# MONITORING SALT IODINATION IN DEVELOPING COUNTRIES

**Roberto A. Molina-Cruz, Universidad del Valle de Guatemala (UVG); David J. Fitch, Instituto de Nutrición de Centro América y Panamá and UVG; James M. Lepkowski, University of Michigan.**
**Roberto A. Molina-Cruz, Apartado Postal No. 82. 01901, Guatemala, Guatemala. rmolina@uvg.edu.gt**

**Key Words: Pooled Samples, Weighting, UNICEF.**

## Introduction

In Guatemala, similar to other countries, there are laws requiring the fortification of some foods. To keep a control on the fortification of salt with iodine - among other things - a program was established in 1994 by the Guatemala's Ministry of Education, with the assistance of the United Nations Children's Fund (UNICEF) and the Institute of Nutrition of Central America and Panama (INCAP). The Program was called Programa de Escuelas Centinelas de Micronutrientes, and it has been surveying students and analyzing salt samples brought from their homes every year since then. One of the Program's principal concerns is the amount of iodine that people - specially children - consume with salt. Its objective with the school surveys is to estimate the distribution of iodine in salt for the nation as a whole, and for its 22 departamentos.

At its installation the Program selected using non-probability methods about 420 schools from the total country's public schools. The Program has been working with the same set of schools, from which 20 students are selected randomly and asked for a small sample of salt - approximately 5 grams - from their homes. Because the chemical analysis are expensive and requires larger amounts of salt, the 20 salt samples are combined into 2 pooled samples of 10 samples each. The number of parts of iodine per million parts of salt (ppm) in each pooled sample is determined. The Program has been reporting these for the nation and each departamento (UNICEF 1995, 1999 and 2000).

In this paper we address three fundamental problems with the Program's school surveys, besides the non-probability sampling of the schools. 1) It is very likely that the reported results are hiding a main issue: much salt contains no iodine. 2) Because the estimates do not take into account the number of people living in the students' home, and of the students' siblings - children living in the same home - attending the same school, the current estimates are biased. 3) By getting salt samples from only homes with children attending public schools, the surveys are using an incomplete frame - especially because of low school attendance in Guatemala.

For the first problem, we look at the published results of previous years, and discuss the estimation of the salt samples' iodine content from the pooled samples data. For the second problem, we outline a proper sampling procedure, and the computation of the mean iodine's ppm that the people consume with salt and its variance. For the last problem, we seek insight on its seriousness from a recent income expenditure survey, and discuss briefly the possibilities of getting national estimates.

Although we don't discuss sequential sampling in this paper it is clear that as the school surveys are carried out every year, a general survey procedure based on sequential sampling should be used.

We want to thank Dr. Charles Proctor for his comments and suggestions during and after the meeting, which we appreciate very much. He pointed out that the pooled sample data may be analyzed using group testing and latent classes, but that the Program could use much simpler analysis techniques and that it may be enough just one pooled sample per school to indicate where the salt is not been adequately iodinated.

## Working with pooled samples

Law requires all salt for human and animal consumption or industrial use be fortified with a minimum of 30 to a maximum of 100 ppm of iodine (Congreso de la República de Guatemala, 1995). Since the regulations have not been strongly enforced, there almost certainly are producers or importers dealing unfortified salt. For instance, the 1999 results shows iodine contents of 477 pooled samples with a minimum 0.2 ppm to a 2nd largest 57.7 ppm (maximum 90.4 ppm), mean 15.7 ppm and median 13.8 ppm. Of which 267 (56.0%) contains 15.0 ppm or less, and 436 (91.4%) contains 30.0 ppm or less. These pooled samples corresponds to 276 schools, 201 (72.8%) schools with two pooled samples - say with $y^{(1)}$ and $y^{(2)}$ ppm of iodine, mean $\bar{y} = \frac{1}{2}\left(y^{(1)} + y^{(2)}\right)$ and variance $\frac{1}{2} \cdot \left(y^{(1)} - y^{(2)}\right)^2$ - and 75 (27.2%) schools with just one pooled sample - say with $\bar{y} = y^{(1)}$ and no variance. The school means range from a minimum 0.3

ppm to a maximum 57.9 ppm, with mean 15.8 ppm and median 14.7 ppm. Of these school means, 255 (92.4%) are 30.0 ppm or less and 142 (51.4%) are 15.0 ppm or less.

Since most school communities may have only one or two sources of salt, we may consider the people of these communities to be partitioned into two groups: those consuming enriched salt and those consuming non-enriched salt. The first group of people may be distributed with some mean $\mu$ and variance $\sigma^2$ - with respect the amount of iodine they consume with salt - and the second group may represent a fraction $\lambda$ of the community. Hence the distribution of the people of a school community may be considered having a general mean $(1-\lambda)\cdot\mu$ and variance $\mu^2\cdot\lambda(1-\lambda)+(1-\lambda)\cdot\sigma^2$ - by conditional expectations - which may be estimated using the school's pooled samples results. It is conceivable to assume that the people of the community consuming fortified salt are normally - truncated at zero - distributed. So, the proportion of people that consume specific interval amounts of iodine with salt may be estimated through the normal distribution with parameters $\hat{\mu}$ and $\hat{\sigma}^2$.

Because the Programs' current pooling procedure combines an equal amount of salt of all samples, the iodine's ppm of each school pooled samples and their average are estimates of the students' - not the school community's - general mean, and $s^2=\frac{10}{2}\cdot\left(y^{(1)}-y^{(2)}\right)^2$ may be used to estimate their general variance. If we assume that all producers and importers that are dealing fortified salt are just trying to meet on average the minimum required, we may assume that $\mu=30$. Then the differences between schools may be only because of different proportions of unfortified salt consumption, that is different $\lambda$s, and the general mean estimate $\bar{y}$ may be used to estimate $\lambda$: $\hat{\lambda}=1-\frac{\bar{y}}{30}$. Using the 1999 $\bar{y}$ values we can compute - under these assumptions - estimates $\hat{\lambda}$ for the 255 (92.4%) schools with $\bar{y}\le 30$, and for the 142 (51.4%) schools with $\bar{y}\le 15$ we have estimates $\frac{1}{2}\le\hat{\lambda}$. That is, half or more of the students in these schools may be getting unfortified salt. If these producers and importers are actually trying to meet the minimum required by more than 50% of their commercial bulk, say for example $\mu=45$, then from the same 1999 data we can compute estimates $\hat{\lambda}$ for 274 (99.3%) schools with $\bar{y}\le 45$, and the same 142 (51.4%) schools with $\bar{y}\le 15$ have estimates $\frac{2}{3}\le\hat{\lambda}$.

If we assume again that $\mu=30$, $\sigma^2$ may be estimated by $\hat{\sigma}^2=\frac{s^2}{1-\hat{\lambda}}-30^2\cdot\hat{\lambda}$. With the 1999 data we can compute estimates $\hat{\sigma}^2$ for 83 (30.0% of the total 276) schools with two pooled samples, $\bar{y}\le 30$ and $\frac{\hat{\lambda}}{1-\hat{\lambda}}\le\frac{s^2}{\bar{y}^2}$. For these schools we can estimate the distribution of students - assuming it normal - with respect their consumption of iodine with salt. In the table below we show the estimated distribution of two schools, an average distribution of the students of the 83 schools - with no weights because the Program does not keep track of the schools' sizes - and the pooled samples' distribution - which is the one reported by UNICEF (UNICEF, 1999). The reported results don't give an explicit estimate of the percentage of students that are consuming salt without iodine, and may be underestimating the amount of iodine that the students consuming fortified salt are getting.

| iodine's ppm: | 0 | (0,15] | (15,30] | (30,60] | (60,∞) |
|---|---|---|---|---|---|
| School A: | 51.83 | 5.84 | 17.93 | 23.77 | 0.63 |
| School B: | 12.50 | 14.80 | 17.41 | 32.21 | 23.08 |
| Average: (on 83 schools) | 49.71 | 6.69 | 13.54 | 20.23 | 9.83 |
| Reported: | 0.00 | 55.97 | 35.43 | 8.60 | 0.00 |

If only the school's pooled samples are analyzed, we can consider the following methods to estimate the three parameters of each school community.

Method 1, and maybe the most natural: First test the sample from each student for any iodine content to estimate $\lambda$, then form pools with only the samples that tested positive. We can get estimates of $\mu$ and $\sigma^2$ by analyzing the pooled samples. There are some low cost iodine content tests but these are of questionable reliability. We want to point out the need of research on the design of an adequate method to discriminate salt with iodine from salt without.

Method 2: Because there are only a few ways to mix iodine with salt, it may be possible to estimate a characteristic $\sigma^2$ by analyzing the making of salt. By forming pools with all students' salt samples and analyzing the pooled samples, we can get estimates $\bar{y}$ and $s^2$ of the communities general mean and variance parameters. And we can get estimates of $\lambda$ and $\mu$ by solving the following system of equations.

$$\bar{y} = (1-\lambda)\cdot\mu$$
$$s^2 = \mu^2\cdot\lambda(1-\lambda) + (1-\lambda)\cdot\hat{\sigma}^2 .$$

**A Sampling Procedure**

It would be desirable to sample first schools and from the selected schools to sample an equal number of students' homes. The selection of the schools should be done with probabilities proportional to their sizes -number of students. The students' homes should be selected trough the selection of students in the school, with equal probability, and taking into account possible absentees and the selection of siblings.

If the salt samples from each selected home were to be analyzed, the mean iodine's ppm that the people of the students' homes of one school consume with salt may be estimated using the common ratio estimator $\bar{y}_r = \dfrac{\frac{1}{m}\sum_{j=1}^{m}\frac{P_j}{M_j}y_j}{\frac{1}{m}\sum_{i=1}^{n}\frac{P_j}{M_j}}$. Where $m$ = number of students selected in the school, $M_j$ = number of children of home $j$ in the school the day of the survey, $P_j$ = number people of home $j$, and $y_j$ = iodine's ppm of home's $j$ salt. The usual variance estimator of $\bar{y}_r$ is

$$v(\bar{y}_r) = \frac{1}{m\cdot(m-1)\cdot\hat{P}^2}\sum_{j=1}^{m}\left(\frac{P_j}{M_j}y_j - \bar{y}_r\frac{P_j}{M_j}\right)^2 ,$$

where $\hat{P} = \frac{1}{m}\sum_{j=1}^{m}\frac{P_j}{M_j}$.

If instead the salt samples are combined into two or more pooled samples - say $n$ pooled samples with $m_i$ = number of individual samples in pool $i$ and $m = \sum_{i=1}^{n}m_i$ - then we can obtain the same $\bar{y}_r$ estimate using the pooled samples data by applying the following pooling procedure. Take from each individual sample a unit amount $z$ of salt per each member of the student's home, so for home $j$ of pool $i$ we have an amount $z\cdot P_{ij}$ of salt, and combine only the fraction $\frac{z\cdot P_{ij}}{M_{ij}}$ which contains $y_{ij}$ - unknown - ppm of iodine. Pooled sample $i$ iodine's ppm will then be $y_i = \dfrac{\sum_{j=1}^{m_i}\frac{z\cdot P_{ij}}{M_{ij}}y_{ij}}{\sum_{j=1}^{m_i}\frac{z\cdot P_{ij}}{M_{ij}}} = \frac{1}{P_i}\sum_{j=1}^{m_i}\frac{P_{ij}}{M_{ij}}y_{ij}$

where $P_i = \sum_{j=1}^{m_i}\frac{P_{ij}}{M_{ij}}$. Therefore

$$\bar{y}_r = \frac{\frac{1}{m}\sum_{i=1}^{n}\sum_{j=1}^{m_i}\frac{P_{ij}}{M_{ij}}y_{ij}}{\frac{1}{m}\sum_{i=1}^{n}\sum_{j=1}^{m_i}\frac{P_{ij}}{M_{ij}}} = \frac{\frac{1}{n}\sum_{i=1}^{n}\frac{n}{m}P_i\cdot y_i}{\frac{1}{n}\sum_{i=1}^{n}\frac{n}{m}P_i} .$$

The same $v(\bar{y}_r)$ estimate can't be computed knowing just the pooled samples iodine's ppm $y_i$, so we look at

$$v'(\bar{y}_r) = \frac{1}{n \cdot (n-1) \cdot \hat{P}^2} \sum_{i=1}^{n} \left( \frac{n}{m} P_i \cdot y_i - \bar{y}_r \frac{n}{m} P_i \right)^2 .$$

Which is the variance estimate of $\bar{y}_r$, but assuming that the students of each school are grouped into groups with $m/n$ students each, and then $n$ groups are selected with equal probability. To evaluate this estimator we simulated the iodine's ppm of the students' homes, the number of people and siblings, and the formation of 2 pooled samples with 10 individual samples each. A run of 1,000 iterations gives for example simulated values of $v(\bar{y}_r)$ and $v'(\bar{y}_r)$ with the following means and variances.

| | Mean | Variance |
|---|---|---|
| $v(\bar{y}_r)$ | 5.44 | 7.34 |
| $v'(\bar{y}_r)$ | 5.75 | 64.46 |

If all the school's salt samples are combined into pooled samples, $\bar{y}_r$ estimates the general mean of the distribution of the people in the community with respect the amount of iodine they consume with salt. If only the salt samples that contains some iodine are combined, $\bar{y}_r$ estimates the mean amount of iodine that people consuming fortified salt are getting with salt -that is $\mu$. To estimate the variance $s_y^2$ of either distribution is not a trivial task. In srs $v(\bar{y}) = \frac{s_y^2}{n}$, and so an estimate of the variance of the people distribution is $s_y^2 = n \cdot v(\bar{y})$. In our case, the mean is a ratio mean $\bar{y}_r$, so the estimated variance of this ratio mean $v(\bar{y}_r)$ is composed of the variances $s_y^2$ and $s_p^2$, and the covariance $s_{yp}^2$ (Fitch, 2001). For practical purposes we propose to estimate $s_y^2$ by a multiple of the estimated variance of the ratio mean $v'(\bar{y}_r)$ that may be $\frac{m}{n} v'(\bar{y}_r)$, where again $\frac{m}{n}$ represents the number of students in each pool.

Above we have specified a method of getting estimates for each school community, through for example $2$ two pools of size $10$. But the bias of estimated variances of ratio means when base on less that 10 PSUs can be a problem. As an approach which is minimally subject to this bias problem we propose the following. The approximately 400 schools should be selected from 20 subregions of the country, 20 schools in each subregion. Ten strata should be formed and 2 schools (PSUs) per stratum selected. From each school just one pool of size $m$ will be formed. Again the same $v(\bar{y}_r)$ will be made but this time with a difference. The $s_y^2$, $s_p^2$ and $s_{yp}^2$ statistics will be computed for each of the 10 strata, summed over the 10 strata, and one $v(\bar{y}_r)$ for the subregion computed. This is the standard method, such as is used by PC CARP. The variance estimate for each subregion will be $m \cdot v(\bar{y}_r)$. Assuming a normal distribution but this time in the subregion, the proportions of the people consuming salt in various iodine categories will be estimated. From such estimates, estimates will be made for each of the 8 regions of Guatemala and for the country as a whole.

**The Study Population**

Since severe iodine deficiencies causes goiter and even mild deficiencies will have health consequences to people of all ages, the Program is interested in getting estimates of all the people in the country - as the reports suggest. However the Program's study population includes only the people living in homes with at least one child attending a primary public school.

The most recent income and expenditure national survey (ENIGFAM 98-99) (INE, 1999a) measured some population characteristics including Edad (VDP05), age of the individual; Nivel de Escolaridad (VDP10), last school grade completed for a person with 7 or more years of age; and Asistencia Escolar (VDP12), which indicates if a person with 7 or more years of age attends school, and in that case also indicates if the school is public or private.

Using these ENIGFAM 98-99 data we estimate the total number of homes in the country as 1,997,537, of which 36.6% have a child in a primary public school. So the estimation procedures we are developing here apply to people living in only 36.6% of the households of Guatemala, and then only when we are able to make a probability selection of schools. People, with regard to their salt, living in households with school age children not in a public school or without a school age child, will differ in unknown ways.

The estimated percent of homes with a primary school age child, i.e. 7 to 14 years old, is 59.1%. Of these homes, 60.4% - that is 35.7% of the total - have at least one child attending a primary public school. The rest (39.6%) homes with no child attending a primary public school are divided into two groups: 35.6% (8.3% of the total) which have at least one child attending a private school, and so may be better off economically, and 64.4% (15.1% of the total) with children just not attending any school, and so the people in these homes - we can speculate - are likely getting unfortified salt.

To solve the problem of the - incomplete - frame of the school surveys, we see several possibilities for the use of multiple frame methodology and regression type models. In the present time INCAP is also analyzing samples of the salt that is been commercialized with a legitimate trade mark. We know that UNICEF and INCAP are preparing a national survey of stores which may be also carried out regularly. Demographic and Health Surveys are now carried in Guatemala every four years. The most recent (ENSMI 98-99) (INE, 1999b) included the measurement of some household characteristics like Type of salt used for cooking (hv222), with possible values: No salt available, Local salt, Pkgd salt with brand, Pkgd salt no brand, Salt for animals, Other and Missing. Salt contains iodine (sh24), with values: No, Yes, Other and Missing. Result of salt test (sh24a ), with values: Negative (white), Positive (purple) and Missing. And Iodized percent(sh24b), with values: 0%, 25%, 50%, 75%, 100%, +100% and Missing.

**References**

Congreso de la República de Guatemala (1995). Ley General de Enriquecimiento de Alimentos y sus Reglamentos, Decreto 44-92. Guatemala.

Fitch, David J. (2001). Some Sampling Theory for Work in Developing Countries. Manuscript. Guatemala.

Instituto Nacional de Estadística (INE) (1999a). Encuesta Nacional de Ingresos y Gastos Familiares 1998-1999. Guatemala.

— (1999b). Encuesta Nacional de Salud Materno Infantil 1998-1999. Guatemala.

UNICEF and Ministerio de Educación (1995). Escuelas Centinelas Micronutrientes. UNICEF, Guatemala.

— (1999). Escuelas Centinelas Micronutrientes. UNICEF, Guatemala.

UNICEF, Comisión Nacional para la Fortificación, Enriquecimiento y/o Equiparación de Alimentos, INCAP, y Organización Panamericana de la Salud (2000). Situación de los Alimentos Fortificados. Guatemala.