Assessment of the PROBIT approach for estimating the prevalence of acute malnutrition from population surveys

Summary of research

Location: N/A

What we know already: Prevalence of GAM is normally estimated using two stage cluster sampled surveys using the SMART method. The PROBIT method is an alternative method that estimates prevalence indirectly and requires a smaller sample size. GAM case definition may be based on weight-for-height z score (WHZ) and mid-upper arm circumference (MUAC) criteria.

What this article adds: The study confirms that the PROBIT method can estimate prevalence of GAM and MAM using MUAC or WHZ. The PROBIT method is more suited when estimating prevalence using small sample sizes (<150), e.g. sentinel-site surveillance systems. The classical method is preferred when estimating prevalence with larger samples.

The most commonly used method for estimating the prevalence of global acute malnutrition (GAM) is by conducting two stage cluster sampled surveys requiring the measurement of several hundreds of children, typically 900 children (thirty clusters of thirty children), for achieving a sufficient precision for decision making.

In 2006, the SMART (Standardised Monitoring and Assessment of Relief and Transitions) method was introduced, addressing the problems of lack of standardisation and lack of methodological rigour in the way nutritional surveys were undertaken. Despite the consistency in methodology and analysis that this method has provided for the nutrition community, there remains concern about the difficulty in obtaining usefully precise estimates of severe acute malnutrition (SAM) prevalence, the large sample size required within the constraints of security and accessibility to villages, and the cost in applications, such as surveillance by repeated cross-sectional surveys.

The classical method of estimating prevalence is to calculate the number of children meeting a case definition in the sample divided by the total number of children in the sample. In 1995, the WHO proposed the PROBIT method as an alternative method for prevalence estimation. The PROBIT method estimates prevalence indirectly using the inverse cumulative normal distribution function, which converts parameters of a normally distributed variable (i.e. the mean and standard deviation) to cumulative probability below any cut-off, which is equivalent to the proportion of individuals below the cut-off. The 1995 WHO document states that the main advantage of the PROBIT method is that it requires a smaller sample size than the classical method, however, no evidence of this is given.

The aim of a recent study was to compare the PROBIT method with the classical method for estimating prevalence of GAM, moderate acute malnutrition (MAM) and SAM using a computer based simulation approach to generate populations from real-world survey data sets and then simulate surveys sampled from these populations. Bias in the estimation of prevalence using the classical and PROBIT methods was investigated. The precision obtained for a given sample size when using the classical and PROBIT methods for estimating prevalence were compared. In addition to weight-for-height, WHO and UNICEF also recommend a SAM case definition based on mid-upper arm circumference (MUAC) of <115mm or the presence of bilateral pitting oedema. Several agencies also use MUAC = 115mm and <125mm as a MAM case definition for programmatic purposes. The study also tested the PROBIT and classical methods using these case definitions.
Method

The classical method and PROBIT method for calculating prevalence of acute malnutrition were compared using computer-based simulations. First, populations were created from a database of existing surveys. Then surveys describing these populations were simulated by sampling from populations created from these original survey data sets. The database used in the analysis consisted of 560 nutritional surveys involving children aged between 6 and 59 months from thirty-one different countries, totalling 459,036 children. The surveys were carried out by eleven different organizations involved in nutrition programmes throughout the world. The surveys included measurements of weight, height, MUAC and assessment of oedema. Weight-for-height z score (WHZ) was calculated using the WHO growth standards.

Each of the 560 surveys in the database was used to create a simulated population of 17,000 children by sampling with replacement from the survey dataset. This size of population was chosen as being typical of the populations in which nutritional anthropometry surveys are commonly performed. Sampling with replacement from the survey data sets was done to create simulated populations of the desired size. Each of the 560 simulated populations was sampled using simple random sampling without replacement. Fifteen different sample sizes (fifty, seventy-five, 100, 125, 150, 175, 200, 225, 250, 275, 300, 350, 400, 450 and 500) were used for these simulated surveys. One hundred and fifty surveys were simulated for each sample size from each population. This process led to a total of 1.26 million simulated surveys: 560 populations x 15 sample sizes x 150 simulations =1.26 million simulated surveys.

The study team calculated the true prevalence in each simulated population by counting the number of children meeting the case definition of SAM or MAM and calculating the ratio of this number to the total population.

First, the team calculated the prevalence using the classical method, by counting the number of children with the case definition of SAM or MAM in the simulated survey and calculating the ratio of this number to the total sample.

Second, the team estimated the prevalence with a PROBIT Method using three different PROBIT approaches. The first approach was based on the approach recommended by WHO that involved the sample median WHZ and assuming SD =1. The second approach used the sample mean and SD. The last approach involved the sample mean and SD calculated from data transformed towards normal, as explained below. Prevalence estimates for all three approaches using the PROBIT function were looked at as the cumulative probability of WHZ < -2 (GAM), WHZ < -3 (SAM) with and without oedema, and -3=WHZ <-2 (MAM). The approach using median and SD = 1 was not applicable to MUAC because the assumption that SD =1 is only applicable to WHZ which is assumed to follow the standard normal distribution. The two other approaches of the PROBIT function were used to calculate prevalence estimates as the cumulative probability of MUAC<125mm (GAM), MUAC<115mm (SAM) with and without oedema and 115mm=MUAC <125mm (MAM).

The team investigated the normality of distributions of anthropometric indices in the simulated surveys using the Shapiro–Wilk test. If there was evidence of non-normality (i.e. P< 0.05 for the Shapiro–Wilk test) then data were transformed towards normality using a power transformation with the transforming power found using the Box–Cox method. Bias was investigated for the PROBIT method by the estimation of mean error (true prevalence minus estimated prevalence). Precision was investigated by the 95% limits of agreement (mean (error) + or- 1.96 x sd (error)). For all methods, the half width of the 95% limits of agreement was calculated for different sample sizes based on the 150 simulated surveys. The analyses were also performed excluding children with oedema as it was suspected that oedema might bias WHZ upwards, leading to downwardly biased estimates of prevalence.

Findings
The analysis indicates that the methods using mean and SD of non-transformed and transformed data are similar, with the method using median and SD=1 inferior for both GAM and SAM but slightly better for MAM. Biases for the two PROBIT methods for GAM, MAM and SAM defined by MUAC again showed similarity between the methods using mean and SD of transformed and nontransformed data. The precision of the PROBIT methods (using the mean and SD of the survey with transformed and nontransformed data) is slightly better than for the classical method for sample sizes n<150 for GAM and SAM for both MUAC and WHZ. However, the precision of the PROBIT method (using the mean and SD of the survey with transformed and nontransformed data) is better for MAM for sample sizes n <300 for both MUAC and WHZ. The method using median and SD =1 is generally inferior to the classical method except for small sample sizes for MAM.

The main limitation of the study was that it was impractical to know the true prevalence of a large number of populations and to perform repeated surveys to estimate bias and precision of different estimators. The only feasible approach to testing the validity of the PROBIT approach was through simulation of surveys. The study confirms that the PROBIT method can estimate prevalence of GAM and MAM using WHZ or MUAC. The PROBIT method provides an estimate of prevalence that is proportional to the true prevalence with a small bias that can be corrected for by simple subtraction of a small value found of bias. The study shows, however, that the PROBIT method is inferior to the classical method for estimating the prevalence for SAM by WHZ or MUAC at sample sizes n >150, although it does seem suitable for small sample sizes which may be useful for applications such as surveillance. These results do not seem to be influenced by a bias resulting from the inclusion of cases with bilateral pitting oedema since the results are similar with or without oedema in the analysis. For WHZ, the PROBIT method of mean with observed SD of the data shows an improvement compared with using SD =1. This suggests that when choosing to use the PROBIT method, it would be useful to use the observed SD to calculate prevalence. Checking for normality and, if necessary, transforming data towards normality may further improve the estimation.

An explanation for the PROBIT method not estimating the prevalence of SAM as well as the classical method may be that perhaps the tail of the distribution of WHZ or MUAC does not follow the normal distribution and relates to children who may have other health problems in addition to primary malnutrition. One could argue that SAM children do not predictably follow the general pattern due to the fact that they are often infected or suffer from a family crisis which makes them shift in an unpredictable way.

In conclusion, the PROBIT method could be useful in sentinel-site surveillance systems using repeated small sample surveys or small spatially stratified samples so as to allow the course mapping of prevalence. The classical method should be preferred when estimating prevalence with larger samples.

1Nancy M Dale, Mark Myatt, Claudine Prudhon and André Briend (2012). Assessment of the PROBIT approach for estimating the prevalence of global, moderate and severe acute malnutrition from population surveys. Public Health Nutrition/FirstView Article/January 2006, pp 16. DOI: 10.1017/S1368980012003345, Published online: 27 July 2012 Link to this article: http://journals.cambridge.org/abstract_S1368980012003345

Taken from Field Exchange 45

www.ennonline.net/fex/45/assessment

© 2017. ENN is a registered charity in the UK no. 1115156, and a limited company no. 4889844.