Food and Nutrition Service

U.S. DEPARTMENT OF AGRICULTURE

WIC Nutrition Assessment and Tailoring Study: Final Report Part 2 – Appendix I

U.S. Department of Agriculture | Food and Nutrition Service | Office of Policy Support

WIC Nutrition Assessment and Tailoring Study: Final Report Part 2 – Appendix I

Submitted to

USDA Food and Nutrition Service
Office of Policy Support
1320 Braddock Place
Alexandria, VA 22314

Project Officers

Dr. Karen Castellanos-Brown
Alexander Bush

This WIC Nutrition Assessment and Tailoring Study was conducted by Westat under Contract No. GS00F009DA, funded by the U.S. Department of Agriculture's Food and Nutrition Service. The findings and conclusions in this report are those of the authors and should not be construed to represent any official USDA or U.S. Government determination or policy.

In approving Ohio WIC's participation in this study, the Ohio Department of Health Institutional Review Board requires that publication of research findings include the following disclaimer: "This study includes data provided by the Ohio Department of Health, which should not be considered an endorsement of this study or its conclusions."

In approving Nebraska WIC's participation in this study, the Nebraska Department of Health and Human Services requires that publication of research findings include the following disclaimer: "These findings and their interpretation are the sole responsibility of the author and do not necessarily reflect the opinions or policy position of DHHS."

Suggested citation

Gabay, M., Wroblewska, K., Borger, C., Manglitz, C., Neenan, R., and Shah, N. (2024). *WIC Nutrition Assessment and Tailoring Study: Final Report Part 2*. Prepared by Westat, Inc., Contract No. GS00F009DA. Alexandria, VA: U.S. Department of Agriculture, Food and Nutrition Service, Office of Policy Support, Project Officer: Karen Castellanos-Brown.

Submitted by

Westat
An Employee-Owned Research Corporation®
1600 Research Boulevard
Rockville, Maryland 20850-3129
(301) 251-1500

Appendix I

Regression Methods

Table of Contents

I.1. Methodological Approach

The approach to regression modeling began with preparing the datasets for analysis. This included the following activities:

- Each set of remote and in-person datasets were appended together – observation-level, participant interview, and staff interview.
- The observation and participant datasets were merged. The staff interview data were merged onto the observation-level data.
- Summary statistics of observation-level data were created for use in staff-level regression models.
- Local Agency Director Survey data were merged onto the staff-level data.

In the course of this preparatory work, analysis of the data contents revealed one observation with over 6 hours total duration. It was deemed an extreme outlier and excluded from all regression models involving time.

Once the datasets had been prepared, Westat constructed the dependent variables necessary to explore the research questions of interest. This included creating the following variables:

- Binary indicators of tailoring: tailoring food package, tailoring education, and tailoring referral.
- Total time participants spent per WIC observation was calculated by summing the minutes of time (including waiting time) across all WIC sessions in one observation. The result was used as a continuous dependent variable.
- Average time spent per WIC observation at the staff level was calculated by summing the minutes of time (including waiting time) across all WIC sessions in each observation, and then calculating the average value for each staff member. The result was used as a continuous dependent variable.
- Average waiting time per WIC observation at the staff level was calculated by summing the minutes of client waiting time across all WIC sessions in each observation and calculating the average value for each staff member. The result was used as a continuous dependent variable.
- Staff efficiency was constructed as the proportion of total WIC observation time staff typically spent with patients as opposed to the waiting time spent by the client, the arithmetic formula being:

  1 – (average staff-level client waiting time / average staff-level observation time).

  This measure of staff efficiency was used as a continuous dependent variable.

- Factor analysis was performed on a set of five staff interview questions to construct a composite measure of staff satisfaction. The result was used as a continuous dependent variable.

Analysis of desired independent variables identified the following:

- Responses to the participant satisfaction interview items did not have sufficient variability to use as a dependent variable in a regression model.

- Responses to the clinic flow items did not have sufficient variability to use as a dependent variable in a regression model.

Consequently, planned regression models using these variables were not pursued.

Once the dependent variables had been created, bivariate analyses was used to assess correlations between each dependent variable and independent variables of interest. When the dependent variable was binary (no/yes), logistic regression was used. When the dependent variable was continuous (non-categorical) we used ordinary least squares (OLS) linear regression. Coefficient significance was examined, and measures of model fit (adjusted R-squared or pseudo-R-squared) values were reviewed to assess whether the independent variable was associated with the dependent variable. This bivariate work informed variables to evaluate for inclusion in regression models.

In addition to variables of focal interest or those identified via bivariate analyses, a few additional independent variables were occasionally included in regression models—for example, if it was of interest to confirm the lack of independent association. In general, including the same set of characteristics in all models allowed for greater comparability between models. For observation-level regression models, these variables included types of nutrition risk codes (Anthropometric, Biochemical, Clinical/Health/Medical, Dietary, and Other), race/ethnicity, whether sessions were held in person or remotely, type of staff conducting the nutrition assessment, years of staff experience for staff conducting the nutrition assessment, size and urbanicity of the WIC clinic, type of WIC clinic, and reason for WIC session (i.e., enrollment or recertification). For staff-level regression models, these variables included whether a participant was asked to complete form(s) prior to appointment, the number of different types of WIC staff that perform the nutrition risk assessment, management information system (MIS) food package tailoring, staff experience, size/urbanicity of WIC clinic, and type of WIC clinic.

All regression models were adjusted for State agency, as unique State effects were assumed to be important. However, coefficient estimates or odds ratios on State agency variables are not discussed within the report.

A p-value of 0.05 determined the statistical significance of coefficients in regression models.

I.2. Regression Analyses Found Not Feasible

Participant satisfaction with their WIC appointment was determined to have insufficient variability to construct a meaningful dependent variable for regression analysis. Tables I-1 and I-2 below show tabulations of responses of the two items of interest to potentially construct a dependent variable. In Table I-1, over 98 percent of cases responded "Strongly Agree" or "Agree" and less than half a percent of cases "Strongly Disagreed" or "Disagreed." In Table I-2, over 98 percent of cases responded "Very satisfied" or "Somewhat satisfied" and no cases responded "Very dissatisfied or Somewhat dissatisfied." It was assumed that estimating factors independently associated with differences in responses of "Strongly Agree" or "Agree" and "Very satisfied" or "Somewhat satisfied" would not be informative or actionable.

Table I-1. Tabulation of participant interview question 19, showing insufficient variation to proceed with regression analysis

Participant Interview Q19: Agree/disagree: the WIC staff made you feel at ease	Number	Percentage
Strongly Agree	423	79.0
Agree	106	19.8
Undecided	4	0.8
Disagree	2	0.4
Strongly Disagree	0	0.0
Missing	1	0.0
Total	536	100

Note: Percentages may not add to 100 due to rounding.
Source: Calculated from Appendix D, Table 5c.

Table I-2. Tabulation of participant interview question 32, showing insufficient variation to proceed with regression analysis

Participant Interview Q32: Satisfaction with the visit/appointment as a whole	Number	Percentage
Very satisfied	489	91.2
Somewhat satisfied	40	7.5
Neither satisfied nor dissatisfied	1	0.2
Somewhat dissatisfied	0	0.0
Very dissatisfied	0	0.0
Missing	6	1.1
Total	536	100

Note: Percentages may not add to 100 due to rounding.
Source: Calculated from Appendix D, Table 7c.

Clinic flow as perceived by staff also did not have a sufficient number of cases across the distribution to construct a meaningful dependent variable for use in a regression model (Table I-3). The question was only asked of staff conducting in-person appointments. About 70 percent responded "Strongly Agree" or "Agree" and about 18 percent responded "Strongly Disagree" or "Disagree," which, on the surface, suggested sufficient variability; however, the 18 percent

represented only 16 cases. There is not one agreed-upon cut-off for an insufficient number of cases when constructing a regression model. This particular regression analysis only used half the sample because the staff had to be in-person to observe clinic flow. After excluding the undecided,[1] there were at most 77 observations for the analysis, a relatively small number for a regression that needed to include State agency and clinic specific variables.[2] When the sample size is small and the model contains several variables, overfitting can result. Overfitting is a modeling error that occurs when there is not enough input data available for the model to find patterns in it. Overfit models are not informative.

Table I-3. Tabulation of staff interview question 42, showing insufficient variation to proceed with regression analysis

Staff Interview Q42: Agree/disagree: The flow of participants through the clinic as they receive services works well	Number	Percentage
Strongly Agree	12	13.8
Agree	49	56.3
Undecided	10	11.5
Disagree	12	13.8
Strongly Disagree	4	4.6
Total	87	100

Note: Not asked of staff conducting remote sessions (n=89).
Source: Calculated from Appendix C, Table 4.

[1] Study analysts briefly considered combining the undecided cases with the strongly disagree/disagree cases. However, a response of undecided seemed to be a meaningfully different response than a disagreement with the statement. Respondents may have selected undecided because flow is not an issue at their clinic (which may be closer to agreeing with the statement), or they may have felt they did not have enough experience to form an opinion.

[2] See Sample Size Issues for Categorical Analyses and Logistic Regression, which indicates under 100 is "risky."