Validation of Measures of Food Insecurity and Hunger

Edward A. Frongillo, Jr.

Division of Nutritional Sciences, Cornell University, Ithaca, NY 14853–6301

ABSTRACT The most recent survey effort to determine the extent of food insecurity and hunger in the United States, the Food Security Supplement, included a series of questions to assess this complex phenomenon. The primary measure developed from this Food Security Supplement was based on measurement concepts, methods and items from two previously developed measures. This paper presents the evidence available that questionnaire-based measures, in particular the national food security measure, provide valid measurement of food insecurity and hunger for population and individual uses. The paper discusses basic ideas about measurement and criteria for establishing validity of measures and then uses these criteria to structure an examination of the research results available to establish the validity of food security measures. The results show that the construction of the national food security measure is well grounded in our understanding of food insecurity and hunger, its performance is consistent with that understanding, it is precise within usual performance standards, dependable, accurate at both group and individual levels within reasonable performance standards, and its accuracy is attributable to the well-grounded understanding. These results provide strong evidence that the Food Security Supplement provides valid measurement of food insecurity and hunger for population and individual uses. Further validation research is required for subgroups of the population, not yet studied for validation purposes, to establish validity for monitoring population changes in prevalence and to develop and validate robust and contextually sensitive measures in a variety of countries that reflect how people experience and think about food insecurity and hunger. J. Nutr. 129: 506S–509S, 1999.

KEY WORDS: • hunger • food insecurity • validity • measurement • humans

Food insecurity clearly demonstrates a range in severity. The least severe form that people experience is the uncertainty in obtaining food in socially acceptable ways; the most severe form occurs when people who simply do not get enough to eat as a result of insufficient resources experience the physical and psychological consequences of hunger. The most recent survey effort was conducted as part of the national nutrition monitoring system to determine the extent of food insecurity and hunger in the United States (U.S.). The survey included a series of questions to assess this complex phenomenon (Carlson et al. 1999, Hamilton et al. 1997, Rose et al. 1995). The primary measure developed from this Food Security Supplement was based on measurement concepts, methods and items from two previously developed measures from Radimer/Cornell (Radimer et al. 1990 and 1992) and the Community Childhood Hunger Identification Project (CCHIP; Wehler et al. 1992).

The purpose of this paper is to present the evidence available that questionnaire-based measures, in particular the national food security measure, provide valid measurement of food insecurity and hunger for population and individual uses. The paper first discusses basic ideas about measurement and criteria for establishing validity of measures. These criteria are then used to structure an examination of the research results available for establishing the validity of food security measures.
University. Because the phenomenon of food insecurity, as experienced by people, was not well understood, a rigorous naturalistic paradigm was chosen to understand, define and measure food insecurity based on the way in which people actually experience it (Radimer et al. 1990 and 1992).

Validation is the process of determining whether a method is suitable for providing useful analytical measurement for a given purpose and context. A method suitable for providing useful analytical measurement for a given purpose and context is one for which the following are true: 1) its construction is well-grounded in an understanding of the phenomenon; 2) its performance is consistent with that understanding; 3) it is precise within specified performance standards; 4) it is dependable within specified performance standards; 5) it is accurate within specified performance standards; and 6) its accuracy is attributable to the well-grounded understanding for that purpose and context. If all of these criteria are fulfilled, then the method is valid for that purpose and context (Habicht et al. 1979, Koch 1987).

Several possible purposes for populations and for individuals exist. Many of these are potentially relevant to a measure of food security. Some purposes for population include the following:

- Estimation of prevalence (How many people are affected?)
- Determination of causes (Why are people affected?)
- Targeting (Who is affected?)
- Monitoring (How is the situation changing?)
- Evaluation of programs (Who has benefitted and how?)

Some purposes for individuals include the following:

- Screening (Is the person at risk?)
- Diagnosis (Does the person have the problem?)
- Monitoring (Is the person's situation improving?)

The Food Security Supplement was specifically intended to be used to estimate prevalence and to monitor populations, but several other population and individual uses can be envisioned, including the identification of the food security status of households or individuals.

**VALIDITY OF FOOD SECURITY MEASURES**

Over the last 10 years, research has focused on the development and validation of questionnaire-based, direct measures of people's experience with food insecurity and hunger. The evidence for the validity of this approach is presented, organized according to the six validation criteria.

**Well-grounded construction.** The most important evidence that the construction of the food security measures is well-grounded in understanding is the formal, in-depth interviews of rural women with children conducted by Radimer et al. (1990 and 1992). An additional basis is the informal knowledge from contact with people who are food insecure, for example, through the CCHP surveys (Wehler et al. 1992). The in-depth interviews of Radimer clarified our understanding of the following facts: 1) food insecurity is experienced differently at the household, adult and child levels; 2) food insecurity has four components; and 3) families experience sequenced levels of severity of food insecurity with hunger as the most extreme consequence of the progression of food insecurity. Two components of food insecurity, quantity and quality of food, are related directly to food. Two components, certainty and acceptability, are psychological and social in nature. Furthermore, food insecurity is a managed process; in their management of food insecurity, families aim to protect children from food insecurity and its consequences. The Radimer/Cornell items were developed from words taken from the in-depth interviews of the women experiencing food insecurity, which contributes to their construct-validity. CCHIP items were developed from the knowledge of project investigators who were in contact with people experiencing food insecurity.

**Performance consistent with understanding.** The consistent manner in which the food security measures meet the well-grounded understanding has been demonstrated in four ways. First, factor analysis has shown that the conceptualized components of food insecurity and hunger are confirmed in empirical data (Hamilton et al. 1997, Kendall et al. 1995, Radimer et al. 1992). Second, the proportion of affirmative responses for items has been examined to determine if the conceptualized sequence of severity is reflected in the sequence of responses in items. Third, extensive cognitive testing of the measured items has been done (Alaimo 1997, Hamilton et al. 1997) to determine and ensure that items ask respondents a meaningful question that they can answer and interpret as the developers intended.

Fourth, the consistency of patterns of affirmative responses across populations has been examined. For example, Figure 1 shows the pattern of affirmative responses for five population surveys that had eight items in common with the Radimer/Cornell measure. Although the prevalence of the items differed markedly across the samples (due in part to differential oversampling of low income households), the patterns were parallel. The only exceptions are two items from the Hispanic Health Survey (Perez-Escamilla et al. 1997). This uncertainty item may have been answered affirmatively more often in the Hispanic Health Survey because of the general economic and social uncertainty in that community at the time of the survey (R. Perez-Escamilla, personal communication) or perhaps for other reasons. Similar examinations of the pattern of responses made for U.S. subpopulations (defined by race/ethnicity, income and household composition) by the national food security measure showed parallel curves, indicating very good consistency of responses.

**Precision and dependability.** Precision is the extent to which repeated measurements yield the same value. The usual...
internal consistency method has demonstrated very good precision (Cronbach's $\alpha \geq 0.85$) for the Radimer/Cornell (Kendall et al. 1995, Radimer et al. 1992), CCHIP (Wehler et al. 1992) and the U.S. food security measures (Hamilton et al. 1997). Dependability is the extent to which differences in a measure consistently reflect differences in the phenomena. For example, fluctuation in hydration is a source of undependability when body weight is used as a measure of nutritional status in infants. Given the relatively long time span of 12 mo relevant for the national food security measure, dependability is not an issue, but it might be for short time spans if, for example, transient events in people's lives influence assessment of their food security status.

Accuracy. Accuracy is the extent to which a measure provides unbiased assessment of the phenomena. Accuracy is achieved by construction, which rests upon the depth of understanding of the phenomena. Accuracy is assessed by in-depth analysis and by relating the measure to a criterion measure, which may be a more definitive measure, determinant or consequence.

The hierarchy of possible measures is the following: definitive, reference and routine (Uriano and Cali, 1977). Definitive measures achieve high accuracy because they rely on first principles, i.e., they reflect in a fundamental way the theoretical structure of the phenomena they purport to represent. Reference measures achieve accuracy because they directly and closely relate to the phenomena of interest, and accuracy is demonstrated by comparison to definitive measures. Routine measures are usually fast and inexpensive; they require relatively unsophisticated personnel, whereas accuracy is demonstrated by comparison to reference measures.

The Food Security Supplement was intended to provide a reference measure. The demonstration of accuracy of a reference measure is best done by comparison to a definitive measure. To develop a definitive measure, a consensus method was used with a rich set of information from a 1993 survey of 193 households with women and children living at home in a rural county. The aim was to define a criterion measure for food security to compare with food security items from Radimer/Cornell and CCHIP (Frongillo et al. 1997). The criterion measure that was developed approximated the most definitive measure possible, one that would be gained from an in-depth understanding of the experience itself through a personal interview with the respondent. Two researchers with very different experiences achieved good agreement after working independently, and excellent consensus after working together, in categorizing the households. The Radimer/Cornell and CCHIP measures had good specificity (i.e., percentage of truly food secure correctly classified, 63–71%) and excellent sensitivity (i.e., percentage of truly food insecure correctly classified, 84–89%) compared with the criterion measure. Estimates of the prevalence of household food insecurity from the criterion, Radimer/Cornell and CCHIP measures were almost identical. Similar results have been obtained for the Radimer/Cornell measure by Anne-Marie Hamelin in Québec (personal communication) and in a study of elderly urban Black and rural White women and men (Wolfe et al. 1998).

Additional evidence for the accuracy of questionnaire-based measures of food security has come from comparisons of the Radimer/Cornell or the national food security measure with a large number of determinants and consequences such as income, education, participation in food assistance programs, having savings, food expenditures and food consumption (Hamilton et al., 1997, Kendall et al. 1994, 1995 and 1996, Olson et al. 1994 and 1997). These studies have demonstrated that the food security status of groups of households is associated with these determinants and consequences in the expected manner.

Attribution of accuracy. It is crucial to establish whether the apparent accuracy of a measure is actually due to the well-grounded understanding. Otherwise, a measure that is supposedly validated may well actually be useless. For example, the apparent accuracy of a measure of lean body mass of small animals has been shown to be attributable to the measure's relation to total body weight and not to its relation to actual lean body mass as had been thought (Bell et al. 1994). This issue for the assessment of food security was addressed by asking whether the food security measure accounted for variability in the definitive criterion measure after accounting for other socioeconomic factors (Frongillo et al. 1997). Multinomial logistic regression was used to examine a sequence of models for the Radimer/Cornell measure. For example, income alone had a model fit ($-2$ log-likelihood) of 233 (6 df). When employment status was added, the fit improved to 201 (18 df). When the Radimer/Cornell measure was added to the other two variables in model, the fit was 182 (24 df), a substantial further improvement ($P < 0.0005$). Nearly identical results were found for the CCHIP measure. These results demonstrate that the food security scale measures food insecurity status beyond what income and employment (and other socioeconomic variables that were examined) can explain.

SUMMARY

These results provide strong evidence that the Food Security Supplement provides valid measurement of food insecurity and hunger for population and individual uses. The construction of the national food security measure is well grounded in our understanding of food insecurity and hunger, its performance is consistent with that understanding, it is precise within usual performance standards, dependable, accurate at both group and individual levels within reasonable performance standards, and its accuracy is attributable to the well-grounded understanding.

To assess accuracy for uses at the level of individuals, a rich set of information available from survey interviews was used to develop a definitive criterion measure. Both the Radimer/Cornell and CCHIP questionnaire-based measures had good specificity and excellent sensitivity compared with the criterion measure.

These results support the use of the items from the Food Security Supplement measure, which is based on the Radimer/Cornell and CCHIP measures, for the purposes of estimating the prevalence of food insecurity and hunger in the U.S. population. Furthermore, the excellent sensitivity of these measures means that they can be used validly to identify households for food insecurity and hunger, and to target portions of the population for food programs. Further validation research is required for subgroups of the population; this may be different from that in Upstate New York, which has been studied most intensively for validation purposes. Continued validation work should be done with the national food security measure to ensure its accuracy for monitoring population changes in prevalence and changes in households and individuals.

This approach of using questionnaire items to assess people's experience of food insecurity has the direct potential to be important in providing a common means for assessing food insecurity in other countries. Food insecurity is a social as well as biological, nutritional and economic phenomenon. This approach provides a way of capturing the social as well as other aspects of this phenomenon. Even the best economic measures
fail to capture important social and contextual variability. Research that uses qualitative and quantitative methods to fully understand how people in a variety of countries experience and think about food insecurity is required. Then, this understanding of the experience of food insecurity should be used to develop systematically robust and contextually sensitive measures that have demonstrated validity, building upon what has been done for the U.S. national food security measure.

ACKNOWLEDGMENT
The criteria for establishing validity of measures was developed with Kathleen M. P. Hanson.

LITERATURE CITED