of the program. The agencies that administer WIC are required to maintain records on the participants in order to document the costs and benefits of the program.

The Senate Committee on Agriculture, Nutrition, and Forestry, having jurisdiction over WIC, has in recent years conducted hearings on the program's effectiveness. Both the Committee and Subcommittee on Nutrition have collected testimony citing WIC evaluations to support various positions. In some instances, evaluation findings have been presented in support of contentions that WIC is effective in improving maternal and child health. In other instances, the evaluations have been criticized as being

methodologically unsound and not national in scope and, therefore, as providing little evidence of WIC's effectiveness.

The Chairman of the Senate Committee asked us to make an objective analysis of WIC evaluations and to determine whether they support the claims that have been based on them. Specifically, he asked us to focus on the nutritional "status" of mothers and on whether WIC prevents miscarriages, stillbirths, and neonatal deaths. He asked us to examine the assertions that pregnancy outcomes are especially positive for "high-risk" WIC mothers and that they become more positive as the length of participation in WIC increases. With regard to infants and children, we were asked to analyze reports of the effect on birthweights and the assertions that the program reduces the chances of anemia and mental retardation.

WIC'S EVALUATION CONTEXT

In the past 10 years, WIC evaluators at local, state, and national levels have faced many challenges and difficulties in designing and implementing studies that could provide useful information. Some of the difficulties are inherent in the evaluation of social programs, and some are associated with the WIC program itself. Evaluation problems that WIC shares with other social interventions include (1) assigning persons randomly to comparison groups, partly because of the ethical issue entailed in withholding services from eligible recipients, (2) constituting adequate comparison groups of people not participating in the program, (3) generalizing findings from state or local studies to the nation because of differences in the way services are delivered and in the way the program is administered and operated, and (4) distinguishing the influence of the program on the participants from other factors that influence their well-being.

A number of evaluation difficulties are particular to WIC. (1) There is a lack of consensus about what the appropriate measures of outcome are. (2) Defining, standardizing, and using such measures are not done consistently. (3) Another difficulty is the disagreement about the underlying assumptions of the WIC program: some people say that undernutrition is a minor problem and does not justify nationwide programs. (4) Evaluating WIC's implementation is difficult, as when it is hard to determine whether food packages are consumed only by WIC participants or included in the preparation of family meals. (5) Furthermore, changes in the program, such as the shift in 1979 toward greater emphasis on serving the persons who have the greater health risks, have complicated evaluation activity. (6) Evaluating the program with an interdisciplinary team made up of at least medical, nutritional, and statistical experts is valuable but not always possible.

Given all this, it is not surprising that WIC's evaluation studies vary greatly in their designs and methods and in the questions they address, the program components and participants they study, the measurements of effects they use, and the periods

for which they collect data. Many evaluations have assessed more than one aspect of WIC, but no one study has addressed all aspects.

PREVIOUS REVIEWS OF WIC EVALUATIONS

The reviewers of WIC evaluation studies have identified different problems with design or methodology that affect the credibility of the findings, but they do not always agree about the severity of the flaws. Furthermore, since WIC evaluations usually look at more than one aspect of the program, several reviewers may refer to one study but base their judgments about it on different parts of it. Reviewers looking at the same set of studies have reached quite different conclusions about WIC's effectiveness, ranging from an opinion, at one extreme, that design and methodology problems and the program's complexity have imposed such severe constraints that a meaningful overall assessment of the WIC program is not possible to the opinion, at the other extreme, that substantial evidence from WIC evaluations indicates that the program is having a significantly positive effect on its participants.

Reviewing WIC evaluation efforts 10 years ago, we pointed to potential problems in the methodology of a medical evaluation that FNS had proposed (GAO, 1973).[1*] The legislative history had indicated that a medical evaluation should assess WIC's effects on the mental development of infants, but we found that physicians and other experts in nutrition and health care questioned whether the proposed methodology could do this. In 1975, we reported on the same study, which was being carried out by Edozien and colleagues at the University of North Carolina for FNS; we discussed problems in assessing the medical benefits of nutritional assistance and identified the specific weaknesses of the evaluation. We concluded that the credibility of its results would be questionable and pointed out that medical evaluations require safeguards to insure adequate methodology and collection periods. Our position was that medical evaluations are useful for decisionmaking only if the data are sound (GAO, 1975), and this position was reaffirmed in a report 2 years later on national nutrition issues (GAO, 1977).

In 1979, we reviewed the available WIC evaluations and concluded that no adequate assessment of WIC's overall results and benefits had been made (GAO, 1979). We included the Edozien study, an analysis of nutrition surveillance data by the Centers for Disease Cotnrol, and an FNS compilation of 12 studies from 10 states and the Virgin Islands. We reported that the results from the state studies could not be projected to the nation and that their quality had not been independently evaluated.

In summary, we have concluded in the past that WIC evaluations are generally disappointing in not providing information about

---

*Notes to each chapter are at the end of that chapter. Page references for the notes are in the table of contents.

WIC's effectiveness that is useful for making national decisions. We have also been cautiously optimistic that better evaluations in the future may be able to provide the Congress with the information it needs.

OBJECTIVES, SCOPE, AND METHODOLOGY

Our objectives were to systematically assess the technical and methodological soundness of all major evaluations of the WIC program, including the studies that were completed after our 1979 review, and to determine the credibility of the claims that have been based on them about the program's effect on certain aspects of the nutrition and health of mothers and their children. It was not our objective to conduct a new study of the WIC program.

The specific topics we were asked to examine in the WIC evaluations are on their findings about the program's effects on (1) improving maternal nutritional status and health, (2) increasing birthweights, (3) reducing the incidence of miscarriage, stillbirth, and neonatal death, (4) reducing the incidence of anemia in infants and children, and (5) reducing the incidence of mental retardation. We were also asked to determine if WIC's effectiveness differs for the various groups of women and children who are considered to have the greater medical risk or for the participants who receive WIC services for longer periods or for the three WIC services--supplemental foods, adjunct health care, and nutrition counseling.[2] (We have reprinted the congressional letter requesting this report in appendix I.)

To produce a draft analysis in 6 months, as we were asked to do, we conducted an evaluation synthesis. By this method, we were able to identify and assemble existing evaluation studies, assess the appropriateness and quality of their design and execution, aggregate their results, determine the level of confidence that can be placed in these findings, and highlight the gaps in the accumulated knowledge (see GAO, 1983).

Developing our evaluation questions

The first step of our review was to determine as precisely as possible which questions would produce answers to the congressional request for information. In table 1, questions 1-5 focus on the overall effectiveness of the WIC program for five outcomes: birthweights, mortality, maternal nutrition, infant and child anemia, and mental retardation. Questions a-c focus on differences in effectiveness with regard to the participants who are served, their length of participation, and the program's components. We applied questions a-c to each of the questions on overall effectiveness in order to determine if there are certain circumstances in which WIC works especially well or not. This is not an exhaustive list of questions for assessing every aspect of the WIC program. The questions are the ones we considered the most relevant to the congressional inquiry.

Table 1

Evaluation Questions Regarding the WIC Program's Effectiveness

On overall effectiveness	On differences in effectiveness
1. Does participating in WIC affect birthweights?	a. Does WIC's effectiveness differ for groups of participants with different risks regarding nutrition or health?
2. Does participating in WIC prevent miscarriages, stillbirths, and the mortality of the newborn?	b. Does WIC's effectiveness differ by the length of participation?
3. Does participating in WIC affect the health and nutrition of pregnant and lactating women?	c. Which of the three WIC services--food supplementation, adjunct health care, and nutrition education--is most effective?
4. Does participating in WIC affect the incidence of anemia in infants and children?	
5. Does participating in WIC affect the incidence of mental retardation in infants and children?	

Identifying the evaluation studies

Since our objective was to identify all documented WIC evaluation studies at national, state, and local levels, we cast a broad net in order to find not only the most frequently cited published WIC studies but also unpublished evaluations at universities and contract research organizations and in the state and local WIC programs. Time restricted us to assessments of the WIC program. We excluded other papers and books that focus generally on nutrition during pregnancy and early childhood and on programs similar to WIC.

We were well assisted by the Office of Analysis and Evaluation and the WIC program staff at FNS. They prepared two chronological bibliographies on the WIC program and made available to us the copies of the evaluation reports that they had on file. Staff of the Congressional Research Service also identified WIC-related reports and legislative testimony. We used all these documents to track down others that are referred to in them.

Finding several evaluation studies that had not been published, we became concerned that we might be missing other

unpublished studies, particularly at state and local levels, and that this could bias the results. Therefore, we made contact with 88 nutritionists, health professionals, researchers, evaluators, and program administrators who were familiar with WIC. Forty-nine experts responded to our request to review our bibliography and to identify other reports and studies that we did not have, and 15 identified additional sources. We found more then 100 documents with this process.

Assessing the evaluation studies

In reviewing the evaluation documents, we made judgments in determining the relevance of each report and in assessing its technical quality. To determine relevance, we reviewed all the documents to see how each addresses our specific evaluation questions. We eliminated those that emphasize special issues not related to our questions, and they do not appear in our bibliography. For example, studies about the breastfeeding practices of the women in the WIC program or the availability of transportation to WIC clinics were not considered relevant. As we show in the accompanying display, we found that many of the remaining 86 documents could not be included in the evaluation synthesis:

	Number
Information related but not directly relevant to outcome of interest	19
Summaries, reviews, or critiques of WIC evaluations	13
Relevant and included in our synthesis	54
	86

We excluded 19 documents that do not contain information about WIC outcomes that would answer the evaluation questions. These are mostly administrative and descriptive reports on the characteristics of the WIC population, certain aspects of WIC service delivery, and management questions regarding quality and efficiency in program operations. We did not formally include these documents in our synthesis, but many of them gave us insight into the WIC program that was useful for interpreting claims about WIC's effectiveness, and they are included in our bibliography. We found 13 more documents to be summaries, reviews, or critiques of WIC evaluations. We used them to inform our judgment about the WIC evaluations but did not formally include them in our synthesis, although they are in our bibliography. The remaining 54 documents, also in our bibliography, are relevant because they contain information on one or more of our questions about health and nutrition outcomes. We found a total of 61 studies in these documents (a few contain more than one study).

To make critical assessments about the methodological quality of the evaluation studies, a team of four raters gave them an overall quality rating that reflects the level of confidence that the raters placed in the reported results. We prepared written

summaries and proposed our own ratings. The major studies were reviewed by more than one rater. Using the summaries and the reports, the raters discussed the strengths and weaknesses of the studies relative to one another and concurred in the final ratings by resolving their disagreements.

Each study's findings on the topics of our evaluation questions were rated according to credibility on the nine-point scale from low (1, 2, 3) to medium (4, 5, 6) to high (7, 8, 9). The criteria we used to judge the appropriateness and soundness of the methodology underlying an evaluation's findings and its implementation included

--the strengths and weaknesses of the research design,

--the adequacy of the comparison groups,

--the type of sample and the sample size,

--the reliability and the validity of the measures used,

--the methods and conditions of data collection,

--the appropriateness of the statistical analyses,

--the disclosure of problems and limitations,

--the completeness of the information reported for understanding and interpreting the data, and

--the relationship of the findings to the conclusions.[3]

For studies that address more than one of our questions and for studies that adopted several investigative strategies, we reviewed each relevant element of the evaluation individually. We found 178 relevant elements among the 61 studies--that is, 178 individual assessments of different aspects of the WIC program that are relevant to the congressional inquiry.

In appendix III, we list the 61 studies, indicate the relevant elements, and note their ratings. We rated 37 elements at four or higher, taking their findings to be relatively credible. We rated the 141 other elements at lower than four. These suffered most often from one or more of the following problems:

--the description of the evaluation design and execution was so brief that an informed judgment could not be made;

--so much data were missing that it was impossible to interpret the results,

--comparison groups were not, in fact, comparable, and

--the analyses of findings were incomplete or incorrect.

Because these problems may mean that results are open to rival interpretations, we did not use the elements with ratings lower than four in our final synthesis. Table 2 summarizes all this.

Table 2

Our Rating of the Quality and Credibility of the Findings of 61 WIC Evaluations

Question on	No. of relevant elements[a] Rated high or medium	Rated low
1. Birthweight		
Overall effects	6	33
Different effects for		
a. Risk groups	6	12
b. Length of participation	4	12
c. WIC components	1	12
2. Perinatal and neonatal mortality		
Overall effects	0	12
Different effects for		
a. Risk groups	0	4
b. Length of participation	0	4
c. WIC components	0	0
3. Maternal nutrition		
Overall effects	6	18
Different effects for		
a. Risk groups	4	3
b. Length of participation	2	1
c. WIC components	1	4
4. Anemia in infants and children		
Overall effects	2	25
Different effects for		
a. Risk groups	2	3
b. Length of participation	2	4
c. WIC components	1	2
5. Mental retardation		
Overall effects	0	1
Different effects for		
a. Risk groups	0	1
b. Length of participation	0	0
c. WIC components	0	0
Total	37	141

[a]Total number of elements. An evaluation is included more than once if its elements are relevant to more than one evaluation question.

Making a synthesis of the results

In synthesizing the results of our analysis for each congressional question, we identified the major findings of the studies rated medium or high in quality and looked for patterns applicable to each evaluation question. We took into account the limitations of the evaluations, such as the sample sizes, the composition of comparison groups, and attrition, as they might have affected the findings. We also took into account the quantity of the evidence and whether it accumulated from study to study. In this way, we assessed both quality and quantity in order to determine the strength of the evidence.

When the data were sufficient and appropriate, we applied quantitative indicators to summarize the estimates of WIC's effects. Several quantitative indicators can be used for this, each providing slightly different information. No one index is entirely suitable as a means of summarizing data. Therefore, where it was possible to characterize the quantitative differences between WIC and non-WIC groups, we calculated the average raw difference between the two groups and the percentage difference. (Our calculation and the meaning of these indicators are described in appendix IV.) Since each indicator translates the findings of each study into a common measure, we were able to summarize the birthweight data from the evaluations included in our synthesis. We discuss these figures in terms of averages, variability, and range (lowest to highest).

The benefits and limitations of our method

An evaluation synthesis is necessarily dependent on the quantity and quality of the data and analyses in the available studies. We relied on documents that could be readily found--items in the FNS files, journals and books, dissertations on file in libraries, and the like. Some of the reports were less than complete. The time restrictions for this review did not allow us to make contact with all authors to clarify ambiguities, request additional information, or obtain primary data. Therefore, we relied primarily on information as it has been reported in the published and unpublished sources we examined.

It is possible that we did not uncover all the available documents, but our careful bibliographic search and survey of experts suggest that any gap is narrow. We believe that we have identified the documentation for all the major, completed evaluation studies of WIC's effectiveness. Our review was performed in accordance with generally accepted auditing standards.

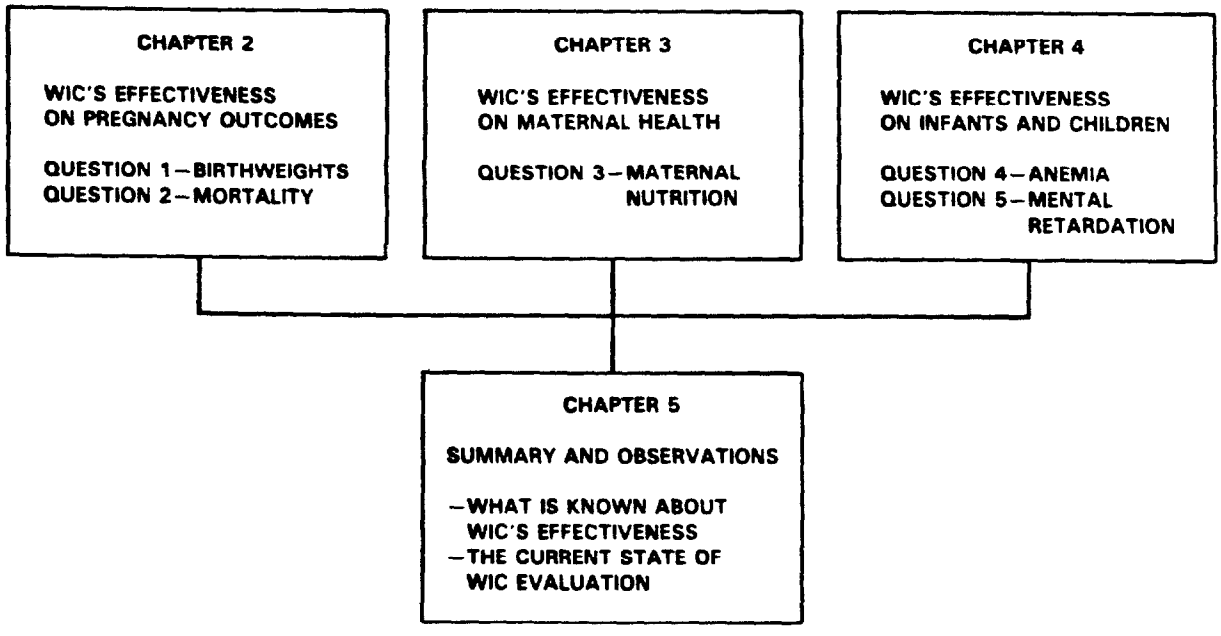
The major benefit of the evaluation synthesis is that, beyond the literature review, it analyzes the quality of each evaluative finding in terms of the evidence supporting it and yields refined information about what is known on a particular topic at a

particular time. General knowledge is strengthened by the findings of several soundly designed and well-executed evaluations when they are consistent, even though they may have used different methods. No matter how high its quality, a single evaluation can rarely do this. Concluding from an evaluation synthesis that many evaluations were not soundly designed or well executed is also beneficial, however. This is partly because the synthesis identifies areas for which there is no firm basis for making policy decisions and, further, because it identifies the problem areas that can be addressed in future evaluations.

The organization of this report

The sequence of topics in the chapters in this report, shown in figure 1, reflects the amount of evaluative information that we found available for answering the five questions on outcomes. Our analysis of WIC's effects on pregnancy outcomes--that is, birth-

FIGURE 1
THE ORGANIZATION OF THIS REPORT



weights and fetal and neonatal mortality--is in chapter 2. In chapter 3, we discuss the evidence about WIC's effects on maternal nutrition. In chapter 4, we describe the evidence on WIC's effectiveness in reducing the incidence of anemia and mental retardation in infants and children. In each chapter, we report on our analysis of the information about overall effects (questions 1-5) and follow this with data on differences in effects (questions a-c). Chapter 5 is a summary of our observations about what is known

about WIC's effectiveness and the state of evaluation of the WIC program.

NOTES

[1]The abbreviated bibliographic citations in parentheses in this report are given in full in appendix II. For example, "GAO, 1973" is the GAO (U.S. General Accounting Office) report entitled Supplemental Food Program and issued in September 1973.

[2]The National Research Council's 1981 report entitled Nutrition Services in Perinatal Care sets forth three sets of maternal risk factors. (1) Risk factors at conception are (a) being adolescent (that is, becoming pregnant less than 3 years after starting to menstruate, (b) having had three or more pregnancies within 2 years, (c) having a history of abortions, pregnancy complications, low-birthweight infants, or perinatal loss, (d) living in economic deprivation, (e) smoking more than 20 cigarettes a day, (f) having a history of binge drinking or of chronically drinking more than 5 ounces of whiskey a day or its equivalent in beer or wine, (g) being addicted to drugs, (h) having unusual dietary practices, (i) having systemic diseases chronically, and (j) weighing less than 85 percent or greater than 120 percent of the standard weight for height. (2) Risk factors during pregnancy are (a) having a hemoglobin level lower than 11 grams per deciliter or a hematocrit level lower than 33 percent (we explain hemoglobin and hematocrit levels in chapter 3, note 5) and (b) gaining weight at less than 1 kilogram per month or, possibly in association with the retention of fluids, at more than 3 kilograms per month. (3) The risk factor following pregnancy involves the nutritional demands of lactation.

[3]A discussion of standards for program evaluation is in the Evaluation Research Society Standards Committee 1982 document cited in appendix II.

CHAPTER 2

DOES PARTICIPATION IN WIC AFFECT PREGNANCY OUTCOMES?

A central premise of the WIC program is that improving the nutrition of pregnant women whose pregnancy outcomes are at risk will help them give birth to healthier infants. While a number of indicators could be used to determine if WIC has had its intended effect, we focused on two: the birthweights and the fetal and neonatal mortality of infants. Birthweight is the most commonly evaluated outcome in WIC studies. The relatively low incidence of death in the several weeks before and after birth has deterred investigation of the WIC program's effect in this area.

The availability of data on the birthweight question is considerable. This evidence indicates that for some segments of the population, WIC can have a direct, positive effect on birthweight. However, our assessment of the quality and credibility of both the data and their analyses does not lead us to the opinion that the strength of that evidence is "conclusive."

We used six studies to analyze birthweight data. Five showed that participation in WIC is associated with a decline in the proportion of low-birthweight infants--that is, infants born weighing less than 2,500 grams. About 7.9 percent of the women who were participating in WIC were reported as having had low-birthweight infants, compared to about 9.5 percent in non-WIC comparison groups. As we discuss in this chapter, we estimate that WIC decreases the proportion of low birthweights for infants born to women eligible for WIC by 16 to 20 percent. WIC's effect on mean birthweights also appears to be positive. A reasonable estimate is that the benefit is in the range of 30 to 50 grams, an increase in mean birthweight of between 1 and 2 percent. There is additional evidence that certain high-risk groups within the eligible population derive proportionally more benefit from the program than other groups.

WIC'S EFFECT ON BIRTHWEIGHTS

The evaluations that examined the effects of WIC on birthweights used two measures for summarizing the evidence: average birthweight and the proportion of the sample (or the population) whose birthweight is below a critical weight. Since these measures present different types of information, we discuss them separately. We have not addressed the clinical significance of the reported results. For example, we have not established whether a 50-gram increase in birthweights because of participation in WIC is clinically meaningful. Most experts agree that 2,500 grams is the weight below which infants are most likely to have health problems.

The application of our questions to the evaluation studies

On WIC's effect on birthweight, we asked four questions. First, is there an effect? For example, have studies shown conclusively that the infants of the participating mothers weigh more at birth than the infants of comparable mothers who did not participate? Second, does the program have a greater effect on women who are considered to have greater health risks? Third, does the length of participation in the program make a difference? Fourth, is it possible to attribute a specific effect to the food supplement, nutrition education, and coordinated prenatal medical care components of WIC?

We applied these questions to the 39 studies reporting that WIC affects birthweight overall. All 39 addressed the birthweight question; 18 addressed the question about high-risk groups, 16 examined the length of participation in WIC, and 3 reported evidence on the effects of WIC's components on birthweight. Our rating system gave us confidence in the conclusions of only 6 of the studies. The remaining 33 were based on casual observations or were poorly documented or are substantially flawed. Table 3 on the next page indicates the methods of the 6 analyses and the key features of their considerably different evaluation designs.[1]

In brief, the 6 studies report evidence on the effects of WIC on mean and low birthweights. They also report on WIC's effect on at least one high-risk condition. The effect of the length of participation on mean or low birthweight was explicitly examined in 3 of the evaluations. One study directly examined the influence of the WIC program's components on birthweight.[2]

The relevance of mean and low birthweights as indicators of WIC's effects

As indicators, mean and low birthweights summarize the same data differently. Mean birthweight is the average birthweight of the infants in a group. The low-birthweight index is the percentage of infants in a group who weigh less than a specified weight.[3] For evaluating the effectiveness of WIC, one or the other may be more important, depending on the question that is being asked. If the overall influence of WIC is the question, then the average or mean is a suitable index. A difference in the average birthweight of the infants of WIC and non-WIC mothers gives an estimate of the extent to which the program has an effect. If the incidence of low birthweight is the question, then the low-birthweight index is more suitable for analysis.

Because the low-birthweight index focuses on the proportion of individuals weighing less than a certain weight, it does not necessarily reveal anything about average birthweight. In other

Table 3

The Focus of the Analysis and Design of Evaluations of WIC's Effect on Birthweight

Study	Overall effects	Risk factors	Months of participation	WIC components	Design remarks
Kotelchuck	Mean, % low	Age, race, education, marital status	1-3 4-6 7-9	Not analyzed	Matched groups, large sample, retrospective data, implementation data, statistical control
Metcoff	Mean, % low	Age, race, income, other low-birthwt infants	4-5	Not analyzed	Random control group, moderate sample, implementation collection, implementation data, statistical control
Stockbauer	Mean, % low	Age, race, medical status	Up to 3 3-5 6+	Not analyzed	Matched groups, large sample, retrospective data, implementation data, multiple statistical models, statistical control
Silverman	Mean, % low	Age, race, other low-birthwt infants, prepregnant wt	Not analyzed	Not analyzed	Pre- vs. post-WIC quasi-time series, retrospective data, multiple statistical models, statistical control
Bailey	Mean, % low	Smoking	Not analyzed	Not analyzed	Comparison group of women eligible for WIC residing in non-WIC area, small sample
Kennedy	Mean % low	Biological and social variables	1-3 4-6 7-9	Nutrition counseling, prenatal care	Unmatched groups, large sample, retrospective medical data, implementation data, multiple statistical models, data quality control

words, giving WIC services to women with the greatest risk may have only minimal influence on the average birthweight of the infants in the group and yet have substantial influence on the low-birthweight index. One implication of this distinction is that, with effective targeting, one might expect to observe greater differences between groups on the low-birthweight than on the mean-birthweight index. On both indicators, the evidence of WIC's effects can be summarized quantitatively.

The overall effect on mean birthweight

Table 4 on the next page gives the key results of the six reports on birthweight and our calculations of average difference, percentage difference, and statistical significance. For all six studies, the mean birthweight exceeds 3,000 grams. Further, five of the studies report higher mean birthweights for WIC participants than for their comparison groups. In four studies, these differences are statistically significant; they are marginally significant in one and not significant in one.

For WIC and non-WIC groups, the simple average of the six means shows a 49.6-gram upward difference for the infants of WIC participants compared to the infants of women in the non-WIC groups. In terms of a percentage difference, the 49.6 grams translates into a 1.55-percent average difference. Study by study, the difference between WIC and non-WIC birthweights ranges from minus 1.4 percent to plus 3.9 percent. The 3.9-percent figure represents Kennedy's unadjusted, and frequently cited, average difference of about 123 grams.

Looking at the sample size in each study reveals considerable variation, ranging from more than 6,000 infants in the WIC group to as few as 37. Since the precision of an estimate is related to the size of the sample, we calculated a weighted average birthweight as a way of obtaining an aggregate figure that accounts for sample size. After weighting each group's average by its sample size, we found that the apparent advantage of WIC participation is only 31.3 grams more at birth. The corresponding weighted figure for the percentage difference is +0.97, which represents about two-thirds of the simple average reported among all six studies.

Since sample size is only one of the many ways in which the studies differ, the weighted average is not necessarily a true summary of the data. It is probably more reasonable to consider it as a plausible lower estimate of WIC's effects. If the health and nutrition of the compared groups within each study were indeed comparable, then the overall pattern of evidence from the six studies suggests these general conclusions:

--on the average, there appears to be a positive benefit from WIC participation; a reasonable estimate is that the average birthweight of WIC infants is higher by somewhere between 30 and 50 grams;[4]

Table 4

Mean Birthweight Quantitative Summary

Study	Year and location	Reported birthweight (grams)[a] WIC	Non-WIC	Quantitative indicators Raw difference	% difference[b]	Statistically significant
Kotelchuck	1978 Mass.	3,281 (4,126)	3,260 (4,126)	21.0	0.6	Marginally
Metcoff	1980-82 Oklahoma City	3,254 (238)	3,163 (172)	91.0[c]	2.9	Yes
Stockbauer	1979-81 Mo.	3,254 (6,657)	3,238 (6,657)	16.0	0.5	Yes
Silverman	1971-77 Allegheny County, Pa.	3,189 (1,047)	3,095 (1,361)	94.0	3.0	Yes
Bailey	1980 2 Fla. counties	3,229 (37)	3,276 (42)	-47.0	-1.4	No
Kennedy	1973-78 Mass.	3,261.4 (897)	3,138.9 (400)	122.5	3.9	Yes
Summary						
Average		3,244.7	3,195.1	49.6	1.55[d]	
Weighted average[e]		3,257.8	3,225.9	31.3	0.97[d]	
Range lowest		3,189.0	3,095.0	-47.0	-1.4	
highest		3,281.0	3,276.0	122.5	3.9	

[a]The numbers in parentheses are sample sizes.
[b]Raw difference divided by non-WIC birthweight.
[c]Adjusted.
[d]Average raw difference divided by average non-WIC birthweight.
[e]Each mean is weighted by the number of participants or controls in its group and an overall average is obtained by dividing by the total number of participants or controls in the six studies. The raw difference is based on the total of participants and controls.

Table 5

Percentage Low-Birthweight Quantitative Summary

Study	Year and location	Reported low birthweight (percent)[a] WIC	Non-WIC	Quantitative indicators: Raw difference	% difference[b]	Statistically significant
Kotelchuck	1978 Mass.	6.9 (4,126)	8.7 (4,126)	-1.8	-20.7	Yes
Metcoff	1980-82 Oklahoma City	8.7 (242)	6.9 (174)	+1.8	+26.1	No
Stockbauer	1979-81 Mo.	8.5 (6,657)	9.4 (6,657)	-0.9	-9.6	Yes
Silverman	1971-77 Allegheny County, Pa.	9.7 (1,047)	13.0 (1,361)	-3.3	-25.4	Yes
Bailey	1980 2 Fla. counties	5.4 (37)	9.5 (42)	-4.1	-43.1	No
Kennedy	1973-78 Mass.	6.0 (833)	8.8 (375)	-2.8	-31.8	Yes
Summary						
Average		7.53	9.38	-1.85	-19.7[c]	
Weighted average[d]		7.92	9.50	-1.58	-16.6	
Range lowest		5.4	6.9	1.8	+26.1	
highest		9.7	13.0	-4.1	-43.1	

[a]Low birthweight = less than 2,500 grams. The numbers in parentheses are sample sizes.
[b]Raw difference divided by non-WIC low-birthweight percentage.
[c]Average raw difference divided by average non-WIC low-birthweight rate.
[d]Each birthweight rate is weighted by the number of participants or controls in its group and an overall average is obtained by dividing by the total number of participants or controls in the six studies.

--relative to the average weight in grams of the non-WIC infants, the difference of 30-50 grams translates into a difference of between 1.0 and 1.6 percent in average weight, although the greatest reported difference is 3.9 percent;

--the average birthweight of WIC infants and comparable infants whose mothers are eligible for WIC is approximately 3,200 grams, or about 7 pounds.

The overall effects on low birthweight

Table 5 provides findings and quantitative indicators for low birthweight similar to those presented in table 4 for mean birthweight. For WIC participants, the low-birthweight percentages range from the low of 5.4 percent to the high of 9.7 percent. In contrast, the percentages of low birthweights for non-WIC groups range from 6.9 percent to 13.0 percent. These differences may reflect differences in the populations that were sampled. Study by study, the raw difference between WIC and non-WIC infants weighing less than 2,500 grams at birth ranges from +1.8 percent to -4.1 percent. As table 5 shows, five studies report that the percentage of low-birthweight infants is smaller for WIC than for non-WIC groups. In the four studies with large sample sizes, the differences are statistically significant.

The simple average of all the studies for low-birthweight rate is 7.53 percent for WIC participants and 9.38 percent for their non-WIC counterparts, a difference of 1.85 percent. This represents an average reduction of 19.7 percent in the proportion of low-birthweight infants when participation in WIC is available.

Since the sample sizes are quite large in some studies and small in others, since the simple average disregards the size and composition of the sample and treats each study equally, and since the accuracy of what is reported depends on sample size and composition, we weighted the results for the low-birthweight index. Our quantitative indicators are given in table 5. We did not correct for sample composition. When the sample size is taken into account, for low-birthweight rates the difference between WIC (7.92 percent) and non-WIC (9.50 percent) drops to -1.58 percent. This reflects a 16-percent reduction in the low-birthweight rate that may be attributable to participation in WIC.

As with the mean-birthweight summary, this type of weighting yields a lower estimate of WIC's overall effect on low birthweights. Weighting by sample size gives greater emphasis to the larger, statewide studies that involved more clinics and a wider diversity of women. The results of studies with smaller sample sizes that may have had better research designs, more homogeneous groups of individuals, and similar treatments are given less emphasis.

The evidence on low birthweight as it is developed in the six studies suggests the following general conclusions:

--participation in WIC appears to have the beneficial effect of reducing the incidence of low birthweights; a reasonable estimate is that the proportion of low-birthweight infants among all who are born declines by 1.58 to 1.85 percentage points;

--relative to the average low-birthweight rate, the 1.58 to 1.85 percentage-point improvement means that 16.6 to 19.7 percent fewer of the infants who are born are born at the low birthweight when their mothers participate in WIC.

WIC's effect on high-risk participants

WIC's eligibility requirements are intended to insure that the women who participate in the program are those who are "at risk" of having poor pregnancy outcomes. However, "risk" is variously defined, and its factors may be classified as socio-economic (family size, marital status, education, income, and so on) or as related to maternal health (chronic disease, low weight before pregnancy, habits of smoking or taking drugs, a history of low-birthweight infants in the family, age, and so on).

Race is generally not considered a risk factor but is sometimes used as a surrogate in analyzing socioeconomic factors. That is, race is used as a way of finding high-risk subgroups, and so is age. Thus, the women who are generally considered to be the most likely to have poor pregnancy outcomes are black and either in their teens or 35 or older. In the discussion that follows, we use the same terms for identifying race as the writers of the evaluation studies we reviewed. We do not assume that the use of the terms "white," "black," "nonwhite," and "nonblack" are totally comparable among the studies.

Because the data we found on particular risk groups are few and not wholly comparable from study to study, we have substantially less confidence in them than in the mean and low birthweights. The data are insufficient for a quantitative synthesis. We analyzed four studies that examined age and race as affecting mean and low birthweights.

Age-related effects

There is some evidence that participation in WIC reduces the incidence of low birthweights for infants born to teenagers. We have been unable to find enough evidence to draw even preliminary conclusions about the other age groups. Kotelchuck found that teenage mothers appear to benefit from WIC. The increase in the average birthweight of infants of mothers 19 and younger exceeded the study average, and the younger the mother, the greater was

Table 6

WIC's Effect on Birthweight by Mother's Age and Race[a]

Study		Mean birthweight	Low birthweight
Age			
Kotelchuck	Up to 15	+	+
	Up to 17	+	+
	Up to 19	+	+
Stockbauer	Up to 18		+
	34+	++	+
Silverman	Up to 15	[b]	+
	16-20	[b]	++
	26-30	[b]	+
Race			
Kotelchuck	Black	+	+
	White	+	++
	Hispanic[c]	++	++
Metcoff	Black	++	
	Nonblack	+	++
Stockbauer	Nonwhite	++	++
Silverman	Nonwhite	[b]	++
	White	[b]	+

[a] + = "observed benefit"; ++ = statistically significant benefit.
[b] Not available.
[c] Hispanic population is a component of white population.

the increase in average birthweight. The low-birthweight rate among younger mothers also seemed to benefit, especially for the mothers 17 and younger for whom low-birthweight rate was 3.2 percentage points lower than for non-WIC teenagers and lower than the average 1.8-percentage-point difference for the study. Kotelchuck found no differences in the mean birthweights of infants of mothers 35 and older. (See table 6.)

Stockbauer found that, among white mothers in the non-WIC comparison group younger than 18, infants had a higher mean birthweight than for the WIC participants. His results for women older than 34 were almost the reverse. The WIC group had a mean birthweight that was significantly 6.3 percent greater than that of the non-WIC comparison group.

Our analysis of low birthweights suggests that this may be a pattern. Mothers younger than 18 appear to benefit from WIC, but the difference is not statistically significant. Mothers older than 34 appear to benefit from WIC, but the difference is statistically significant only for the nonwhite groups. Metcoff presented no data on age but reported a significant correlation between birthweight and age and education but not income. Silverman reported a significant difference in the low-birthweight rate for women between 16 and 20. He found that these WIC participants had a 5.1-percentage-point advantage in the lower proportion of infants born to them at low birthweights compared to non-WIC mothers in the comparison group.

Race-related effects

There is some evidence that black women who participate in WIC give birth to infants with a higher mean birthweight and have a lower proportion of the infants who are born at the low birthweight than comparable black women who do not participate (see table 6). Kotelchuck reported that the mean birthweight of infants born to black WIC participants was 37.8 grams greater than for the comparison group. Although this is more than twice the 16.1-gram advantage reported for white participants, neither figure is statistically significant. Kotelchuck's data show a significant difference of 1.7 percentage points fewer in the proportion of infants at low birthweight for white WIC participants and an even greater, but not significant, difference of 2.1 percentage points fewer for black WIC participants.

Analysis of data for hispanic women shows significant increases in mean birthweight. Their infants weighed an average of 70.8 grams more than those in one comparison group and 106.1 grams more than those in a more closely matched group. Adjusting these data for the age of the fetus when it is born reveals a statistically significant mean-birthweight advantage from WIC participation of 39.7 and 102.3 grams, respectively. The comparison between WIC participants and the unmatched group showed a significant 2.5-percentage-point advantage in the low-birthweight rate from WIC participation.

Silverman reported a significant difference of 4.2 percentage points in the proportion of infants at low birthweights for black mothers in WIC compared to black mothers not in WIC. Metcoff found strong evidence favoring WIC participation for black women. The difference in birthweight for the infants of black WIC mothers compared to the infants of similar black mothers not in WIC was 199 grams. The advantage of being in WIC for nonblack mothers was a difference of 118 grams in birthweight.

Support for the belief that black women have greater risk than white women was developed by Stockbauer in a study of the Missouri WIC program. His data show that nonwhite women older

than 34 gained a significant advantage by participating in WIC. The mean birthweight of their infants was significantly higher, and their proportion of infants at low birthweights was significantly lower, compared to mothers in a nonwhite, non-WIC group. Stockbauer did not define race itself as a risk factor, suggesting instead that the risk that nonwhite WIC participants have may be more influenced by nutrition than that of white WIC participants. He created a broad "risk measure," using information from the Missouri birth certificates. His risk factors include age and indicators of poor health. Using his broad measure, he found a significant difference in the low-birthweight rate for nonwhite mothers. The rate was 12.9 percent, almost 25 percent lower than the 16.7-percent rate for the at-risk non-WIC comparison group.

With the exception of age and race, we found few results reported for socioeconomic risk factors. Metcoff investigated the relationship between birthweight and income and did not find a significant correlation. He concluded that using a poverty income as the sole certifying criterion for admitting pregnant women to WIC is unlikely to raise birthweights.

Health-related effects

Evaluation studies provide little analysis of the effect of WIC on women with health-related risks. What is available suggests that black women with health-related risks benefit from participating in WIC. Participating in WIC may mitigate some of the effect of a mother's smoking, demonstrably harmful to infant birthweights.

Stockbauer's analysis of health-related risk factors appears to be the most comprehensive. He examined the effects of age, intervals of giving birth of less than 18 months, having had four or more pregnancies, having a history of stillbirths and neonatal deaths, having had three or more spontaneous abortions, having medical complications, and being disproportionately heavy or light in weight compared to height. Taken one by one, only two factors were significantly affected by participation in WIC: mean birthweight improved for nonwhite mothers with medical complications and for all mothers older than 34, especially nonwhite mothers. When Stockbauer analyzed mean birthweight with his broad risk measure, he found that the mean birthweight of the infants of at-risk participants was 50 grams greater than that of the infants of the at-risk non-WIC comparison group. This difference was statistically significant. However, this advantage was observed for only the nonwhite women. Silverman found that an increase in the number of risk factors increases WIC's ability to reduce the low-birthweight rate. Stockbauer's finding that nonwhites with one or more health-related risk factors benefit from WIC participation may lend some support to Silverman's results, but there is no reason to expect that supplemental food will affect risk factors that are not nutritionally related.

The effect of smoking has been examined by several evaluators. Metcoff found a significant benefit from WIC for women who smoke. He reported a difference of 115 grams in adjusted mean birthweight when comparing smoking WIC and non-WIC mothers. There is a direct relationship between low infant birthweight and the amount of smoking a mother does, but the effect on birthweight appears to be less for WIC participants. Bailey reported that among WIC participants, the mean birthweight does not significantly differ between smokers and nonsmokers, but there is a significantly lower mean birthweight for the infants of smokers in the non-WIC comparison groups. Kennedy suggests that smoking reduces the mother's weight gain and that it is a lack of weight gain that is associated with lower birthweights.

The effect of length of participation in WIC

Relating outcomes to the length of time women participate in WIC is useful in understanding WIC's effectiveness. There is some evidence that both mean and low birthweights rise significantly when program participation extends beyond 6 months. However, some severe design problems in the studies were not completely addressed when this effect was being measured. For example, the women who begin to participate in WIC early in their pregnancy may be more conscientious than women who enroll later. Performing the analysis for women enrolled for 8 or 9 months may, by definition, exclude women who miscarry or give birth prematurely.

The program's service of supplemental food would lead one to expect that, if it were helpful, the benefits to health and nutrition would increase with longer participation. Several of the better studies examined the effect of time. Kotelchuck found that the mean birthweight is 111 grams greater for the infants of women in WIC for 7-9 months than for their comparison group. Low-birthweight rates were 5.1 percentage points lower for these WIC women. There was no significant difference in the mean birthweight of the children of women in WIC for less than 7 months, but the proportion of infants at low birthweights was lower for women participating for 4-6 months. Metcoff did not examine the effect of time because all the WIC participants in his study entered at midpregnancy (about 19 weeks after conception). We note with interest that he found clearly positive results for WIC, which he reported as having been achieved in a participation period of 4-6 months, although Kotelchuck found minimal benefits for this period.

For WIC participants in Missouri, Stockbauer found greater mean birthweights and lower low-birthweight rates, both statistically significant, for mothers participating in WIC longer than 6 months compared to the non-WIC group. The difference in mean birthweight was 113 grams (125 grams for nonwhite participants and 88 grams for white participants). For both nonwhite and white participants, the low-birthweight rates (4.7 percent for nonwhite

and 3.4 percent for white) were less than one-half of the rates for comparable non-WIC mothers. No significant birthweight advantages were found for the infants of women participating in WIC for 6 months or less.

Kennedy found that the number of food vouchers that a participant receives is strongly and significantly related to mean birthweight. Her statistical analyses indicate that each voucher, representing a month in the WIC program, is associated with a gain of 23 grams. The mean birthweight of the infants whose mothers participated 7-9 months was almost 225 grams greater than that in the comparison group. Kennedy did not report on whether time discernibly affects low-birthweight rate.

The effect of WIC's components

We found only one study that attempted to analyze the effects of WIC's food supplements, nutrition education, and adjunct health care. Kennedy used the number of times mothers received prenatal care and nutrition counseling to try to determine the effect on birthweight. Neither one affected birthweight significantly. Attributing specific effects to each of the three components may, however, be impossible, since most evaluations collect data from many or all counties or from several states, where homogeneous food distribution practices, nutrition education efforts, and medical procedures are unlikely.

Kotelchuck, Kennedy, and Stockbauer used the number of food vouchers received or cashed to look at the effect of food supplements, but this information was related to the length of participation and did not differentiate the effects of food supplements from those of the other components of the program. In general, all the evaluations gave very little information on the particular procedures by which food was provided, nutrition counseling was given, and participants were referred for medical care.

WIC'S EFFECT ON FETAL AND NEONATAL MORTALITY

Our ability to analyze reported effects on miscarriages and stillbirths or neonatal deaths is hampered by the shortage of data and the substantial variability in the infant mortality measures that the researchers have used.[5] The results that are available indicate that participation in WIC may have a positive effect on newborn mortality, but we believe that the evidence is insufficient to support the claims that have been made about this.

One of the important goals of the WIC program is to reduce the incidence of fetal and neonatal mortality, but many of the evaluations have not addressed its success. Many WIC studies simply do not work with a sufficiently large number of cases to develop statistically meaningful data on mortality. The

combined fetal and neonatal death rate in the United States is about 20 deaths in each 1,000 live births. This is low from a statistical viewpoint, which complicates the detection of WIC's value in reducing the number of deaths. Comparison among studies is difficult, because some have examined miscarriages and stillbirths while others have examined fetal, perinatal, neonatal, or infant mortality, and still others have taken up other variations. It is clearly not possible to measure the rate of early miscarriages, for example, if they have not been recognized or reported.

Kotelchuck's analysis of WIC's effects for more than 4,000 women in WIC compared to an equal number not in WIC found 12 neonatal deaths among the WIC participants and 35 among the others. There were no neonatal deaths for women in the WIC program for 6 months or more but there were 12 neonatal deaths in the comparison group. We believe that one problem with the Kotelchuck study is that the WIC participants and the comparison group were not matched for health variables. In testimony before the Subcommittee on Nutrition in 1983, Dr. David Rush noted that approximately 350 women (7 percent of the original sample) were terminated from the program and that, as a group with especially high risks of poor outcomes, their experience was lost to the study. He added that the 12 neonatal deaths make a rate of 3 deaths in every 1,000 births and that this is half the rate in Holland and Sweden, countries known for their exceptionally low neonatal mortality. This is such an extraordinary outcome that it suggests there are unexplained design problems in the Kotelchuck study. Metcoff worked with a smaller population and found no significant differences in the rates of stillbirths and abortions.

Stockbauer reported a significantly lower perinatal death rate for nonwhite participants than for nonwhite nonparticipants. He found 19.3 deaths in every 1,000 births, a rate that is one third lower than the corresponding rate of 28.8 for the non-WIC group. Among white participants, however, he found that the non-WIC group has a perinatal death rate that is significantly lower than that of the WIC group. The result of these findings is that they cancel each other out, so that it is possible to say that the overall perinatal death rate of the total population is essentially the same for both WIC and non-WIC mothers. This raises a serious question about the reliability of the mortality data.

SUMMARY

The mean birthweight and the proportion of low-birthweight infants are sufficiently common and accepted measures that they allow for comparable results from study to study. However, our assessment of the quality and credibility of the data that have been collected and the analyses that have been made of them does not lead us to conclude that the evidence is conclusive one way or the other.

The general results indicate that participation in WIC is associated with a 16-20 percent decrease in the proportion of low-birthweight infants born to women eligible for WIC. Mean birthweight appears to rise about 1 to 2 percent. WIC mothers appear to experience greater benefit the longer they participate.

There is relatively little information available on WIC's effects on socioeconomic and health-related risk factors. We conclude tentatively that teenage women and black women who participate in WIC have better birth outcomes than comparable women who do not participate in WIC.

No data that are available enable us to differentiate between the effects of the WIC program's components. Information is insufficient to warrant conclusions regarding WIC's effect on perinatal and neonatal mortality.

NOTES

[1]The bibliographical data for these studies are in appendix II. The studies are Kotelchuck et al. (1981), Metcoff et al. (1982), Stockbauer (1983), Silverman (1981), Bailey et al. (1983), and Kennedy et al. (1982).

[2]Only one study employed a randomized control group design with control over data collection. The others used various types of nonrandomized control groups and retrospective data collection, such as the examination of administrative records. These methods are not usually as dependable, but the authors of the five studies did attempt to match the comparison groups carefully and used statistical techniques to adjust for differences between them. Two studies used matching strategies to devise a non-WIC comparison, one used a modified time-series approach, one set up several comparison groups, and one made up the comparison group of mothers who would have met WIC's eligibility requirements except that they resided in a nearby community where WIC was not available. Four of the studies collected information on the extent to which WIC services were delivered. All the authors gave evidence on the appropriateness of their methods. A major strength of these studies is that the investigators made a concerted effort to show how the receipt of services was connected to the observed outcomes.

[3]Low-birthweight rate is customarily reported as a percentage of the infants in the sample, or in a population, who weigh less than 2,500 grams at birth. Since the WIC legislation and regulations require that WIC services be directed to women who risk having poor pregnancy outcomes, it might be expected that proper delivery of services would lead to a decline in the percentage of infants who are born weighing less than 2,500 grams. However, the ability to detect WIC's effect on the low-birthweight population depends either on the prevalence of infants weighing less than 2,500 grams in the comparison group or on the availability of area-wide or national norms. If there are few infants

weighing less than 2,500 grams in a non-WIC comparison group, it will be difficult to detect a positive effect from the WIC program unless the WIC group is exceptionally large in number.

[4]Kennedy and Silverman reported WIC's effects on mean birthweights after adjusting for sociological and biological factors. Kennedy's adjusted weight is 60 grams, and Silverman's is 39.3 grams. If these figures are substituted in table 4, the average effect is 30.1 grams. It is 24.7 grams when weighted by sample size, which is the lowest and most conservative of the several estimates of WIC's effect on mean birthweights.

[5]Several different terms are used to describe the periods shortly before and after birth. "Fetal" refers to the period between birth and 7-8 weeks after conception, but the fetal death rate is based on deaths 20 weeks, or 5 months, after conception. "Neonatal" refers to the period between birth and 28 days after birth. The "perinatal" period begins with the completion of 28 weeks of gestation and is defined variously as ending from 1 to 4 weeks after birth. "Infant mortality" refers to a number of deaths among every 1,000 infants who are born alive but who live less than 52 weeks.

CHAPTER 3

DOES PARTICIPATION IN WIC AFFECT MATERNAL NUTRITIONAL STATUS?

A fundamental premise of the WIC legislation is that adequate maternal nutrition is essential for maintaining maternal body tissues and for insuring the optimal growth and development of the unborn. The WIC program was designed to provide adequate nourishment for women who have low incomes and who are pregnant or postpartum or breastfeeding their infants and who need nutrition they are not able to get. As an adjunct to health care, WIC provides for foods that are especially rich in protein, iron, calcium, and vitamin C and gives nutrition counseling, in an attempt to provide an adequate diet and reverse the course of anemia and abnormal weight gain.

Overall, the evaluations of WIC's effects on maternal nutrition are fewer in number and lower in quality than what is available for birthweights. The six best WIC studies on maternal nutrition are moderate in quality. They differ in the research methods they used, the way they controlled for alternative explanations of the findings, and the measurements they reported. Thus, we find little direct comparability in the findings and very little accumulation of evaluative evidence. Definitive conclusions are not possible at this time, but the results suggest that participation in WIC by pregnant women may be associated with a greater intake of calories and of iron, protein, and other nutrients and greater weight gain.

The evaluators who have assessed maternal nutritional status have used four methods. They have (1) estimated dietary intake, (2) analyzed biochemical tests, (3) examined body measurements, and (4) looked at the results of clinical examinations. Table 7 indicates some of the problems of precision and meaning in these methods. While there is consensus that a "low-birthweight infant" weighs less than 2,500 grams at birth, we found no universally accepted standard for maternal nutritional status. The four measurement methods in table 7 assess different aspects of nutrition. Individually, they cannot fully describe nutritional well-being, but taken together they can provide more complete information about the adequacy of dietary intake and the use of nutrients.

We found no single procedure or set of procedures in use consistently in the WIC evaluations. Even when similar measures have been used, the results have been reported in formats that do not easily allow aggregation. Although we have not synthesized the studies quantitatively, we present the data and make observations about trends in WIC's effect on dietary intake, biochemical measurements, and weight. For each of these measures, we address the question of the overall effect of the WIC program on maternal nutrition and then present the evidence that is available on

Table 7

Methods of Assessing Maternal Nutritional Status

Method	Focus	Measures	Measurements	Design remarks
Dietary intake	Self-reported, observed	Actual intake (24-hour recall, dietary records), usual intake (dietary history, food frequency)	Frequency of food consumption, mean nutrient and energy intake, nutrient intake/1,000 kcal, % RDA consumed	Direct observation usually infeasible; no generally accepted indirect measures; commonly used 24-hour recall of actual intake varies in validity and reliability
Biochemical tests	Nutrients in blood or urine	Blood plasma	Hemoglobin g/dl, % hematocrit, other iron indicators (transferrin saturation, mean corpuscular hemoglobin concentration); vitamin levels	No universally accepted standards of what hemoglobin, hematocrit, and other iron measures indicate; quality control of testing and lab procedures varies
Body measurements	Physical change	Anthropometric	Total weight gain, weight-gain rate during pregnancy, weight-gain grids, skinfold thickness	Verifiable baseline usually absent (most women cannot accurately recall prepregnancy weight); measurement of rates requires at least 2 consecutive visits but frequency and no. of prenatal visits vary
Clinical examinations	Nutrition-related illness or complications during pregnancy	Reported symptoms or signs such as blood pressure, use of health services	Incidence of toxemia, hospitalization rate, no. of prenatal visits or hospital days	Clinical indexes, difficult to attribute to WIC, are more directly related to quality of prenatal care; incidence of illness often over- and under-diagnosed

high-risk participants, the length of program participation, and the services from the separate WIC components. As with birthweights, we do not address the clinical importance of these findings.

Twenty-four WIC studies have addressed some aspect of the overall effect on maternal nutrition, but our procedure for rating their quality gives us confidence in only the six studies listed in table 8. As the table shows, five studies include data on dietary intake, four contain biochemical indicators, and three report information about weight gain. Four looked at one or more groups of women considered to have greater health and nutrition risk, two provide a little information about three aspects of WIC's implementation--the length of participation, the receipt of food supplementation, and the number of visits made to a clinic. None of the reports indicated that the relative effects of the three separate WIC components had been analyzed.[1]

WIC'S EFFECT ON MATERNAL DIET

Of the studies listed in table 8, all but the one by Kennedy and Gershoff included 24-hour dietary information recalled by women in WIC and by similar women not receiving WIC services. Endres presents Nutrient Dietary Data Analysis (NDDA) data for an early study, done in fiscal year 1978, and the NDDA study includes data separately for fiscal years 1979 and 1980. All the studies differ somewhat in the nutrients and in the types of measurements that were reported. The differences make it difficult to aggregate the results, although we have tried in tables 9 and 10 (on pages 32-33) to summarize the data reported on group means for the energy and protein that the women had ingested and on the average percentage of the recommended daily allowance (RDA) that they had consumed. Several observations can be made about these two key facets of the women's diets:

--in five of six studies, the WIC participants reported consuming more calories than the women in the non-WIC groups; in the sixth study, both groups reported an average intake greater than 100 percent of the RDA;

--in the studies reporting an advantage for WIC for the percentage of the RDA of calories consumed, the greatest difference between the WIC and non-WIC groups was 13.2 percent;

--perhaps because of the larger sample sizes, the only statistically significant differences in the mean percentage of RDA for energy intake were reported by the NDDA evaluators;

--all six studies reported that the women consumed more than 90 percent of the RDA for protein and five reported that the WIC women took in significantly more protein, ranging from 8 to 15 percent more than the non-WIC women;

Table 8

The Focus of the Analysis and Design of Evaluations of WIC's Effect on Maternal Nutritional Status

Study	Overall effects	Risk factors	Months of participation	WIC components	Design remarks
Metcoff	Dietary intake, biochemical and body measures	Not analyzed	Not analyzed	Not analyzed	Random control group, moderate sample, controlled data collection, WIC implementation data, statistical control
Bailey	Dietary intake, biochemical measures	Race, age	Not analyzed	Not analyzed	Comparison group of women eligible for WIC residing in non-WIC area, small sample
Endres	Dietary intake	Not analyzed	Not analyzed	Not analyzed	10% sample of women in WIC 6+ months compared to new entrants, moderate sample, trained data collectors, data analyzed centrally, group means compared
NDDA	Dietary intake	Age	Not analyzed	Not analyzed	Women receiving WIC supplements at home in prior 6 months compared to new entrants, relatively large sample, trained data collectors, data analyzed centrally, group means compared
Kennedy and Gershoff	Hemoglobin, hematocrit, weight measures	Biological and social variables	1-3 4-6 7-9	No. of clinical visits	Unmatched comparison group not fully described, small sample from larger study, retrospective WIC and medical data, WIC implementation data, statistical control
Edozien	Dietary intake, biochemical and weight measures	Race, age	Less than 3 months, 3 or more months, post-partum	Not analyzed	WIC participants compared to new entrants at 19 WIC clinics in 14 states, large sample, group data, substantial attrition in groups, complex multiple and other analyses not fully documented, initial and followup data for individuals not linked, questionable lab procedures and data quality control, WIC implementation data, statistical control

Table 9

Energy Intake Reported on 24-Hour Recall

Study	Year and location	Measure[a]	Reported data[b] WIC	Reported data[b] Non-WIC	Quantitative indicators: Raw difference	Quantitative indicators: % difference[c]	Quantitative indicators: Statistically significant
Bailey	1980	Mean kcal	2,390	2,496	-106	-4.2	No
	Two Fla.	Mean % RDA	104	108	-4	-3.7	No
	counties		(41)	(37)			
Metcoff	1980-82	Mean kcal	1,965	1,883	82	4.4	No
	Oklahoma	Mean % RDA	--	--	--	--	--
	City		(145)	(125)			
Endres	FY 1978	Mean kcal	--	--	--	--	--
	Ill.	Mean % RDA	77	68	9	13.2	Yes
			(115)	(651)			
NDDA	FY 1979	Mean kcal	--	--	--	--	--
	Ill.	Mean % RDA	76	70	6	8.6	Yes
			(341)	(1,064)			
NDDA	FY 1980	Mean kcal	1,888	1,780	108	6.1	--
	Ill.	Mean % RDA	80	75	5	6.7	Yes
			(873)	(2,277)			

[a] % RDA calculated with 1974 RDA standards, except 1980 NDDA calculated with 1980 RDA standards.

[b] The numbers in parentheses are sample sizes.

[c] The difference attributable to WIC or what would have been expected in the absence of WIC: (WIC - non-WIC)/non-WIC.

Table 10

Protein Intake Reported on 24-Hour Recall

Study	Year and location	Measure[a]	Reported data[b]		Quantitative indicators		
			WIC	Non-WIC	Raw difference	% difference[c]	Statistically significant
Bailey	1980	Mean grams	90	105	-15	-14.3	No
	Two Fla.	Mean % RDA	118	138	-20	-14.5	No
	counties		(41)	(37)			
Metcoff	1980-82	Mean grams	79.3	71.8	7.5	10.4	Yes
	Oklahoma	Mean % RDA	--	--	--	--	--
	City		(145)	(124)			
Endres	FY 1978	Mean grams	--	--	--	--	--
	Ill.	Mean % RDA	(105)	91	14	15.4	Yes
			(115)	(651)			
NDDA	FY 1979	Mean grams	--	--	--	--	--
	Ill.	Mean % RDA	101	93	8	8.6	Yes
			(341)	(1,064)			
NDDA	FY 1980	Mean grams	79	73	6	8.2	--
	Ill.	Mean % RDA	106	98	8	8.2	Yes
			(873)	(2,277)			

[a] % RDA calculated with 1974 RDA standards, except 1980 NDDA calculated with 1980 RDA standards.

[b] The numbers in parentheses are sample sizes.

[c] The difference attributable to WIC or what would have been expected in the absence of WIC: (WIC - non/WIC)/non-WIC.

--in the one study that reported greater protein intake for the non-WIC group, the difference was not significant, and both WIC and non-WIC groups reported intakes greater than 100 percent of the RDA.

The studies reported on nutrients other than protein that are supplied by the WIC food packages and that are strongly associated with growth and development. WIC women took in more iron, calcium, and vitamin C than non-WIC women. Both the WIC and non-WIC groups indicated a consumption of more than 130 percent of the recommended daily allowances of vitamin C but less than 85 percent of the recommended daily allowance of calcium. Iron intake in the Endres and NDDA studies was 85 percent of the recommended daily allowance or less, but Bailey found it to be greater than 90 percent.[2]

Overall, WIC women reported consuming greater quantities of calories and nutrients more often than non-WIC women. On the average, the diet of neither group appeared to be greatly inadequate, but relative deficiencies were reported for iron and calcium. However, Edozien pointed out that

> "even for those nutrients where comparison of mean values to the RDA suggested an adequacy of intake, there was still a high proportion of individuals who consumed inadequate amounts of the nutrient." (Edozien, 1976, vol. II, p. 226)

Thus, another way of looking at dietary intake is to compare the measurements of WIC participants who have very poor diets at their first WIC visit with their measurements after a period of participation. Edozien did this by calculating the percentage of participants who recalled consuming less than a predetermined quantity of each nutrient.[3] As is indicated in table 11, the Edozien data show that a higher proportion of the WIC participants reported lower dietary intakes (except for energy) at the initial visit than after 3 months of participating in WIC. The greatest improvement was for calcium--almost 20 percent. Although these data are severely limited, they suggest that more investigation is required if we are to find out whether WIC effectively improves the nutritional intake of pregnant women with very poor diets.

The NDDA study presented information about the dietary intake of women whose age meant that they were considered to have greater risks because of nutrition or health. The diets of pregnant women younger than 18 and older than 34 who were not in WIC appeared to be less dense for some nutrients (measured per 1,000 kilocalories) than the diets of WIC participants.[4] The WIC and non-WIC pregnant teenagers reported consuming similar percentages of the recommended daily allowance of calories, but among the older pregnant women, those in WIC consumed more energy (400 kilocalories) than those not in WIC.

Table 11

The Percentage of Pregnant Women in WIC Reporting Low Daily Intakes of Energy and Nutrients[a]

Nutrient	On entering[b]	After 3+ months
Energy	79-83 (2,536)	83-84 (256)
Protein	61-64 (2,536)	52-54 (256)
Iron	48-54 (2,754)	43-44 (280)
Calcium	49-54 (2,754)	28-36 (280)
Ascorbic acid	38-43 (2,754)	33-36 (280)

SOURCE: J.C. Edozien, B.R. Switzer, and R.B. Bryan, Medical Evaluation of the Special Supplemental Food Program for Women, Infants and Children (WIC) 6 vols. (Chapel Hill, N.C.: University of North Carolina School of Public Health, July 15, 1976.

[a]Low daily intake = less than 70 percent of RDA for energy and less than 54 percent of RDA for other nutrients.
[b]The numbers in parentheses are sample sizes.

WIC'S EFFECT ON MATERNAL BIOCHEMISTRY

The biochemical indicators

Four of the six studies--Metcoff, Bailey, Kennedy and Gershoff, and Edozien--included biochemical indicators but did not always report similar information. The most frequently used measures of biochemical status were the hematocrit and hemoglobin levels.[5] The four evaluations reported information about group means or the proportion of the groups with the low levels that are considered to indicate anemia. Three studies reported some differences between the WIC and non-WIC groups, and two of these studies found the differences to be statistically significant.

Metcoff measured hematocrit over time and found no significant increase in average levels at 24 or 32 weeks of pregnancy or at delivery in either high-risk WIC or high-risk non-WIC women. Metcoff also found no significant difference between these two high-risk groups and a group of low-risk pregnant women.[6]

Bailey reported similar average hematocrit levels, at 35 percent, for WIC and non-WIC women who were 30 weeks pregnant. However, the incidence of anemia (defined as a hematocrit level of less than 32 percent) was lower for the WIC participants. Seventeen percent of the WIC group, compared to 29 percent of the non-WIC group, were considered to be anemic.

Kennedy and Gershoff found that although there were no significant differences in the hematocrit levels of WIC and non-WIC women at the time of their first prenatal visit (at approximately 18 to 22 weeks of pregnancy), WIC mothers had significantly higher iron levels during the last trimester of pregnancy (usually at 34 weeks or later).[7] These findings differ from Metcoff's finding that there is no significant change in hematocrit over time. Part of the reason for the difference may be that Metcoff adjusted for the stage of pregnancy and the interval between measurements but Kennedy and Gershoff did not.

Defining anemia as either a hemoglobin level lower than or equal to 11 grams per deciliter (g/dl) or a hematocrit level lower than 34 percent, Kennedy and Gershoff compared women in their third trimester who had been classified as anemic on their first clinical visit. They found that women who were participating in WIC had significantly higher hemoglobin levels than women who were not. Final hematocrit levels for the two groups were not significantly different.[8]

Edozien found that the WIC program had no effect on the mean hemoglobin or hematocrit levels for women who had been pregnant for less than 28 weeks. For women more than 28 weeks pregnant and for postpartum women, WIC participation was associated with higher levels of hemoglobin and hematocrit.[9] WIC participation was similarly associated with a reduction in the incidence of anemia (that

is, hemoglobin less than 11 g/dl for pregnant women and less than 12 g/dl for postpartum women). The proportion of WIC women who either had been pregnant for more than 28 weeks or were postpartum and who were classified as anemic declined more than 30 percent.[10]

Three of the studies used other biochemical indicators (see table 7). While Metcoff did not find significant differences in the average hematocrit levels, he and his colleagues did find several significant differences on other biochemical indexes between women who were in WIC and women who were eligible for but not in WIC, all of whom were about 36 weeks pregnant and in high-risk groups. In another analysis, Metcoff found that several other maternal, nutrient-related blood levels measured at about 19 weeks of pregnancy were associated with the birth of both small and large infants--that is, infants considered by these evaluators to have the risk of health problems.[11] Although the two Metcoff analyses used different portions of the data, it is interesting that only two of the blood nutrients that differed significantly between the WIC and non-WIC women were also modestly related to infant birthweights. This suggests that more investigation might reveal whether the biochemical variables that are associated with pregnancy outcomes are, in fact, similar to those that are related to the mothers' nutritional status and affected by WIC.

Bailey reported that WIC participants were significantly better off on some biochemical measures, while non-WIC women were better off on others.[12] Edozien found that of 17 biochemical values that were analyzed in his study, 4 changed consistently and significantly for WIC women after their participation in the program.[13] However, Edozien reported conflicting estimates of iron deficiency for different measures of iron.[14]

In summary, two studies reported greater, statistically significant average hemoglobin and hematocrit levels for WIC participants, and two studies reported no significant differences. One study reporting no significant average differences found that a lower proportion of WIC women than non-WIC women were anemic. Two other studies reported similar findings about the reduction in the incidence in anemia. The information that has been reported on other biochemical changes is not sufficiently consistent to allow general conclusions. Overall, the evidence is modest and far from conclusive but does suggest that WIC may improve the biochemical status of pregnant women under certain circumstances.

The data for high-risk mothers

Unlike the information on birthweights, the information on WIC's effects on the biochemical status of high-risk women is relatively sparse. Three studies analyzed biochemical measures for women whose characteristics are most often associated with greater risk because of nutrition and health. Only two studies investigated the relationship between various risk factors and

changes in mothers' biochemical measures stemming from their participation in WIC.

In Bailey's study, no difference was observed in biochemical measures because of race or age. Kennedy and Gershoff found age to be significantly and positively associated with hemoglobin and hematocrit values during the last trimester of pregnancy--as age increased, so did these blood measurements. Kennedy and Gershoff did not include race in their analyses.

Edozien analyzed only the differences in blood levels by age and ethnic background for women at the time they enrolled in WIC. The finding was that women older than 40 had significantly higher hemoglobin and hematocrit levels than the younger women. Pregnant black women had the lowest mean hemoglobin levels while pregnant white women had the highest.[15]

Kennedy and Gershoff found the number of earlier pregnancies to be significantly and negatively associated with the hemoglobin and hematocrit measurements. In addition, they reported that WIC women with higher initial hematological levels had higher levels during the last trimester of pregnancy.

In summary, we believe that the information on biochemical levels for high-risk women is too sparse and inconsistent for making an informed judgment about WIC's effect on them.

Other effects

Only two of the six studies reported information about the effect of the intensity or length of participation in WIC. Kennedy and Gershoff found that WIC participants who received more food vouchers than others had higher hemoglobin and hematocrit levels.[16] Edozien reported that mean hemoglobin levels went up and the anemia rate went down for women who had been pregnant more than 28 weeks or were postpartum and who had received food supplements. The difference between the initial and the followup levels at both "less than 3 months" and "3 or more months" of WIC food supplementation were statistically significant.

Kennedy and Gershoff also looked at the effect of prenatal care on the final hemoglobin and hematocrit measurements. They found that as the number of clinical visits increased, hemoglobin and hematocrit levels decreased.[17]

In summary, there is too little information to draw any firm conclusions, but some evidence suggests that the longer that a woman participates in WIC, the better the levels of iron in her blood will be. The data regarding the negative association between biochemical measurements and the number of clinical visits are difficult to interpret without additional data about the health of the women who were studied and the kind of health care they received.

WIC'S EFFECT ON MATERNAL BODY MEASUREMENTS

How much weight women gain during pregnancy has been demonstrated to be strongly associated with infant birthweights. Two of the WIC evaluations present information about weight gain, and one of these studies includes data on other body measurements. A third study includes information about the relationships between a woman's weight before pregnancy, her weight gain during pregnancy, and her infant's birthweight.

Metcoff examined the effect of WIC by analyzing the measurements taken at 36 weeks of pregnancy. It was found that WIC women weighed 79.3 kilograms and non-WIC women weighed 76.8 kilograms, a difference of about 2.5 kilograms (5.5 pounds). This approached but did not reach statistical significance. It was also found that the women differed in thigh circumference and biceps skinfold.[18]

Edozien reported similar data for weight gain, finding that women who received WIC foods gained about 2 kilograms (4.4 pounds) more than non-WIC women at 24-27 and 28-31 weeks of pregnancy. The differences between the WIC women and those in the comparison group were not statistically significant at the other measurement times, but the WIC participants always had a greater weight gain.

In trying to determine what factors contributed to the higher birthweights of infants born to WIC mothers, which we discussed in chapter 2, Kennedy found that a mother's weight gain was a significant influence. Birthweight was also affected by her weight before pregnancy, an indicator of nutritional status before conception. Kennedy found also that the black women in this study had the lowest weight gains.

In summary, the data in two studies suggest that WIC food supplements may lead to greater, but not always statistically significant, weight gain during pregnancy. Another study shows that the mothers' increase in weight gain is related to increased birthweights.

SUMMARY

The available evaluative evidence is modest and preliminary but suggests that participation in WIC improves the intake of energy, protein, and some other nutrients for pregnant women, enhances the iron in their blood, and increases their weight gain. However, it should be noted that, on the average, the diets and iron of the non-WIC comparison groups were reported as not greatly inadequate. There are indications that longer participation in WIC may be associated with better iron levels. Information on the effects of WIC on maternal nutritional status given the nutrition and health risk factors is too inconsistent for making informed judgments.

Evaluations of maternal nutritional status have not attempted to discern whether there are different effects from the three separate WIC components. Information is incomplete on how WIC improves the nutrition of mothers relative to the birthweights and health of their infants. Strong statements cannot be made about the effects of WIC on maternal nutritional status, primarily because there is no consensus about the precise measures and standards for judging nutritional deficiency, there are limitations in the design and execution of the evaluations that have been performed, including the lack of information about WIC's implementation, and there are no comparable or cumulative findings among the studies.

NOTES

[1]The bibliographic data for these studies are in appendix II. The studies are Metcoff et al. (1982), Bailey et al. (1983), Endres, Sawicki, and Casper (1981), NDDA Laboratory (1980), Kennedy and Gershoff (1982), and Edozien, Switzer, and Bryan (1976). Metcoff's study is the only one that used a prospective, randomized control group design. The other studies used nonrandomized comparison groups and attempted to compensate for differences in the background and other characteristics of the women in these groups by means of the sample selection or statistical adjustments. The sizes of the samples were relatively large in two studies, moderate in one, and small in three. The methods for making statistical adjustments differed. For example, Metcoff adjusted for the normal increase and decrease of blood values throughout pregnancy by including in the analyses the week of gestation when measurements were taken, the initial level of the measurements, and the time interval between two measurements. The NDDA studies found significant differences in the consumption of calories between the comparison groups and carried out additional analyses of the mean nutrient intake per 1,000 kilocalories of food energy in order to determine the nutrient density of the women's diets and thus control for differences that should otherwise be attributable simply to greater caloric intake. Edozien adjusted for age, ethnic origin, and income (using an income-poverty ratio) in most of the statistical analyses.

[2]More specifically, iron intake was reported by Bailey, Endres, NDDA, and Edozien. They all reported statistically significant differences between the WIC and non-WIC groups. The largest differences for the percentage of the RDA were reported for the larger NDDA samples--56-66 percent for the non-WIC group and 65-85 percent for the WIC participants. As for calcium, the NDDA WIC groups consumed significantly more than the non-WIC groups--78-81 percent of the RDA compared to 66-70 percent. According to NDDA, WIC women took in 148-194 percent of the RDA of vitamin C, non-WIC women 132-150 percent. Reporting mean intakes for these three nutrients, after adjusting for intake per 1,000 kilocalories, NDDA found that the differences between the WIC and non-WIC groups remained statistically significant for

iron but not for calcium and vitamin C. Edozien reported that WIC participation increased the pregnant women's intake of protein, iron, calcium, and vitamin C but not energy, but the data are not directly comparable to data in the other studies.

[3]The quantity that Edozien used (in the study reported in 1976) to categorize participants as having "low" dietary intake was the quantity that would ordinarily meet the needs of only 2.5 percent of the population; that is, only 2.5 percent of the participants would normally be in this category.

[4]The non-WIC teenagers had slightly higher mean nutrient intakes per 1,000 kilocalories for iron, thiamin, niacin, and vitamins A and C. The non-WIC women older than 34 had greater intakes of protein, calcium, iron, niacin, and vitamins A and C.

[5]"Hemoglobin" level, or concentration, refers to the oxygen-carrying capacity of red blood cells and is expressed in grams per deciliter (g/dl). "Hematocrit" level refers to the volume of red blood cells and is expressed as a percentage of the total blood volume. These measurements record two separate hematological characteristics, and there is no acceptable method of converting either measurement into the other. Anemia is the reduction of the hemoglobin concentration, the hematocrit, or the number of red cells to a level below that which is normal for a given individual. This level differed in the studies: for Bailey, it was hematocrit less than 32 percent; for Kennedy and Gershoff, it was hematocrit less than 34 percent or hemoglobin less than or equal to 11 g/dl; for Edozien, it was hemoglobin less than 11 g/dl during pregnancy and less than 12 g/dl postpartum.

[6]In the Metcoff study, the average hematocrit level ranged from 36.3 to 37.1 percent.

[7]Kennedy and Gershoff reported the following mean levels: hemoglobin at 12.6 g/dl (WIC) and 11.7 g/dl (non-WIC) and hematocrit at 36.7 percent (WIC) and 35.1 percent (non-WIC).

[8]Kennedy and Gershoff reported significant hemoglobin findings--12.1 g/dl for WIC women and 11.5 g/dl for non-WIC women. They also noted that people with nutritional anemia respond to iron therapy by absorbing relatively more iron and that blood hemoglobin levels respond quickly to iron medication.

[9]Edozien reported that the average hemoglobin concentration for WIC participants was 0.31 to 0.33 g/dl greater than that for women just entering the WIC program. WIC women more than 28 weeks pregnant also had hematocrit levels 0.5-percent higher.

[10]About 27 percent of the women pregnant for more than 28 weeks were classified as anemic when they entered WIC. After the women had received WIC services for 3 months or more, about 18 percent were classified as anemic--about a 33-percent

reduction in the incidence of anemia. For postpartum women, 23.3 percent were initially classified as anemic. About 13.6 percent were classified as anemic after 3 months or more in WIC--a 41.6-percent reduction.

[11]Metcoff reported that WIC supplementation is related, in a statistically significant way, to increased leukocyte protein synthesis, higher plasma beta-globulin levels, and lower levels of riboflavin and the two plasma amino acids alanine and cystine. In a separate analysis, Metcoff reported that the levels of the two plasma amino acids leucine and phenylalanine, the iron-building capacity, and the cholesterol level were significantly and positively related to giving birth to small or large infants (weighing less than 3,000 grams or more than 3,600 grams). However, the levels of riboflavin (GRST), the plasma amino acids arginine and tyrosine, and leukocyte protein synthesis approached, but did not reach, statistical significance.

[12]Bailey reported that iron, as measured by transferrin saturation levels, was significantly better in WIC than in non-WIC women. Serum folacin levels were significantly higher in non-WIC women. Their red blood cell folacin levels were also higher but not significantly. Serum iron levels were higher in the WIC women but not significantly different in the two groups.

[13]Edozien found that, when all pregnant women were considered together, the only consistent, significant biochemical changes were an increase in hemoglobin, in mean corpuscular hemoglobin concentration, and in vitamin A and a decrease in total protein.

[14]For women more than 28 weeks pregnant, for instance, Edozien found the estimates of 36.5 percent for saturation of transferrin below 15 percent, 10.6 percent for mean corpuscular hemoglobin concentration below 30 g/dl, and 4.1 percent for plasma iron below 40 micro-g/dl.

[15]Edozien reported a difference of about 1.0 g/dl between black and white women.

[16]In the Kennedy and Gershoff study, each additional WIC voucher that was received corresponded to the statistically significant differences of 0.12 g/dl in the final hemoglobin level and 0.28 percent in the final hematocrit level.

[17]Kennedy and Gershoff reported that for each increase in the number of visits for prenatal care, hemoglobin decreased by 0.087 g/dl and hematocrit decreased by 0.22 percent.

[18]Metcoff adjusted for the initial measurements (which were recorded at approximately 19 weeks), the actual week of gestation when those measurements were taken, and the interval between the initial and final measurements.

CHAPTER 4

DOES PARTICIPATION IN WIC AFFECT THE INCIDENCE OF ANEMIA AND MENTAL RETARDATION IN INFANTS AND CHILDREN?

The Senate Committee inquiring about the effectiveness of the WIC program on the health of infants and children asked us to look at the problems of anemia and mental retardation.[1] Two assumptions of the program are that it can have a direct effect on anemia and an indirect effect on mental retardation. Since iron deficiency in the older infants and young children who are enrolled in WIC is fairly widespread, it is expected that supplementing their diets with iron-rich foods will be beneficial. Changes in the incidence of mental retardation, however, are expected to be conditional on WIC's raising the birthweight of infants above a risk threshold and on its improving their health and nutrition.

We found that the evidence is insufficient to support an assertion that the WIC program reduces the chances that infants and children will be anemic or mentally retarded. Twenty-five evaluations have reported information about the overall effects of WIC on iron in the blood of infants and children, but we have marginal confidence in only two. No evaluation has focused on the incidence of mental retardation, although one study examined the cognitive development of WIC participants. This study is so limited, however, that we do not have confidence in its conclusions.

The limited evidence on anemia from the two studies of moderate quality suggests that WIC may reduce the incidence of anemia among infants and children. The program seems to have been especially helpful for those who were classified as anemic when they entered the program and for those who remained in the program for at least 6 months. However, since no well-designed comparison groups were used in the evaluations, we cannot confidently attribute improvements in iron to the WIC intervention. We do not know the extent to which other factors affected the data. One such factor might be that some families were more highly motivated than others to continue in WIC because it seemed to be helping a child. Since information is missing for a large number of participants who were not available for the duration of the evaluation, we cannot determine the range of improvements in iron for all WIC participants. We cannot draw any conclusions at all about the incidence of mental retardation among WIC participants.

WIC'S EFFECT ON ANEMIA

Problems in measuring anemia

Since anemia is one of the conditions that signifies eligibility for participation in the WIC program, hemoglobin and

hematocrit data are frequently available in its records. However, the use of these data for assessing whether WIC can change the anemia of infants and young children is complicated by the fact that iron levels change with age, the lack of agreement about appropriate measures and standards for determining anemia, and the problems in performing evaluations with credible research designs.

Iron levels in children are not constant, especially during the first and second years.[2] Iron levels can be a useful indicator of anemia from 6 months after birth and possibly even from 4 months. However, evaluative measurements must control for each child's exact age, so that natural maturation can be taken into account.

That there is no uniformly accepted method for measuring iron further complicates the assessments of WIC's effect on anemia. Hemoglobin and hematocrit are the most common measures, but they are general, rather than specific, indicators of the presence of anemia. Some consider them to be only a crude means of detecting iron deficiency. Many children may grow deficient in iron stores but exhibit no change in hemoglobin. Using these and other biochemical tests and different laboratory methods can lead to differing estimates of anemia for the same population.[3]

Even when the same test is used, estimates of anemia can differ because different criteria are used to define iron deficiency. Measurements below certain points on different biochemical scales indicate a risk of nutritional deficiency and disease, but expert opinions disagree about precisely which points indicate risk and about the clinical implications of using different points to indicate iron deficiency.[4]

The studies of WIC's effect on anemia in infants and children are particularly beset by design problems. The loss of participants from a study between measurement-takings has been as great as 70-90 percent, so that the data cannot be interpreted conclusively. They may be biased if the persons who remain in the program differ in important ways from those who leave. For example, it may be that some participants in the program were not benefiting from WIC and, realizing this, chose not to continue. Final data would be missing for them, and a study's results could overestimate the overall effect of WIC. Similarly, if comparison groups are not carefully constituted, it may be unwise to attribute improvements to WIC that are better attributable to other factors, such as maturation. Finally, one might question the quality of any biochemical data that are obtained from clinics whose testing and recording are not uniform and consistent.

The evidence on anemia

The two WIC evaluations in table 12 reported data on anemia in infants and children. The data were collected from several states in 1974-76, the early years of WIC's implementation. The evaluations are flawed but offer some interesting results. The

Table 12

The Focus of the Analysis and Design of Evaluations of WIC's Effect on Anemia in Infants and Children

Study	Overall effects	Risk factors	Months of participation	WIC components	Design remarks
Edozien	Biochemical measures	Anemia	1- 6 7-11	% of food supplements received	WIC participants compared to new WIC entrants at 19 WIC clinics in 14 states, large sample, group data, 56% attrition, complex multiple regression and other analyses not fully documented, initial and followup data for individuals not linked, questionable lab procedures and data quality control, WIC implementation data, statistical control
CDC	Biochemical measures	Anemia	1- 6 7-12	Not analyzed	No comparison group, moderate sample from several states not nationally representative, retrospective and longitudinal data

studies are inconclusive and hampered by the general problems we mentioned above but have some strengths that the other evaluations we reviewed do not have.

Data for the two studies were available for a large number of infants and children at the time they entered WIC but for only a portion of them after a year's participation. In both studies, one fifth or fewer of the WIC youths were black. In the CDC study, more than half were white. In the Edozien study, more than half were Spanish American. Edozien collected and analyzed data on WIC's implementation; CDC did not.[5]

Both studies collected the data on anemia for children entering the WIC program. When anemia was defined at 10 grams of hemoglobin per deciliter, the prevalence of anemia in children 6 to 23 months old ranged from 10 to 14 percent. Hemoglobin at 11 g/dl for children 24 months old or older gave a range of 12 to 24 percent.[6]

Using a subset of their data bases and the same definition of anemia for the respective age groups, both studies showed a reduction in the rate of anemia after one year of participation in

Table 13

The Percentage of WIC Children with Low Hemoglobin at the Start of Participation and After 12 Months

Study and age group	Visit[a] 1	2	3
Edozien[b]			
6-23 months[c]	12.9 (8,996)	6.6 (5,437)	5.3 (1,961)
24-47 months[d]	18.8 (9,326)	10.3 (4,876)	10.0 (2,949)
CDC			
6-23 months[c]	14.2 (450)	6.1 (450)	2.7 (450)
24-47 months[d]	28.9 (260)	8.9 (260)	9.6 (260)

[a]Visit 1 was the first clinical visit; visits 2 and 3 were about 6 and 11-12 months after that. The numbers in parentheses are the total number of participants.
[b]The composition of the groups changed; the analysis adjusted for age, gender, race, and income.
[c]Hemoglobin lower than 10 g/dl.
[d]Hemoglobin lower than 11 g/dl.

WIC. As we show in table 13, the proportion of WIC children with low hemoglobin dropped, decreasing at higher rates for the younger children. Of the children 6 to 23 months old, about 13 to 14 percent were classified as anemic at their first visit, but only about 3 to 5 percent were anemic at their third (second followup) visit. The decline this represents is about 59 percent for Edozien's study and 81 percent for CDC's. A greater proportion of the children 24 months old or older were anemic at their first visit --19 percent in Edozien and 29 percent in CDC--but about 10 percent of these children remained anemic at their third (second followup) visit in both studies. This is a decline of 47 percent for Edozien and 67 percent for CDC. Most of the improvement was during the first 6 months of participation in WIC.

Edozien also compared the average mean hemoglobin levels of children who had participated in WIC for 6 and 11 months with children who were just enrolling in WIC, making statistical adjustments for differences between the groups in age, gender, ethnic origin, and family income. Edozien reported that the WIC children who were 12 months old or older had greater, statistically significant hemoglobin levels at the 6-month and at the 11-month visit.[7] However, since the comparison group was not a true control group, and since very little information was given about its members, we believe that the data indicate only that WIC may possibly make a difference.

In another analysis, CDC looked at changes in hemoglobin and hematocrit levels for children who were considered anemic at their first clinical visit. Of anemic children 6 to 23 months old--64 children with low hemoglobin and 162 with low hematocrit--fewer than 10 percent were still anemic after 12 months of participating in WIC. Of anemic children 24 to 47 months old--75 children with low hemoglobin and 242 with low hematocrit--fewer than 20 percent were still anemic after 12 months. Most of the improvement was in the first 6 months of participation. The CDC evaluators concluded that the changes were the most dramatic for the children who had the lowest levels to start with. However, the changes may be attributable, in part, to such problems as errors in laboratory measurements and a regression to the mean.[8]

Edozien also reported that hemoglobin levels were related to the proportion of food supplements received. The adjusted mean hemoglobin for the participants who received less than 76 percent of their food packages was lower than that for the participants who received 76 to 100 percent of their packages.[9] Since mean hemoglobin was considered adequate for both groups, an additional analysis of initially low hemoglobin levels might be more informative, but these data were not reported.

Summary

Since no groups were suitably designed for comparing what happens without WIC, it is not possible to ascribe increases in iron levels to the WIC program with certainty. Using two different

levels to account for normal maturational changes, CDC and Edozien showed similar results. The data from both studies suggest that WIC is somewhat helpful in reducing anemia for children whose iron is low at the beginning of participation, especially during the first 6 months of enrollment. Part of the observed improvement, however, may be the result of other factors that were not accounted for in the evaluations.

WIC'S EFFECT ON MENTAL RETARDATION

Problems in measuring mental retardation

Determining the effects of WIC on mental retardation is very difficult. It is made difficult by limits to the general understanding of how to measure mental processes. The three major problems in evaluating the cognitive development of infants and children are the variation in what is measured, the difficulty of detecting real effects, and the inability to disentangle the effects of nutrition from other influences.

The phenomena of mental development are so complex that no test measures them all. Since the various tests measure different aspects, the results are often not comparable. For example, standard tests for early development measure reflexes and sensorimotor alertness, but standard intelligence tests measure abilities in abstraction, verbal and spatial reasoning, and problem solving.

No measure may be sufficiently sensitive to demonstrate changes in cognitive, emotional, and sensorimotor development in areas where malnutrition is not severe. Even where it is, the causal link between severe malnutrition and poor cognition is only suggestive, rarely conclusive. A relationship between mild or moderate malnutrition and cognitive development has not been well established.

Finally, studies have not been able to isolate nutrition as either the only or a major contributor to mental deficiencies diagnosed in later childhood. This is because malnutrition often occurs in a milieu where socioeconomic status is low, education is limited, sanitation is poor, and infection recurs. Many children who are nutritionally deprived during infancy are also exposed during their early, formative years to other complex environmental problems. Since malnourished children are especially deprived, it is difficult to select appropriately matched control groups. There is no uniformity in the criteria that are used either for diagnosing malnutrition or for evaluating its relationship to other social, environmental, and economic conditions.

The evidence on mental retardation

Given these problems, it is not surprising that we found no WIC evaluation studies that focus specifically on mental retardation, and we found only one that attempted to assess cognitive

development. Hicks and colleagues compared 21 pairs of siblings from rural Louisiana who started participating in WIC at different times. The children who were born to women who received WIC food supplements during the last 3 months of pregnancy (called the "early" group) scored significantly higher on most intellectual and behavioral measures than their older siblings, whose food supplements began after they were a year old (called the "late" group). However, we have little confidence in this conclusion, for the following reasons:

--the sample of 21 pairs of siblings was small (although the use of sibling pairs helped control for demographic variables),

--the length of treatment for the two groups (about 56 months for the "early" group and 31 months for the "late" group) may have affected the results but was not controlled for,

--the children in the "late" group may have been less healthy or had greater nutritional risk than their siblings in the "early" group,

--the study reports no information about what WIC services were provided,

--the results may be biased because the testing psychologist knew which groups the children belonged to, and

--it is not clear that appropriate statistical analyses were used.

The limitations of the Hicks study prevent us from firmly concluding that the mostly favorable results that were reported about WIC's increasing children's mental abilities are attributable to early participation in the program.[10] The study provides insufficient evidence for making any reasonable judgment. For one thing, alternative explanations were not sufficiently ruled out: it may be that lengthy, as much as early, participation contributed to the differences that were observed. For another, the Hicks study did not make comparisons with children who were not in WIC: it cannot be said with confidence that the changes that were reported were caused by participation in WIC. The Hicks study may be viewed as a first step in an attempt to assess WIC's effects on mental growth; further research is required before any conclusions are possible.

SUMMARY

We found that the evidence is insufficient to confirm assertions that the WIC program reduces the incidence of anemia in infants and children. Most of the data that are available come from studies of dubious research design. The limited evidence from the two studies of moderate quality suggests that participation in WIC

decreases anemia, especially during the first 6 months of enrollment in the program, but crucial flaws make the finding questionable. No information is available on the incidence of mental retardation among WIC participants. One study providing meager information about the cognitive development of WIC children suggests that WIC has some positive effects under some circumstances, but conclusive information clearly depends on more research.

NOTES

[1]Other common indicators of the nutritional status of infants and children include body measurements of height, weight, weight for height, height for age, head circumference, skinfold thickness, and the like. We focused only on the indicators specified in the congressional request.

[2]The fetus absorbs iron from the mother during the last trimester before birth, and at birth a transition from fetal toward adult hemoglobin begins. At about the third or fourth month of life, the infant's hemoglobin concentration reaches about 10 g/dl. It rises to 11-12 g/dl at 12-15 months. Hemoglobin levels remain low, at 9-10 g/dl or less, when the amount of biologically useful iron in ingested foods is inadequate.

[3]In a nutritional study in the northwestern United States, the prevalence of anemia from iron deficiency was reported to be 8.3 percent when measured by hemoglobin but 17.5 percent when measured by free erythrocyte protoporphyrin. Therefore, it has been suggested that more than one test be used to detect iron deficiency.

[4]Such points, sometimes called "cutoff" points, are derived from reference-population norms that are established by means of national nutrition surveys or a consultation of experts, and they can differ from study to study. The use of a single, arbitrary point assumes that iron levels do not depend on such things as age and race. Regarding age, this assumption is false, especially for the early years of life. Ethnic differences, such as that blacks have lower hemoglobin levels than whites, are well documented, but the basis for them is disputed. Some evidence suggests that at least 50 percent of the anemia that has been diagnosed for blacks may actually be a systematic difference of 0.5 to 1.0 g/dl in hemoglobin between blacks and whites of all ages and incomes and, thus, that the cutoff point for whites may not indicate true iron deficiency in blacks. This suggests the need for separate criteria for each race. However, other evidence of the successful treatment of black infants with iron supplements supports the use of uniform definitions of anemia for both races. To overcome some of these difficulties, the Centers for Disease Control (1982) has developed reference curves from National Health and Nutrition Examination Survey I data.

[5]There may be some overlap in the two data bases; both included Arizona data for overlapping time periods. Edozien

(Edozien, Switzer, Bryan, 1976) reported that 41,000 infants and children were examined but that only 44 percent of those who could have returned for their 11-month visit during the study did so. In CDC (Centers for Disease Control, 1977), Nutrition Surveillance System records, reported mainly by six states, were examined for 115,249 WIC infants and children at their first visit to WIC clinics, and data for approximately 12 months of WIC participation for 5,692 infants and children in four states were taken from three linked records. In the Edozien study, the participants who returned for their 11-month visit were 57 percent Spanish American, 21 percent black, and 18 percent white. In CDC, approximately 18 percent were Spanish American, 13 percent were black, and 55 percent were white.

[6]Edozien used hemoglobin at 11 g/dl to define anemia for all children and, therefore, reported a higher incidence of anemia for the youngest group (6-23 months of age). We report data at 10 g/dl in order to make them comparable to the CDC data and to account for normal growth.

[7]Edozien reported that the adjusted mean hemoglobin levels at 6 months and 11 months were greater by 0.34 g/dl and 0.42 g/dl.

[8]Regression to the mean occurs if an individual's score is at the bottom (or top) of a measurement index, because there is no place to go but up (or down). The implication for evaluation studies is that, if an individual's score is below (or above) the 50th percentile, for example, at a first reading, a change upward (or downward) may be expected in that individual's percentile standing at a second or subsequent reading, even in the absence of any intervention. This should not be overlooked in the search for small improvements in the percentile standings of populations at the extremes of a distribution (which happen frequently in health interventions).

[9]Edozien used regression analysis to investigate the relationship between changes in hemoglobin and the amount of food that participants received from WIC. He defined the amount of food that was received in terms of a percentage--1-49 percent, 50-75 percent, and 76-100 percent--of the total amount that could be issued by the WIC projects. The adjusted mean hemoglobin levels were 12.7 g/dl for children who received less than 76 percent of their food packages and 12.9 g/dl for children who received 76-100 percent.

[10]Hicks reported that the children who were born to mothers participating in WIC during the last trimester of pregnancy scored significantly higher at home and at school on most intellectual and behavioral measures--including IQ, attention span, visual-motor synthesis, and school grade-point average--than their siblings who began WIC supplements after reaching one year of age. Most of the health measures reflected this difference, but only the relationship between height and age was significantly different for the two groups.

CHAPTER 5

SUMMARY OBSERVATIONS

Evaluators have been studying WIC for nearly a decade at the local, state, and national levels. They have used a variety of research designs and focused on different aspects of the program, and their evaluations exhibit a range of methodological quality. The studies generally report positive findings about WIC's effect for several health and nutrition measures. Proponents of the program often cite these findings as evidence that WIC is effective, while others often contend that the evaluation methods from which the findings derive are so flawed that drawing conclusions is not possible.

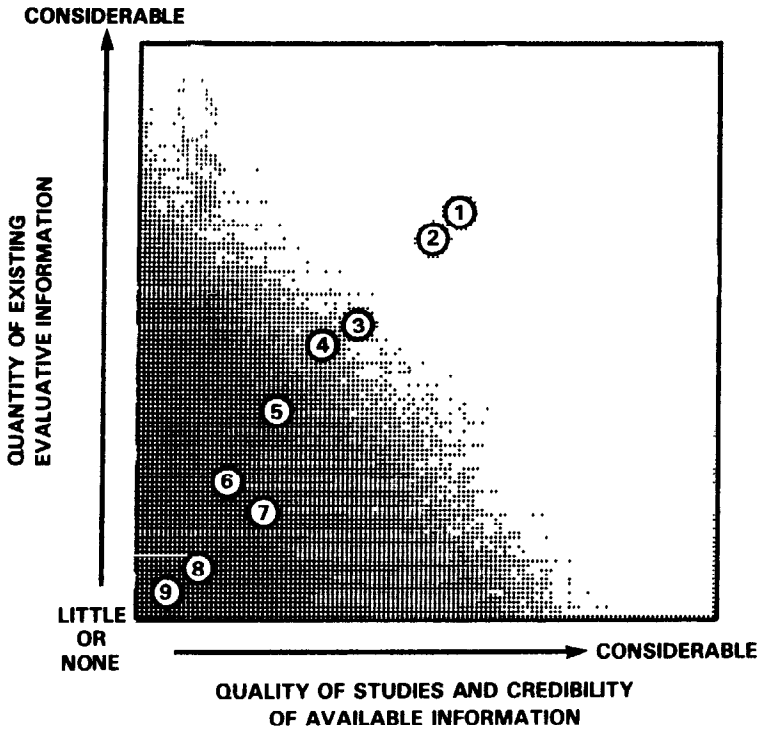
The Senate Committee on Agriculture, Nutrition, and Forestry asked us to perform an objective review and analysis of the evaluations of WIC to determine whether they support the assertions that are made about the program's effectiveness. We formulated five major questions on WIC's overall effects on (1) the birthweights of infants, (2) miscarriages, stillbirths, and neonatal deaths, (3) maternal nutrition, (4) anemia in infants and children, and (5) mental retardation in infants and children. We also looked for evidence of WIC's effect on groups having a variety of health and nutrition risks and on groups participating in WIC for different lengths of time and for evidence of the separate effects of the three WIC services--food supplements, nutrition education, and adjunct health care. We looked for answers to our questions from all the evaluation studies whose findings were relevant and credible. We attempted to determine what is known about WIC's effectiveness and to synthesize this information, and we attempted to assess the problems in the current state of WIC evaluation efforts.

Our critical review of the evaluation designs and their execution leads us to believe that the information is insufficient for making any general or conclusive judgments about WIC's effectiveness. However, in a limited way, the information indicates the likelihood that WIC has modestly positive effects in some areas.

WHAT IS KNOWN ABOUT WIC'S EFFECTIVENESS

Any attempt to assess the strength of the evidence given for WIC evaluation findings must take into account both its quantity and its quality. To say that the reported evidence is conclusive regarding a specific outcome would require an adequate amount of evaluative information of high quality. Even a single evaluation of good quality can rarely have this kind of power. The presence of a number of evaluations that are sound in design and execution would give strength to what is generally known, if the findings of the different studies were consistent and cumulative. Figure 2 shows the importance of looking at the evidence from a number of studies. When many provide little

FIGURE 2
OUR ASSESSMENT OF THE STRENGTH OF THE EVALUATIVE EVIDENCE ABOUT THE WIC PROGRAM'S EFFECTS



LEGEND.

- ☐ CONCLUSIVE EVIDENCE
- ◘ SOME OR MODERATE EVIDENCE
- ■ GAPS IN KNOWLEDGE

KEY

1. INCREASE IN MEAN BIRTHWEIGHTS
2. DECREASE IN PERCENTAGE OF LOW-BIRTHWEIGHT INFANTS
3. EFFECTS, FOR HIGH-RISK GROUPS AND FOR THOSE PARTICIPATING LONGER THAN 6 MONTHS, ON BIRTHWEIGHTS
4. IMPROVEMENT IN MATERNAL NUTRITION
5. DECREASE IN INCIDENCE OF ANEMIA IN INFANTS AND CHILDREN
6. DECREASE IN INCIDENCE OF FETAL AND NEONATAL MORTALITY
7. EFFECTS, BY LENGTH OF PARTICIPATION AND FOR HIGH-RISK GROUPS, ON MATERNAL NUTRITION, FETAL AND NEONATAL MORTALITY, AND ANEMIA IN INFANTS AND CHILDREN
8. DECREASE IN INCIDENCE OF MENTAL RETARDATION IN INFANTS AND CHILDREN
9. EFFECTS OF THE THREE SEPARATE WIC COMPONENTS

information or their quality is poor, we have an indication of where there are gaps in our knowledge.

Figure 2 displays our assessment of the strength of the evidence on WIC's effectiveness as we found it in the relevant and credible evaluation reports. That there are no findings in the most lightly shaded area indicates that none of the evaluations we reviewed give us evidence that is conclusive. The data on the birthweights are substantial, but our assessment of their quality and credibility lead us to the statement that their strength as evidence is only moderate. For the remaining evaluation questions, the information that is available pushes the findings toward the "gaps in knowledge" corner of figure 2, indicated by the darkest shading. The two areas for which we found particularly little, if any, information are the incidence of mental retardation and the separate effects of the three WIC components.

Does participating in WIC affect infant birthweights?

On birthweights, we found six studies that are of high or medium quality and that somewhat support, but not conclusively, the assertion that WIC has the positive effect of increasing the birthweights of the infants of mothers who participate in WIC. Five of the six studies that examined the proportion of low-birthweight infants--that is, infants weighing less than 2,500 grams at birth--show that participation in WIC is associated with some improvement. About 7.9 percent of the women participating in WIC gave birth to low-birthweight infants compared to about 9.5 percent of the women in the non-WIC comparison groups. Related calculations suggest a decrease of 16 to 20 percent in the proportion of infants who are thought to have health risks at birth because of their weight. The effect of WIC on mean birthweights seems positive also. Our estimate is that the average benefit for WIC participants is 30 to 50 grams, which is a 1-2 percent increase in mean birthweight. However, both WIC and non-WIC infants averaged about 3,200 grams at birth, which exceeds the 2,500-gram boundary below which neonatal and infant health problems are expected.

Does participating in WIC affect fetal and neonatal mortality?

Both the quantity and the credibility of the results on fetal and neonatal mortality are substantially lower than those on birthweight. The favorable results reported from several evaluations are low in credibility. We consider them to be insufficient to support the assertion that WIC reduces the incidence of fetal and neonatal deaths.

Does participating in WIC affect maternal nutrition?

The quality and the quantity of evidence from WIC evaluations on how WIC changes maternal nutrition are lower than those

on birthweight. Six studies, of moderate quality, differ in many important aspects, including the rigor with which they rule out alternative explanations and the measurements they report. Therefore, it is difficult to synthesize the results of these studies. It is not yet possible to make firm conclusions, but there is some evidence to suggest that participation in WIC is associated with some improvements in nutritional well-being, especially in diet, iron, and weight.

Does participating in WIC affect anemia in infants and children?

We found very little usable evidence to give support to the claims that the WIC program reduces the chances that infants and children will have anemia. Limited evidence from two studies of only moderate quality suggests that WIC is associated with improving the levels of iron in the blood of children classified as anemic when they enter the program. This evidence is inconclusive.

Does participating in WIC affect mental retardation in infants and children?

Virtually nothing is known about whether WIC does or does not have an effect on the incidence of mental retardation. No WIC evaluation has specifically addressed the question. One study focused on the cognitive development of infants and children in WIC, but limitations in its study design and execution lower our confidence in its favorable conclusions.

Does participating in WIC benefit some groups more than others?

Regarding the different effects that WIC may be having for different groups of WIC participants, we found some information in which we have moderate, but not high, confidence. WIC appears to have greater positive effects on infant birthweights among pregnant teenagers, black women, and women with multiple nutritional and health-related risks. The lack of sufficient and consistent information prohibits making informed judgments about the differences in WIC's effect on fetal and neonatal mortality, maternal nutrition, and anemia in infants and children.

Does WIC's effectiveness depend on length of participation?

We found some evidence suggesting that participating in WIC for longer than 6 months is associated with increases in average birthweights and decreases in the proportion of infants who are born at low birthweights. Some evidence suggests that longer participation improves iron levels in a mother's blood. As for anemia in children, the limited evidence suggests that its incidence is reduced the most during the first 6 months of participation. However, there are flaws in the evaluations that make this evidence inconclusive.

What are the different effects of the three individual WIC components?

We found very little information about the separate effects of the three components of the WIC program. Most of the evaluations identify WIC participation from unvalidated listings on the WIC rolls and give no description of the WIC intervention being analyzed. Only a few studies give information about the effect of the receipt of food supplement vouchers. A few studies mention the number of visits to a clinic or the number of times nutrition-education sessions are attended, but the analysis of this information is incomplete. The separate effects of the three WIC program components have not been adequately examined.

THE CURRENT STATE OF EVALUATION EFFORTS

What is known about the WIC program's effectiveness is predicated, at least in part, on the current state of WIC's evaluation, about which we have several observations. We found that no study or group of studies provides conclusive evidence for or against WIC's effectiveness with respect to the outcomes that the Senate Committee asked us to examine. We found that even when several evaluations have focused on the same outcomes, the evidence still falls short of being conclusive.

Several studies are of high or moderate quality, but they are not sufficient for drawing conclusions regarding the overall effectiveness of the nationwide WIC program. WIC is funded by the federal government but is administered at the state and local levels. The state programs and local projects can differ considerably in how they carry out the WIC program and in the types of populations they serve. The majority of the studies we reviewed focused on only one or a few projects or on specific geographic areas. Therefore, even where their results are positive, these evaluations provide evidence only that WIC can be effective under the conditions that were studied.

The shortage of credible evaluative information does not mean that the WIC program is not effective. It means that there is not enough clearly indisputable evidence from which to draw a firm conclusion. It is possible that WIC has certain positive effects on its participants but that the designs, sample sizes, and measures that were used in the evaluations that have been performed were not always sensitive enough to detect beneficial changes in women and children who receive WIC services. Even where statistically significant effects have been found, questions remain about whether they are clinically meaningful.

The absence of the kind of evaluative information that is necessary for making strong general judgments about WIC's effectiveness does not mean that the existing reports are not useful. Some WIC studies may have been useful in collecting information for decisionmakers on the implementation and operational issues

of local programs. It was not our objective to examine them for this information.

Many of the documents we reviewed did not adequately describe the design, execution, and analyses of the evaluation effort. That this information was missing made it difficult for us to determine the technical adequacy of these studies. In some instances, it may have excessively limited our confidence in the findings.

We did not combine the studies in which we have little confidence with those that we rated moderate or high in quality. The studies that we rated low in quality are so severely flawed that they simply do not provide useful findings for our synthesis. To include their findings could be misleading.

The following are the methodological problems that are generally characteristic of WIC evaluations:

--they lack research designs that are adequate for establishing cause and effect (such as a causal relationship between participating in WIC and positive outcomes);

--the indexes they use for measuring nutrition are neither precise nor standardized, and experts do not agree on what the indicators of nutritional inadequacy are;

--the data are of questionable quality because collection and reporting are not sufficiently uniform and consistent;

--they do not present firm details about the WIC intervention that was studied;

--they do not separate the effects of the individual components of WIC or of WIC from the effects of other programs;

--they do not analyze the relationships between a mother's nutrition, her pregnancy, and the health of her children during the early years of life;

--they do not build upon past research and are not designed to enable subsequent studies to use their results.

These problems have been identified before (see, for example, GAO, 1975, 1979, and 1981; FNS, 1976; Lawrence, 1981; Hayes, 1982; and Dwyer, 1983). They continue to be problems, but some progress can be seen in the improved designs and methodologies that have been used in some recent evaluation efforts. Metcoff's 1982 study, for example, was able to implement a stronger experimental design by randomly assigning pregnant women to WIC and to a control group. The Centers for Disease Control has developed age-specific reference curves for identifying anemia, an improvement over the earlier practice of using arbitrary cutoff points as indicators.

Some of these advances might not have been made if the difficulties we have described had not been identified in WIC evaluations and received attention. The Research Triangle Institute's national WIC evaluation for FNS now under way has given considerable emphasis to reviewing past evaluation difficulties in order to guide the design of the new assessment. Its findings are to be reported during 1984.

In summary, previous reviewers of WIC evaluation studies have identified problems with their design or methodology that affect the credibility of the findings. The conclusions about WIC's effectiveness have ranged between two extremes. Either

--design and methodology problems and program complexity impose such severe constraints that a meaningful overall assessment of the WIC program is not really possible or

--a substantial body of evidence from WIC evaluations now exists and indicates that the WIC program is having a positive and significant effect on its participants.

Our review takes a position between these two extremes. We find some sound, but not conclusive, evaluative evidence of favorable program effects on birthweights and little credible evidence for several other measures of effectiveness. That the evaluations are not adequate for determining whether or not the WIC program is having the effect that was intended in the legislation underscores the necessity of designing and implementing evaluations that can provide the information that the Congress needs. It seems likely that past experience will make it possible to produce more precise information that will fill in the gaps we have identified.

APPENDIX I

United States Senate

COMMITTEE ON
AGRICULTURE, NUTRITION, AND FORESTRY
WASHINGTON, D.C. 20510

June 30, 1983

The Honorable Charles Bowsher
Comptroller General of the United States
General Accounting Office
Washington, D. C. 20548

Dear Mr. Bowsher,

As you know, the Senate Committee on Agriculture, Nutrition, and Forestry has jurisdiction over the Special Supplemental Food Program for Women, Infants, and Children, often referred to as the WIC program. The Committee and the Subcommittee on Nutrition have conducted hearings in recent years on the program's effectiveness. Increasingly, there appears to be conflicting testimony about what is actually known about the program's effectiveness from evaluations which have been conducted for this purpose.

Specifically, findings from various WIC evaluations have been cited to support contentions that the program is effective in improving a variety of maternal and child health conditions among participants. However, others have critized the studies as being methodologically unsound and the findings insufficient for national representation.

The current authorization for WIC expires at the end of fiscal year 1984. Next year, the Committee will have among its highest priorities consideration on reauthorizing the WIC program.

In anticipation of these important deliberations, I would request that the General Accounting Office undertake an objective analysis of the evaluations which have attempted to assess the WIC program and the extent to which the evaluations support the claims being made. Because the GAO analysis will be an important background for the Committee's deliberations which will begin early in 1984, I would ask that this request be given high priority by GAO, and that the written analysis be completed by the week of January 9, 1984.

I want to emphasize that I am not requesting a new study of the WIC program itself, but rather a careful examination of existing research to determine the technical and methodological soundness of these evaluations and the credibility of the claims which have been made based on them.

APPENDIX I

It is my understanding that the evaluations on this subject are relatively few in number. However, if you should discover additional ones, concentration should be given to those which you determine are more methodologically sound and reflect more recent data.

In analyzing the evaluations, it would be useful to distinguish between the impact from nutrition provided by the WIC food supplement versus the improved health care which some contend accompanies WIC participation.

Specifically with regard to maternal health, the analysis should focus on the impact of WIC on miscarriages and still births (neonatal deaths) and on the nutritional status of mothers. Additionally, the analysis should examine the claims that positive pregnancy outcome is especially strong in "high risk" WIC mothers and that as the length of participation in WIC increases, the positive effects also increase.

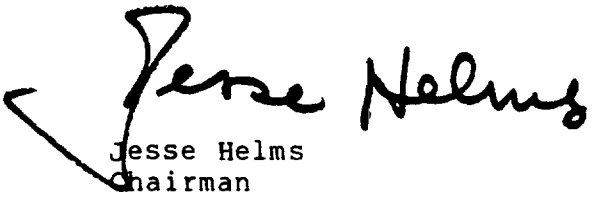
In relation to infant and child outcomes, the analysis should examine the evalutions which purport to find that WIC has a positive effect on increasing the birth weights of infants and the claims that the program reduces the chances of anemia and, if feasible to examine, mental retardation.

Members of the Committee staff have been in touch for several weeks with GAO staff from the Institute for Program Evaluation. I understand that GAO has developed a process of review called evaluation synthesis, in response to similar requests, which sounds as if it would meet our need to learn what is really known from the evaluations about the impact of WIC program participation. The staff will, of course, be available to assist in establishing the initial parameters of the analysis and other assistance as you deem appropriate.

Both because of the time constraint and because this analysis is not a review of actual program operations, I do not regard it as essential to obtain USDA comment prior to publication.

Many thanks for your consideration of this request.

Sincerely,



Jesse Helms
Chairman

BIBLIOGRAPHY

Argeanas, S., and I. Harrille. "Nutrient Intake of Lactating Women Participating in the Colorado WIC Program." Nutr. Repts. International, December 1979, pp. 805-10.

Arizona WIC data. 1976-77 survey of participant nutrition and pregnancy outcomes, transmitted to GAO from Food and Nutrition Service, Washington, D.C., August 1983.

-----. 1977-78 evaluation of health intervention and nutrition education, transmitted to GAO from Food and Nutrition Service, Washington, D.C., August 1983.

Bailey, L.B., C.S. Mahan, and D. Dimperio. "Folacin and Iron Status in Low-Income Pregnant Adolescents and Mature Women." Amer. J. Clin. Nutr., 33 (September 1980), pp. 1997-2001.

Bailey, L.B., et al. "Vitamin B_6, Iron and Folacin Status of Pregnant Women." Nutrition Research, 3 (1983), 783-93.

Baxter, Jane (Tennessee Department of Public Health, Nashville). "Reduced Frequency of Low Weight Births Among Women Receiving WIC in Tennessee." Attachment to letter to National WIC Evaluation, Food and Nutrition Service, Alexandria, Va., June 21, 1983.

Belshaw, J. "WIC Among the Navajos." Community Nutritionist, 1:3 (May-June 1982), 10-12.

Bendick, M., et al. Efficiency and Effectiveness in the WIC Program Delivery System. Washington, D.C.: The Urban Institute, September 1976.

Berkerfield, J., and J.B. Schwartz. "Nutrition Intervention in the Community: The WIC Program." New Eng. J. Med., 302:10 (March 1980), 579-81.

Brevard County Health Department. "Report of WIC Data." Rockledge, Fla., July 25, 1977.

Caan, B., and S. Marger. "Evaluating Evaluations." University of California, Department of Social and Administrative Health Sciences, School of Public Health, Berkeley, Calif., 1983.

California WIC data. 1980 survey of nutrition education, transmitted to GAO from Food and Nutrition Service, Washington, D.C., August 1983.

Carabello, D., et al. "An Evaluation of WIC." Masters thesis, Yale School of Medicine, Department of Epidemiology and Public Health, New Haven, Conn., May 1978.

Centers for Disease Control. Analysis of Nutrition Indices for Selected WIC Participants. Atlanta: December 1977.

Centers for Disease Control Nutrition Surveillance: Annual Summary 1980. Atlanta: November 1982.

Christie, D.D., and L.B. Gale. "WIC Program Involvement in the Prevention of Mental Retardation." New Jersey State Department of Health, Trenton, N.J., 1979.

Collins, T., D. Leeper, and S. DeMellier. "Integration of WIC Program with Other Infant Mortality Programs." Final report, Appalachian Region Commission Report, University of Alabama, University, Ala., September 30, 1981.

Cook, J.D. "Iron Deficiency: Methods to Measure Prevalence and Evaluate Interventions." Nutrition Intervention Strategies in National Development, pp. 257-63. New York: Academic Press, 1983.

Deterding, J., A. Wickiser, and J.L. Smith. "The Benefits of the WIC Program in Three Indian Communities." University of Nebraska Medical Center, Omaha, Neb., May 1983.

Development Associates. "Evaluation of the WIC Migrant Demonstration Project: A Final Report." N.p., May 1979.

Drayton, P.K.D. "Evaluation of the WIC Nutrition Education Intervention Program for High-Risk Pregnant Women in Illinois." Ph.D. thesis, Southern Illinois University, Carbondale, Ill., 1982.

Dwyer, J. "Case Study of a National Supplementary Feeding Program: The WIC Program in the United States." Nutrition Intervention Strategies in National Development, ed. Barbara A. Underwood. Cambridge, Mass.: Massachusetts Institute of Technology, 1983.

Edozien, J.C., B.R. Switzer, and R.B. Bryan. Medical Evaluation of the Special Supplemental Food Program for Women, Infants and Children (WIC), 6 vols. Chapel Hill, N.C.: University of North Carolina, School of Public Health, July 15, 1976.

-----. "UNC Medical Evaluation of WIC." Amer. J. Clin. Nutr., 32 (March 1979), 677-92.

Endres, J., and M. Sawicki. "Food and Nutrient Intake of 7,728 Illinois Infants and Children, 1978-1979." Supplementary report, Southern Illinois University, NDDA Laboratory, Carbondale, Ill., August 1980.

-----, and J. Casper. "Dietary Assessment of Pregnant Women in a Supplemental Food Program." Amer. Diet. Assoc. J., 79:8 (August 1981), 121-26.

Evaluation Research Society. Standards for Program Evaluation. Washington, D.C.: May 1980.

Fleshood, H., et al. "Is WIC Reducing the Prevalence of LBW and Infant Mortality?" Presented at the annual meeting of the American Public Health Association, Los Angeles, October 1978.

Food and Nutrition Service. Evaluating the Nutrition and Health Benefits of the Special Supplemental Food Program for Women, Infants and Children. Washington, D.C.: U.S. Department of Agriculture, Advisory Committee on Nutrition Evaluation, November 1977.

-----. "State and Local Agency Evaluations of the WIC Program." Report on the Special Supplemental Food Program, U.S. Department of Agriculture, Washington, D.C., [November 1977].

-----. Evaluation of the Effectiveness of WIC. Washington, D.C.: U.S. Department of Agriculture, 1981.

-----. A Response to Graham's Review of the Supplemental Food Program for Women, Infants, and Children (WIC). Washington, D.C.: U.S. Department of Agriculture, Office of Policy, Planning, and Evaluation, December 1981.

Friends of the Earth. Would the Federal Government Make a Profit by Doubling the Budget of the WIC Program for Pregnancy? Washington, D.C.: February 1983.

GAO (U.S. General Accounting Office). Preliminary Report on the Special Supplemental Food Program, B-176994. Washington, D.C.: September 1973.

-----. Observations on Evaluation of the Special Supplemental Food Program, Food and Nutrition Service, CED-75-310. Washington, D.C.: December 18, 1975.

-----. National Nutrition Issues, CED-78-7. Washington, D.C.: December 1977.

-----. The Special Supplemental Food Program for Women, Infants, and Children (WIC): How Can It Work Better? CED-79-55. Washington, D.C.: February 1979.

-----. "Comments on Evaluation Studies of WIC." Testimony for the Chairman, Senate Subcommittee on Agriculture, Rural Development and Related Agencies, Washington, D.C., 1981.

-----. The Evaluation Synthesis, Institute for Program Evaluation Methods Paper I. Washington, D.C.: April 1983.

George, N.N. "Prepregnancy Weights, Weight Gains and Other Factors Related to Birthweight of Infants Born to Overweight Women." Masters thesis, Bowling Green State University, Bowling Green, Ohio, June 1982.

Georgia Department of Human Resources. "WIC Nutrition Survey." Atlanta, April 1982.

Goldberg, H. "An Evaluation of the Effectiveness of the WIC Program in Terms of Height, Birthweight, Weight, and Hematocrit." Paper on Harlem Hospital Medical Center (N.Y.) WIC program, privately circulated, May 17, 1982.

Hawaii WIC data. 1982 survey of participant nutrition, transmitted to GAO from Food and Nutrition Service, Washington, D.C., August 1983.

Hayes, C.D. Making Policies for Children: A Study of the Federal Process. Washington, D.C.: National Academy Press, 1982.

Healthwise, Inc. "WIC Nutrition Education Evaluation." Final report, Boise, Idaho, September 30, 1980.

Heimendinger, J. "The Effect of WIC on Growth of Children." Ph.D. thesis, Harvard University School of Public Health, Nutrition Department, Boston, 1981.

-----, and N. Lairde. "Growth Changes: Measuring the Effect of an Intervention." Evaluation Review, 7:1 (February 1983), 80-95.

Hicks, L.E., R.A. Langham, and J. Takenaka. "Cognitive and Health Measures Following Early Nutritional Supplementation: A Sibling Study." Amer. J. Pub. Health, 72 (October 1982), 1110-18.

-----. "Interpretation of Behavioral Findings in Studies of Nutritional Supplementation." Amer. J. Pub. Health, 73 (June 1983), 695-97.

Horgen, D.M., P.A. Loris, and D.M. Rose. "Retrospective Study of the Special Supplemental Food Program for Women, Infants, and Children (WIC)." California Department of Health Services, Community Health Services Division, Sacramento, Calif., April 1982.

Hunterdon Medical Center. "Innovative Workshops in Nutrition Education." Final report, Flemington, N.J., October 1982.

Jarka, E., J. Maass, and S. Elridge. "Nutritional Risks Identified in Pregnant Adolescents Participating in the Illinois WIC Program." Illinois Department of Public Health, Springfield, Ill., 1981.

Kautz, L., and G.G. Harrison. "Comparison of Body Proportions of One-Year-Old Mexican American and Anglo Children." Amer. J. Pub. Health, 71 (March 1981), 280-82.

Kennedy, E.T., J.E. Austin, and C.P. Timmer. "Cost/Benefit and Cost/Effectiveness of WIC." Privately circulated, n.p., n.d.

Kennedy, E.T., and S. Gershoff. "Effect of WIC Supplemental Feeding on Hemoglobin and Hemocrit of Prenatal Patients." J. Amer. Diet Assoc., 80:3 (March 1982), 227-30.

Kennedy, E.T., et al. "Evaluation of the Effect of WIC Supplemental Feeding on Birth Weight." Amer. Diet. Assoc. J., 80:3 (March 1982), 220-27.

Kotelchuck, M., et al. "1980 Massachusetts Special Supplemental Food Program for Women, Infants, and Children (WIC) Evaluation Project." Massachusetts Department of Public Health, Division of Family Health Services, Boston 1981.

-----. "Massachusetts Special Supplemental Food Program for Women, Infants, and Children (WIC) Follow-Up Study." Massachusetts Department of Public Health, Division of Family Health Services, Boston 1982.

Langham, R.A., et al. "Impact of the WIC Program in Louisiana." Privately circulated, n.p., n.d.

Lawrence, J.E.S., et al. Evaluation of the WIC Program: Predesign Activities Phase I Final Report. Washington, D.C.: U.S. Department of Agriculture, Food and Nutrition Service, and Research Triangle Institute, August 1981.

Leddy, P. "Improving Diet Assessment Methods of Pregnant and Breastfeeding Women." University of Rhode Island, Kingston, R.I., 1982.

Mahan, C.S. "Revolution in Obstetrics: Pregnancy Nutrition." J. Florida M.A., 66:4 (April 1979), 367-72.

-----, and C. Sharbaugh. "North Central Florida WIC Evaluation." Centers for Disease Control, Nutrition Surveillance, Atlanta, June 1976.

Massachusetts Department of Public Health. "1983 Massachusetts Nutrition Survey." Boston, October 1983.

Mauer, A.M. "The WIC Program Tying Supplemental Foods to Nutritional Needs." J. Florida M.A., 66:4 (April 1979), 453-56.

Metcoff, J., et al. "Nutrition in Pregnancy (NIP)." Final report, University of Oklahoma Health Science Center, Oklahoma City, Okla., December 1982.

Michigan Department of Public Health. "Evaluation of Health Services Utilization in Rural Areas and Among Migrant Farmworkers." Lansing, Mich., 1980.

Montana Department of Health and Environmental Services. "WIC Teenage Pregnancy Outcome Project." Report on U.S. Department of Agriculture Grant 59-3198-9-82, Helena, Mont., January 1982.

National Research Council. Nutrition Services in Perinatal Care. Washington, D.C.: National Academy Press, 1981.

NDDA Laboratory, Human Development (Food/Nutrition). "Report of Nutritional Factors of Pregnant Women in the Illinois Department of Public Health Special Supplemental Food Program for WIC." Southern Illinois University, Carbondale, Ill., August 1980.

Norad, D., et al. Evaluation of the Commodity Supplemental Food Program Final Report: Health and Nutrition of Three Local Projects. Washington, D.C.: U.S. Department of Agriculture, Food and Nutrition Service, 1982.

Northrop, P. "Perceptions and Knowledge of Breastfeeding Among WIC and non-WIC Pregnant Women." University of Alabama, University, Ala., 1982.

Nutt, P.C., M. Wheeler, and R.A. Wheeler. "Social Program Evaluation Revisited: The WIC Program." Ohio State University, Graduate Program in Hospital and Health Services Administration, Columbus, Ohio, n.d.

Paige, D.M. "Medical Assistance Cost and Utilization Patterns in WIC Enrollees." A joint study of John Hopkins University, School of Hygiene and Public Health, and Maryland Department of Health and Mental Hygiene, Baltimore, Md., June 1983.

-----. "Evaluation of the WIC Program for Infants on the Eastern Shore of Maryland." Johns Hopkins University, Baltimore, Md., 1983.

Pelto, J.M. "Results of a Nutrition Intervention Program: The WIC Program in Alaska." Alaska Medicine, 24:2 (March-April 1982), 14-17.

Pestronk, R.M. "Reasons for Enrollment in the Michigan WIC Program." Michigan Department of Public Health, Lansing, Mich., 1982.

Pollitt, E., and Lorimor, R. "Effects of WIC on Cognitive Development." Amer. J. Pub. Health, 73 (June 1983), 698-700.

Robbins, G.A. "Surveillance of Nutritional Status in the United States." Public Health Currents (Ross Labs), 20:2 (March-April 1980), 5-8.

Rosenberg, M.J., B.J. McCarthy, and J.S. Terry. "The Effect of the WIC Program on Infant Mortality in Rural Georgia." Centers for Disease Control, Atlanta, 1981.

Rush, D. "Effects of Changes in Protein and Calorie Intake During Pregnancy on the Growth of the Human Fetus." Effectiveness and Satisfaction in Antenatal Care, eds. M. Enkin and I. Chalmers. Philadelphia: Lippincott, 1982.

-----. "The Behavioral Consequences of Protein-Energy Deprivation and Supplementation in Early Life: An Epidemiologic Perspective." Nutrition and Behavior, ed. J.R. Galler. New York: Plenum Publishing, 1982.

-----. "Is WIC Worthwhile?" Amer. J. Pub. Health, 72 (October 1982), 1101-03.

-----. "In Response to Hicks, et al." Amer. J. Pub. Health, 73 (June 1983), 700-01.

Rye, J., M. White, and M. Majchrzak. "Does Objective Based Health Education Effect Positive Changes in the Health Status of WIC Program Clients?" Preliminary report, State Department of Health Service, Bureau of Nutrition Services, Phoenix, Ariz., April 7, 1978.

Schelzel, G., and M.A. Britton. "An Assessment of the WIC Program in Pennsylvania." Pennsylvania Health Department, Pittsburgh, Pa., January 1978.

Schramm, W.F. "WIC Prenatal Participation and Its Relationship to Newborn Medicaid Costs in Missouri: A Cost/Benefit Analysis." Draft report, Missouri Center of Health Statistics, Jefferson City, Mo., April 26, 1983.

Schuster, K., L. Bailey, and C. Mahan. "Vitamin B_6 Status of Low-Income Adolescent and Adult Pregnant Women and the Condition of Their Infants at Birth." Amer. J. Clin. Nutr., 34 (September 1981), 1731-35.

Shanklin, D.S. "What Has WIC Accomplished?" U.S. Department of Agriculture, Food and Nutrition Service, Washington, D.C., October 1981.

Silverman, P.R. "Allegheny County, Pa., Health Department WIC Evaluation." Pittsburgh, Pa., December 1981.

Slonim, A., K. Kolasa, and M. Bass. "The Cultural Appropriateness of the WIC Program in Cherokee, North Carolina." Amer. Diet. Assoc. J., 79:8 (August 1981), 164-68.

Stockbauer, J. "Evaluation of the Prenatal Participation Component of the Missouri WIC Program." Draft report, Missouri Division of Health, Jefferson City, Mo., April 26, 1983.

Thenen, S.W. "Folacin Content of Supplemental Foods for Pregnancy." Amer. Diet. Assoc. J., 80:3 (March 1982), 237-41.

Thomason, C.F. (Louisiana Department of Health and Human Resources, New Orleans). "Low Birth Weight Study." Attachment to letter to National WIC Evaluation, Food and Nutrition Service, Alexandria, Va., May 16, 1983.

U.S. Senate, Committee on Agriculture, Nutrition, and Forestry. Compilation of Selected Federal Nutrition Studies, 96th Cong., 1st sess. Washington, D.C.: January 1979. Pp. 77-80.

-----, Committee on Agriculture, Nutrition, and Forestry, Subcommittee on Nutrition. Oversight on Federal Nutrition Programs, 97th Cong., 2nd sess. Washington, D.C.: February 22-23, 1982.

-----. Oversight on Nutritional Status of Low-Income Americans in the 1980's, 98th Cong., 1st sess. Washington, D.C.: April 1983.

-----, Subcommittee on Rural Development, Oversight, and Investigations. Oversight on Federal Nutrition Programs, 98th Cong., 1st sess. Washington, D.C.: March 1983.

-----, Select Committee on Nutrition and Human Needs. "WIC Program Survey 1975." Working paper, 94th Cong., 1st sess., Washington, D.C., April 1975.

-----. "Medical Evaluation of the Special Supplemental Food Program for Women, Infants and Children." 94th Cong., 2nd sess., Washington, D.C., August 1976.

Weiler, P., et al. "Anemia as a Criterion for Evaluation of WIC." Pediatrics, 63 (April 1979), 584-90.

Wholey, J.S., and M.S. Wholey. "Toward Improving the Outcome of Pregnancy: Implications for the Statewide Prenatal Program." Report for the Tennessee Department of Public Health, Wholey Associates, Arlington, Va., June 1982.

Williams, J. "Wyoming WIC Evaluation." State Department of Health and Social Services, Cheyenne, Wyo., 1982.

THE WIC EVALUATIONS RELEVANT TO OUR QUESTIONS

The chart on pages 70-75 lists the 61 evaluation studies that contain information that answers our evaluation questions. We have included studies whose primary focus was not on the outcomes of interest to the Committee if the studies contain other information that is relevant to them. As the chart shows, we looked at each element of the studies that used several investigative strategies or that provide information for more than one of our questions.

Our questions are about the effects of the WIC program on birthweights, fetal and neonatal mortality, maternal nutrition, and anemia and mental retardation in infants and children. (See chapter 1, table 1.) In the evaluation reports, we looked first for evidence about overall effects and then for information about the differences in effectiveness with regard to the participants with greater health and nutrition risks ("risk groups"), the length of participation in WIC ("time in program"), and the three individual WIC services--food supplementation, nutrition counseling, and adjunct health care ("WIC components").

We use "++" and "+" to indicate the elements of the evaluations that are relevant to our evaluation questions. The "++" indicates that we gave the element a rating of high or medium in quality and credibility; the "+" indicates that we gave the element a low rating.

APPENDIX III

Report	Birthweight				Mortality				Maternal nutrition				Childhood anemia				Mental retardation			
	Overall	Risk groups	Time in program	WIC components	Overall	Risk groups	Time in program	WIC components	Overall	Risk groups	Time in program	WIC components	Overall	Risk groups	Time in program	WIC components	Overall	Risk groups	Time in program	WIC components
Argeanas and Harrille (1979)									+											
Arizona WIC Data 1976-77 (1983)	+												+							
Arizona WIC Data 1977-78 (1983)	+												+							
Bailey, Mahan, and Dimperio (1980)									+											
Bailey et al. (1983)	++	++							++	++										
Baxter (1983)	+	+	+																	
Belshaw (1982)													+			+				
Brevard County Health Dept. (1977)	+								+				+							
Carabello et al. (1978)																				
Part 1					+	+														
Part 2	+												+							

Report	Birthweight				Mortality				Maternal nutrition				Childhood anemia				Mental retardation			
	Overall	Risk groups	Time in program	WIC components	Overall	Risk groups	Time in program	WIC components	Overall	Risk groups	Time in program	WIC components	Overall	Risk groups	Time in program	WIC components	Overall	Risk groups	Time in program	WIC components
Centers for Disease Control (1977)	+												++	++	++					
Centers for Disease Control (1982)	+	+																		
Christie and Gale (1979)													+	+						
Collins, Leeper, and DeMellier (1981)	+	+							+											
Deterding, Wickiser, and Smith (1983)													+							
Drayton (1982)									+			+								
Edozien, Switzer, and Bryan (1976)	+	+	+		+				++	++	++		++	++	++	++				
Endres, Sawicki, and Casper (1981)									++											
Fleshood et al. (1978)	+		+		+		+													

APPENDIX III

Report	Birthweight: Overall	Birthweight: Risk groups	Birthweight: Time in program	Birthweight: WIC components	Mortality: Overall	Mortality: Risk groups	Mortality: Time in program	Mortality: WIC components	Maternal nutrition: Overall	Maternal nutrition: Risk groups	Maternal nutrition: Time in program	Maternal nutrition: WIC components	Childhood anemia: Overall	Childhood anemia: Risk groups	Childhood anemia: Time in program	Childhood anemia: WIC components	Mental retardation: Overall	Mental retardation: Risk groups	Mental retardation: Time in program	Mental retardation: WIC components
Food and Nutrition Service (1977)																				
Arizona	+												+							
Michigan	+								+				+							
Oregon													+							
Texas	+								+	+			+	+						
Virgin Islands	+		+										+		+					
George (1982)	+	+							+	+										
Georgia Dept. of Human Resources (1982)													+							
Goldberg (1982)	+												+		+					
Hawaii WIC Data (1983)	+								+				+		+					
Heimendinger (1981)	+	+																		
Hicks, Langham, and Takenaka (1982)	+												+				+[a]	+[a]		

[a]This study focused on cognitive development, not the incidence of mental retardation.

Report	Birthweight				Mortality				Maternal nutrition				Childhood anemia				Mental retardation			
	Overall	Risk groups	Time in program	WIC components	Overall	Risk groups	Time in program	WIC components	Overall	Risk groups	Time in program	WIC components	Overall	Risk groups	Time in program	WIC components	Overall	Risk groups	Time in program	WIC components
Horgen, Loris, and Rose (1982)	+	+	+		+	+			+	+			+	+						
Hunterdon Medical Center (1982)									+			+	+			+				
Jarka, Maass, and Elridge (1981)		+	+						+											
Kennedy, Austin, and Timmer (n.d.)	+								+											
Kennedy and Gershoff (1982)									‡	‡	‡	‡								
Kennedy et al. (1982)	‡	‡	‡	‡					b	b										
Kotelchuck et al. (1981)	‡	‡	‡		+		+													
Kotelchuck et al. (1982)	+	+	+		+															
Langham et al. (n.d.)													+		+					

[b]In our report, we included data from Kennedy et al. (1982) on maternal weight in our discussion of data from Kennedy and Gershoff (1982) on maternal nutrition.

APPENDIX III

Report	Birthweight: Overall	Birthweight: Risk groups	Birthweight: Time in program	Birthweight: WIC components	Mortality: Overall	Mortality: Risk groups	Mortality: Time in program	Mortality: WIC components	Maternal nutrition: Overall	Maternal nutrition: Risk groups	Maternal nutrition: Time in program	Maternal nutrition: WIC components	Childhood anemia: Overall	Childhood anemia: Risk groups	Childhood anemia: Time in program	Childhood anemia: WIC components	Mental retardation: Overall	Mental retardation: Risk groups	Mental retardation: Time in program	Mental retardation: WIC components
Mahan and Sharbaugh (1976)																				
Part 1	+																			
Part 2	+																			
Part 3	+		+																	
Metcoff et al. (1982)	‡	‡	‡		+				‡											
Montana Dept. of Health and Environmental Services (1982)	+	+		+					+			+								
NDDA Laboratory (1980)									‡	‡										
Nutt, Wheeler, and Wheeler (n.d.)	+	+	+																	
Paige (1983)													+							
Pelto (1982)	+												+							
Robbins (1980)													+							
Rosenberg, McCarthy, and Terry (1981)					+	+	+													

APPENDIX III

Report	Birthweight				Mortality				Maternal nutrition				Childhood anemia				Mental retardation			
	Overall	Risk groups	Time in program	WIC components	Overall	Risk groups	Time in program	WIC components	Overall	Risk groups	Time in program	WIC components	Overall	Risk groups	Time in program	WIC components	Overall	Risk groups	Time in program	WIC components
Rye, White, and Majchrzak (1978)	+		+	+								+								
Schelzel and Britton (1978)					+		+		+		+		+							
Schramm (1983)	+				+															
Schuster, Bailey, and Mahan (1981)									+											
Silverman (1981)	++	++																		
Stockbauer (1983)	++	++	++		+	+														
Thenen (1982)									+											
Thomason (1983)	+	+	+																	
U.S. Senate Select Committee on Nutrition and Human Needs (1975)	+				+				+				+							
Weiler (1979)													+							
Williams (1982)	+		+																	