Methods for estimating comparable prevalence rates of food insecurity experienced by adults in 147 countries and areas

Mark Nord, Carlo Cafiero and Sara Viviani
Statistics Division, Food and Agriculture Organization of the United Nations, Viale delle Terme di Caracalla 00153, Rome, Italy.

E-mail: mark.nord@fao.org

Abstract. Statistical methods based on item response theory are applied to experiential food insecurity survey data from 147 countries, areas, and territories to assess data quality and develop methods to estimate national prevalence rates of moderate and severe food insecurity at equal levels of severity across countries. Data were collected from nationally representative samples of 1,000 adults in each country. A Rasch-model-based scale was estimated for each country, and data were assessed for consistency with model assumptions. A global reference scale was calculated based on item parameters from all countries. Each country’s scale was adjusted to the global standard, allowing for up to 3 of the 8 scale items to be considered unique in that country if their deviance from the global standard exceeded a set tolerance. With very few exceptions, data from all countries were sufficiently consistent with model assumptions to constitute reasonably reliable measures of food insecurity and were adjustable to the global standard with fair confidence. National prevalence rates of moderate-or-severe food insecurity assessed over a 12-month recall period ranged from 3 percent to 92 percent. The correlations of national prevalence rates with national income, health, and well-being indicators provide external validation of the food security measure.

1. Introduction
Food security is said to exist when all people, at all times, have physical, social and economic access to sufficient safe and nutritious food that meets their dietary needs and food preferences for an active and healthy life [1]. Food security is inherently multi-dimensional, but one critical dimension is consistent economic access to adequate food. The United Nations Food and Agriculture Organization (FAO) has undertaken the Voices of the Hungry (VoH) project to develop and support a survey-based experiential measure of this dimension of food security. The methodology is similar to that used in the United States, Brazil, Canada, Mexico, and several other countries [2]. The project has taken on increased importance as the experiential-based measure has been endorsed by the United Nations Statistical Commission to monitor Target 2.1 of the recently agreed 2030 Agenda for Sustainable Development, which states: “By 2030, end
hunger and ensure access by all people, in particular the poor and people in vulnerable situations, including infants, to safe, nutritious and sufficient food all year round” [3].

The VoH methodology measures the severity of food insecurity (lack of economic access to adequate food), based on conditions and behaviours reported in response to an 8-question survey module, the Food Insecurity Experience Scale Survey Module (FIES-SM; table 1). Response data are assessed and combined to create a measure using statistical methods based on Item Response Theory, in which the severity of food insecurity experienced by an individual or household is modelled as a latent trait—a characteristic not directly observable. Observable conditions assumed to be caused by the latent trait and elicited in response to the FIES-SM questions provide the information for constructing the measure. Based on the measured severity of food insecurity of a nationally representative sample of individual adults or households, a national prevalence rate of food insecurity beyond a specified threshold of severity is calculated. Thresholds and, thus, prevalence rates of food insecurity are made comparable across countries using information on the severity each item from a statistical model of each country’s scale.

Table 1. English version of the Food Insecurity Experience Scale

| Num. | Short reference | Question wording |
|------|-----------------|------------------|
| 1    | WORRIED         | During the last 12 MONTHS, was there a time when you were worried you would not have enough food to eat because of a lack of money or other resources? |
| 2    | HEALTHY         | Still thinking about the last 12 MONTHS, was there a time when you were unable to eat healthy and nutritious food because of a lack of money or other resources? |
| 3    | FEWFOODS        | Was there a time when you ate only a few kinds of foods because of a lack of money or other resources? |
| 4    | SKIPPED         | Was there a time when you had to skip a meal because there was not enough money or other resources to get food? |
| 5    | ATELESS         | Still thinking about the last 12 MONTHS, was there a time when you ate less than you thought you should because of a lack of money or other resources? |
| 6    | RANOUT          | Was there a time when your household ran out of food because of a lack of money or other resources? |
| 7    | HUNGRY          | Was there a time when you were hungry but did not eat because there was not enough money or other resources for food? |
| 8    | WHOLEDAY        | During the last 12 MONTHS, was there a time when you went without eating for a whole day because of a lack of money or other resources? |

2. Data
FAO expects that national prevalence rates of food insecurity for monitoring progress toward the Sustainable Development Goal Target 2.1 will eventually be based on data from national surveys conducted by national statistical agencies in each country. However, to obtain initial baselines on a large number of countries and to provide data for developing methods to make prevalence rates across countries comparable, FAO
contracted with the Gallup Organization, as a data collection service provider. The 8-question FIES-SM was thereby added as a client module to the Gallup World Poll beginning in 2014. The Gallup World Poll has been conducted annually since 2006 in about 150 countries, interviewing nationally representative samples of 1,000 adults (ages 15 and older) in each country on a range of topics including family economics, employment, human development, and well-being [4]. In 2014, FIES data were collected in the Gallup World Poll in 147 countries, areas, and territories.

This paper reports on the psychometric assessment of the FIES data collected in the 2014 Gallup World Poll, the development of a methodology to make prevalence estimates comparable across countries, and initial estimated prevalence rates of food insecurity at two levels of severity.

3. Psychometric Assessment of FIES Data
The Item Response Theory measurement model used for FIES data assessment and scale construction is the single-parameter logistic measurement model commonly known as the Rasch Model [5]. The Rasch model assumes that the log-odds of respondent \( r \) saying “yes” to item \( i \) is a linear function of the difference between the severity of food insecurity experienced by \( r \) and the severity of item \( i \). The severity of an item is defined as the severity experienced by respondents who are equally likely to affirm or deny the item. Two further important assumptions of the model are: 1) that all items discriminate equally—that is, that the slopes of the curves of log-odds of affirmation against respondent severity are equal for all items; and 2) that items are conditionally independent—that is, that correlations between items are entirely due to their common association with the latent trait. Given these assumptions, and a few additional assumptions less likely to be problematic for these data, the raw score of a respondent (the number of affirmative responses to the set of items) is a sufficient statistic for, and an ordinal indicator of, the severity of food insecurity he or she has experienced.

Maximum likelihood methods are used to estimate the severity of each item and the severity experienced by each respondent. Model estimation also includes calculation of item and respondent fit statistics, conditional correlations across items, and measurement reliability of the scale. Item severity parameter estimation, item-fit statistics, respondent-fit statistics, and inter-item conditional correlations are based on conditional maximum likelihood (CML) estimation. That is, the likelihood of item affirmation is calculated conditional on raw score. Respondent severity parameters and measurement error (uncertainty) are calculated as the maximum likelihood values given the item severity parameters.

The Rasch Model assumption of equal discrimination of items was assessed primarily by the item infit statistic, an information-weighted chi-square-type statistic that compares observed with expected misfit of each item. The expected value of all infits is 1.0; higher values indicate weaker discrimination; and values between 0.7 and 1.3 were considered to adequately meet the assumption of equal discrimination. In spite of the wide range of cultures and languages in which the FIES-SM was administered and the attendant challenges of translation, the fit of all the items to the measurement model was remarkably good in almost all countries. In 93 percent of countries, infit statistics for every item were in the acceptable range. Only 7 countries had any item with an infit higher than 1.4, and with one exception, those were countries with small numbers of non-extreme cases and sizeable resulting estimation errors for item-infits. (Only non-extreme cases—those with raw scores of 1 to 7—are used to estimate the measurement model. In highly food-secure populations, a large majority of respondents say “no” to all 8 questions, and the estimation analysis sample is small.)

Conditional correlations were not found to be excessive for any pairs of items in countries with sufficient sample size of non-extreme cases to produce reliable assessments.

Rasch reliability is the proportion of total variance in the population that is accounted for by the measurement model. A modified Rasch reliability, in which calculations of error variance and variance accounted for by the model were weighted equally across raw scores, was used in this study. The standard
Rasch reliability statistic is affected both by model fit and by the distribution of severity of food insecurity in the sample. The equal-weighted reliability statistic avoids the confounding effect of differences in the distribution of severity of food insecurity across countries. Mean Rasch reliability across countries (using the modified statistic) was .740, with a range from .676 to .847. Reliability was between .70 and .80 for 79 percent of countries. These levels of reliability for a scale comprising just eight items reflect reasonably good model fit, and result in measurement errors in national prevalence estimates that are small compared with sampling errors.

4. Making the Measure Comparable across Countries

Measures were made comparable across countries by first calculating a global standard scale based on item severity parameters of almost all countries; second, specifying thresholds for moderate and severe food insecurity on the global scale, and third, adjusting each country’s scale to the global standard based on item severity parameters in order to transform the global thresholds to that country’s scale. These processes are based on the assumption that, with identifiable exceptions, the objective conditions elicited by responses to FIES questions are the same in all countries and languages. However, due to differences in linguistic nuances and differences in the ways in which food insecurity is experienced and managed in different cultures, one or more items in a country’s data collection may, in fact, not represent the same objective conditions as in the majority of countries. Therefore, in both the calculation of the global standard scale and the adjustment of each country’s scale to the global standard, an item or items in a country’s scale were considered unique to that country if they differed in severity from the global standard in excess of a specified tolerance (about 0.5 logits). Unique items were retained in the country’s scale, but were not considered equivalent to the corresponding items in other countries’ scales or in the global standard. Making these adjustments is complicated by the fact that Rasch-model based scales are undefined with respect to location (i.e., they have no natural zero point) and that dispersion of items is affected by the extent of memory and reporting error, which can differ from country to country.

The global standard scale was calculated based on item severity parameters of almost all countries. A few countries were omitted due to small estimation sample sizes or poorly fitting data. An iterative process of estimating the global standard item parameters allowed up to 3 items to be considered unique in each country. In the first iteration, each country’s scale was divided by the standard deviation of its item parameters to equate the dispersion of items in all countries; no items were considered to be unique, and the global standard parameter for each item was calculated as the median parameter for that item across countries. In the second iteration, each country’s scale was adjusted by a linear mapping to equate the mean and standard deviation of all items to those of the global standard. Items were specified as unique if their differential item function (DIF) from the global standard exceeded a set tolerance. In each succeeding iteration, each country’s scale was adjusted by a linear mapping to equate the mean and standard deviation of the non-unique items to the corresponding items in the global standard. Unique items were also omitted from the calculation of the median for that parameter—the global standard parameter to be used in the following iteration. Iteration continued until the DIF of all non-unique items in all countries was within the set tolerance. Countries in which more than three items were unique were removed from the set of countries used to calculate the global standard.

The threshold for moderate food insecurity—the level of severity beyond which respondents would be classified as having moderate or severe food insecurity—was specified at the severity of the item ATELESS on the global standard scale. The threshold for severe food insecurity was specified at the severity of the item WHOLEDAY on the global standard scale (see Table 1 for the full wording of these questions). These specifications give substantive meaning to the main statistics reported for each country. Individuals with “moderate or severe food insecurity” have typically eaten less than they thought they should at times during the year because they lacked sufficient resources for food, and most will have experienced more severe...
conditions. Those with severe food insecurity have typically gone whole days without eating at times due to lack of resources.

For country scales with three or fewer unique items (93 percent of countries), the linear mapping required to adjust the country scale to the global standard was already specified in the final step of the development of the global standard. By reversing that mapping, the thresholds for moderate and severe food insecurity on the global standard scale were back-transformed to the metric of each country. Adjustment of the remaining countries to the global standard—those with DIF of one or more items still out of tolerance after specifying the most discrepant three items as unique—was conducted on a case by case basis, and greater uncertainty about prevalence rates is noted in published results. For four countries, no defensible adjustment was found, and raw-score parameters as well as thresholds were based on the global standard scale.

Finally, two prevalence rates were calculated for each country: the percentage of the population with moderate-or-severe food insecurity and the proportion with severe food insecurity. These calculations are based on the assumption that within each raw score, the true severity of food insecurity is normally distributed with mean at the value of the respondent parameter for that raw score and standard deviation equal to the estimated measurement standard error (uncertainty) of severity associated with that raw score. The proportion of the population beyond a threshold was then calculated by summing across raw scores the proportion of the sample with that raw score multiplied by the proportion of the normal distribution of severity for that raw score falling beyond the threshold. Survey sampling weights were used for these calculations so that the prevalence rates are representative of the national population.

Special considerations had to be given to respondents with raw scores 0 and 8. The conditional maximum likelihood methods used to estimate the measurement model do not provide statistically rigorous severity parameters or measurement errors for these extreme scores. Respondents with raw-score 0 were classified as highly food secure. Although no severity parameter or measurement error was specified for those respondents, their classification as highly food secure implicitly assumes that the proportion with severity beyond the moderate threshold is negligible. For respondents with raw-score 8, a severity parameter was calculated based on a pseudo raw score that ranged from 7.5 to 7.7, depending on the ratio of respondents with raw-score 8 to respondents with raw scores 1 to 7. Measurement error (uncertainty) for these respondents was calculated based on a pseudo raw score of 7.5. These ad-hoc specifications were developed and assessed by analysis of additional, more severe items available for a few countries in the 2014 data and a larger number of countries in the 2015 data. Any imprecision due to imperfect specification of the distribution of severity in raw-score 8 affects only a relatively few countries that have a substantial proportion of cases with raw score 8.

5. National Prevalence Rates

National prevalence rates of moderate-or-severe food insecurity assessed over a 12-month recall period ranged across countries, areas, and territories from 3 percent to 92 percent. [6]. The prevalence of severe food insecurity (characterized by individuals not eating for a whole day at times during the year) ranged from below 0.5 percent to 76 percent.

Regional prevalence rates for FAO statistical regions were calculated as population-weighted means across countries, areas, and territories for which data are available (more than 90 percent of population in most regions). Food insecurity was most prevalent in Sub-Saharan Africa and least prevalent in East Asia and Developed Regions (figure 1).

Preliminary external validation was provided by associations of national prevalence rates of food insecurity with expected correlates such as the World Bank Poverty Rate (Pearson r = .76), Human Development Index (r = -.82), FAO Prevalence of Undernourishment (r = .76), and the World Health Organization Under-5 Mortality Rate (r = .83). To date, few of these associations have been explored in
detail, but the association with Under-5 Mortality Rate was found to remain substantial and statistically significant with controls for Poverty Rate and Prevalence of Undernutrition.

**Figure 1.** Prevalence rates of food insecurity by FAO Statistical Region, 2014

**Acknowledgements**

Financial support for the Voices of the Hungry project has been provided by the United Kingdom Department for International Development (DFID) and by the Kingdom of Belgium through the FAO Multipartner Programme Support Mechanism (FMM). The comments, opinions and judgments expressed in this paper are those of the authors and do not imply any official position by FAO, the FAO Statistics Division, or the funding partners.

**References**

[1] Declaration of the World Summit on Food Security, World Summit on Food Security, Rome, 16-18 November 2009

[2] Coleman-Jensen A, Rabbitt M, Gregory C and Singh A 2015 Household Food Security in the United States in 2014 (Washington, DC: U.S. Department of Agriculture, Economic Research Service, Economic Research Report No. ERR-194)

[3] United Nations 2015 Transforming Our World: the 2030 Agenda for Sustainable Development (United Nations, Resolution adopted by the General Assembly on 25 September 2015)

[4] Gallup World Poll. http://www.gallup.com/services/170945/world-poll.aspx

[5] Fischer G H and Molenaar I W 1995 Rasch Models: Foundations, Recent Developments, and Applications (New York: Verlag Springer)

[6] United Nations Food and Agriculture Organization 2016 Methods for Estimating Comparable Rates of Food Insecurity Experienced by Adults throughout the World. (Rome: FAO)