Monitoring and evaluation in flour fortification programs: design and implementation considerations

Juan Pablo Pena-Rosas, Ibrahim Parvanta, Frits van der Haar, and Thomas J Chapel

Designing and implementing effective monitoring and evaluation (M&E) is an integral element of wheat flour fortification programs. This review provides practical guidance for designing a M&E system for a flour fortification program. The Centers for Disease Control and Prevention’s Framework for Program Evaluation in Public Health has been adapted to identify key issues in the development of an integrated M&E system. A clear understanding of 1) the stakeholders in flour fortification and their needs, 2) the description and context of the fortification program, 3) the country’s wheat flour and flour products market, and 4) the resources available for the M&E component are critical and should be considered early in a program’s design.

© 2008 International Life Sciences Institute

INTRODUCTION

Increasing emphasis is being placed on designing and implementing effective monitoring and evaluation (M&E) as an integral element of any nutritional intervention. The M&E of flour fortification or other food fortification programs may serve some or all of the following important purposes: 1) to generate a shared understanding of the program among the stakeholders, 2) to document the linkages between the activities of the different stakeholders and overall program operations (successes and weaknesses) and outcomes, 3) to assess the availability and accessibility of the fortified food or foods across the country, 4) to assess the impact (intended or unintended) of the fortification program on the nutritional and health status of the target population over time, 5) to demonstrate accountability for resources expended toward the efforts, 6) to educate stakeholders about realistic expectations for change, and 7) to provide a platform for discussion, exchange of ideas, and decision-making among diverse stakeholders.

Given its many potential purposes, it is important to understand that M&E is situation-specific and varies with the intended purpose, the end user of the findings, and the intended use of the information. Each of these purposes may entail different users and uses of the findings, may measure different aspects of the program, and may vary in study design, data sources, and rigor. Hence, there is no single best M&E system for a flour fortification program.

Further complicating matters are the varied ways in which the core terminology is used in the literature and among nutrition program M&E specialists. Recently, an expert committee meeting of representatives from several international agencies working in food fortification programs agreed on the necessity for some operational definitions of program monitoring as compared with program evaluation.1

Affiliations: JP Pena-Rosas is with the International Micronutrient Malnutrition Prevention and Control Program, Division of Nutrition and Physical Activity, Centers for Disease Control and Prevention, Atlanta, Georgia, USA. I Parvanta is with the Infants and Toddlers Goal Team at the Coordinating Center for Health Promotion at the Centers for Disease Control and Prevention, Atlanta, Georgia, USA. F van der Haar is Visiting Associate Professor of Global Health at the Rollins School of Public Health, Emory University, Atlanta, Georgia, USA. TJ Chapel is with the Office of Workforce and Career Development, Centers for Disease Control and Prevention, Atlanta, Georgia, USA.

Correspondence: JP Pena-Rosas, International Micronutrient Malnutrition Prevention and Control Program, Division of Nutrition and Physical Activity, Centers for Disease Control and Prevention, 4770 Buford Highway NE, MS K-25, Atlanta, GA 30341-3724, USA. Phone: +1-770-488-5183, Fax: +1-770-488-5369, E-mail: jpenarosas@cdc.gov

Key words: evaluation, flour fortification, monitoring, sentinel monitoring
Program monitoring was defined as the continuous, ongoing collection, review, analysis, and use of information on program inputs, implemented activities, outputs, and outcomes, to assess how the program is performing against predefined criteria. A main aim of monitoring is to identify problems, such as inadequate expenditures, insufficient training, noncompliance, and similar factors, so that corrective actions can be taken to improve program performance. Thus, the purpose of monitoring is to reveal what is happening in the programs, and monitoring should be conducted often enough that the problems can be identified and addressed on a timely basis. By contrast, program evaluation is the systematic and objective assessment of the intervention to determine the relevance and fulfillment of program objectives, performance quality, outcome achievements, cost-effectiveness, and sustainability. In this context, evaluation helps to answer why certain things are or are not happening in a program. Program evaluation is most concerned with providing evidence for policy makers and program managers to make decisions regarding continuation, modification, expansion, or even termination of a program.

As defined, monitoring and evaluation are complementary processes. Typically, program evaluation is done less frequently than the ongoing and regular data collection for monitoring; indeed, it often targets problems that are identified through routine program monitoring. Given the different approaches that can be taken when designing monitoring or evaluation plans, we will refer to any flour or food fortification program assessment in the remainder of this article as M&E, with the understanding that each has a place in the program assessment.

Frontline programs should not always be held responsible for both the monitoring and the evaluation of flour fortification programs. Few grantees (i.e. some countries funded for implementation of strengthening of a fortification program or project) have the capacity or expertise to undertake rigorous program evaluation, especially of downstream outcomes. Therefore, some would argue that the M&E of flour fortification programs should be limited to the monitoring of inputs, activities, outputs, and trends in primary nutritional and health status indicators. Program evaluation, if done at all, could be limited to a process evaluation of program implementation (which is sometimes called “implementation fidelity”) – i.e., was the program implemented as intended, and what factors are related to strong or weak program implementation? Others, including donor agencies, however, view evaluation methods more broadly and see grantees and implementers of food fortification projects and programs as ready, willing, and able to undertake management-oriented evaluation of short-term, mid-term, and even long-term outcomes.

Furthermore, they see the findings of these evaluations as crucial to establishing a sense of continuous quality improvement – not only what is working or not working but why it is working or not working and what can be done about it. Here, we provide practical guidance on designing a M&E system for a flour fortification program that will support improving program performance and reaching the desired health and nutritional goals while assuring the availability of a safe, high-quality, fortified product.

**STEPS IN DESIGNING A FLOUR FORTIFICATION M&E SYSTEM**

The World Health Organization (WHO) M&E framework for food fortification programs identifies two components of food fortification programs that require different assessment approaches: 1) food control and regulatory monitoring to assess the supply of adequately fortified foods (nationally produced or imported), and 2) program-based M&E to track the population’s access to and use of fortified foods and the impact of the intervention on the health and nutrition of the population (Figure 1).

When planning a M&E system for a flour fortification program, it is important to understand some key aspects of the relevant food industry and the population’s consumption habits of flour and flour products to determine how best to assess and track the implementation and impact of fortification. For example, the approaches to determining whether fortified flour or flour products are readily available in the market in a country with a few flour mills (e.g., Kuwait, which has one mill, or Indonesia, which has four milling companies in five cities) will differ from those in a country with hundreds or thousands of roller mills (e.g., China). The objective in planning an M&E system is to design the least complicated, most efficient, and most sustainable data collection and reporting system possible.

In designing a M&E framework for flour fortification that would be responsive to these two components identified by WHO, we used the Centers for Disease Control and Prevention’s (CDC) Framework for Program Evaluation in Public Health to identify key issues (Figure 2). The CDC Framework involves six steps:

- Step 1: engage stakeholders;
- Step 2: describe the program;
- Step 3: focus the M&E design;
- Step 4: collect credible data;
- Step 5: justify the conclusions; and
- Step 6: ensure that the lessons learned are shared and used.
Following these steps will help to ensure that the time and effort invested in M&E pays off in findings that are relevant and likely to be used.

Each of the six steps in the M&E framework are further guided by standards that are organized into four groups: 1) utility – to determine at each step who will use the information and for what purpose; 2) feasibility – to determine constraints at each step, including technical, logistical, financial, and other available resource constraints; 3) propriety – to ensure that the M&E system is carried out legally, ethically, and with due regard for the welfare of those involved in implementing the system, as well as those affected by the results; and 4) accuracy – to ensure that the indicators, methods, and sources used will be accurate enough given the intended use of the information.

The data collected and the resulting information should adequately assess changes in the population coverage (percentage of the population that is regularly consuming sufficient quantities of fortified flour and flour products) and the effectiveness of the flour fortification program on the basis of improvements in the population’s nutrient status. In national flour fortification programs it is also important that the M&E system can detect any dramatic changes in dietary patterns of the population and that their overall micronutrient intake is within recommended limits. It is also critical that the flour is adequately fortified using good manufacturing practices. Below, we present in more detail the application of these steps for creating a strong M&E system for use in a flour fortification program.

**Step 1: Engage stakeholders**

Stakeholders are people or organizations who are invested in the program, are interested in the results of M&E, and have a stake (or tangible interest) in what occurs with the results. Stakeholders in a national flour fortification program might include the following: ministries of health, education, commerce, and industry along
with their relevant local branches; local governments and national and local legislators or judiciary; professional and scientific groups; civil society organizations; scientific and academic groups; food industries (millers, bakers, pasta and confectionary makers, and others); fortificant suppliers; media; and the United Nations and relevant donor agencies. Engagement of the public sector, health professionals, and civil society provide legitimacy to the fortification program; the industry and business sectors deliver the needed vitamins and minerals to the population by producing and distributing the fortified flour and flour products.5

For a multifaceted effort like food fortification, the stakeholders must cooperate for the program to be sustained. The coordinators of flour fortification programs depend on the stakeholders and partners to ensure program quality and efficiency, to ensure that political commitment is regularly renewed and invigorated, to ensure that communications are persistent over time, to ensure that the most effective channels are used to influence key decision makers, and to ensure that resources are available for training and other support activities, such as inspection and laboratory-based measurements.

Note that many of the stakeholders might have been engaged already in the planning and implementation of the program; here, we are extending their involvement to include the M&E of that program. The role that the stakeholders will play in M&E will vary. Some may or may not play a day-to-day role in the implementation of the M&E, but should be consulted to ensure consensus about the activities and outcomes that should be included in the M&E system. Their advice and concurrence are important because they may be the ones who fund the M&E component, who can advocate for its expansion and sustainability, or who may need to change their activities on the basis of the M&E. Other stakeholders may play direct roles in M&E, for example, by serving as members of a technical M&E committee or consulting on the development of tools for data collection, analysis, and interpretation. Involving the stakeholders from the start improves the likelihood that they will support the monitoring system and accept and use the resulting information. Stakeholder analysis allows the managers of the M&E system to identify the interests of different groups and to find ways to leverage the support of some while managing the risks that are posed by others.6

Step 2: Describe the program

The following is required: clarity and consensus on the need for the program, its components, its expected effects, the proposed activities, the resources available, and the context in which it will be implemented. Logic models are one way to describe the main elements of a program and how they work together to reach a specific goal.7 Logic models are useful for connecting the planning of a flour fortification program with its M&E component. Logic models frame M&E by helping to communicate the fundamental purpose of the program, helping to serve as a basis for determining whether the planned activities will lead to the expected impact, helping to identify potential operational and logistical problems, helping to help solve such problems, and helping to improve everyone’s understanding of the “big picture”. An overall visual representation of a flour fortification program is shown as a logic model in Figure 3. An additional component in the program may be the industry’s role in flour marketing and strategies.

In addition to displaying the inputs, activities, and outcomes of the program, the program description step also considers the context in which the program is to be implemented. Taken together, this information helps to predict whether and how the activities in a specific country can be implemented and to project the likelihood that implementing the activities will lead to the intended outcomes. Several contextual factors need to be assessed for flour fortification programs. First is the prevalence of vitamin and mineral deficiencies in the area (regional and national data as available) and other current and future nutritional interventions, including their goals and objectives, expected start and end (if any) dates, the target population outcome indicators, M&E activities, and the partners involved. If at all possible, the dietary patterns of the target population should be assessed, including regular, if any, consumption of other fortified foods. An analysis of the flour industry is needed. This should include assessing the characteristics of the flour milling industry and its production capacity (industrial vs. small scale); factors that may affect flour demand, such as price elasticity and various uses of the flour (commercially produced vs. homemade flour products); and the scope of the program as it is expected to affect the market. Finally, but equally important is whether the fortification program is to be mandatory or voluntary, and the regulatory requirements of the program including the level and type of fortificant to be used and the final product quality specifications.

Specific contextual questions to consider when designing the M&E system in flour fortification programs include the following:

a) What proportion of the total flour or flour product supply is expected to be fortified? If <100%, will that proportion of flour be fortified and marketed throughout the year (i.e., will consumers have access to the fortified flour and flour products regularly for sufficient periods to expect a meaningful change in nutrient status)? Is the amount of fortified flour or flour products sufficient to regularly
Figure 3  Logic model of a flour fortification program aimed at improving iron and folate status and eventually women's and children's health and well-being.
meet the average consumption needs of the entire population or only a fraction of the population (i.e., what proportion of the population would be expected to regularly consume adequate quantities of the target nutrients?) If only a fraction of the population is expected to have access to fortified flour or flour products, then would that fraction of the population be evenly distributed throughout the entire country or primarily in some geographic areas (i.e., will nutritional impact be expected across the country or mainly in some areas)?

b) Does the average consumer eat primarily commercially produced primary flour products (e.g., breads and pastas) or do consumers primarily make flour products at home? If consumers make flour products at home, do they store fortified flour in the original package or another container (i.e., if not all the flour is fortified, is it possible for the consumer to mix fortified and unfortified flour during storage)? Do households mix fortified and unfortified forms of flour (e.g., white and whole-wheat flour) when preparing flour products?

c) Can fortified flour or the flour products be easily identified in the markets and households by a product label or a fortification logo? Will all types and grades of flour be fortified? If not, which ones?

d) Do households typically store flour in the original package or transfer it to another container that would not carry a fortification label or logo?

The logic model of the flour fortification program in the specific context in a country should be consistent with the goals and specific objectives and strategies for food fortification in that country. Where such a strategic plan exists already, it should be the main source document for the logic model. Where the strategic plan does not exist, the logic model makes it easy to create one, because the model presents an overview of what activities lead to the logic model. Where the strategic plan does not exist, the logic model makes it easy to create one, because the model presents an overview of what activities lead to the same kind of clarity. A log frame is an analytic tool that derives its name from the logical linkages set out by the planners to connect a means with its ends.

**Step 3: Focus the M&E design**

Once consensus has been reached on the program description, the next step is to determine which aspects of the program should be tracked through the M&E component. This is discerned from the purpose of the flour fortification program and the intended uses and users of the M&E data and the resultant information.

Because the focus of the M&E system will shift with the maturity of the flour fortification program, the focusing step must be repeated at certain intervals. The length of the intervals will vary with the speed at which results can be obtained compared with the process and outcome indicators that are defined for each stage of the program. The M&E design is focused by determining 1) the purpose of the M&E component and 2) the intended users and users of the M&E data.

**Purpose of the M&E system.** The M&E system may have several purposes, including the following: 1) to enable the
program managers to take actions to rapidly correct problems or overcome unexpected difficulties encountered during program implementation; 2) to assess program effectiveness by providing information on how well the expected nutritional impacts are achieved; and 3) to demonstrate whether the financial and human resources expended on flour fortification should be sustained.

Users and uses of the M&E results. The users and uses of the M&E data must be determined as early as possible. The potential users must be given the opportunity to provide feedback on the design of the M&E plan as early as possible to ensure support for and future use of the data gathered.

The stage of implementation of the flour fortification effort will drive the focus of the M&E system at any point in time; during the early stages of flour fortification programs, the focus of the M&E system should be on monitoring program implementation, especially ensuring safe and adequately fortified flour, and achieving maximum population coverage. With more mature programs, the M&E system can focus on assessing and evaluating the impact of the program on the nutritional status of the population. Therefore, to develop realistic, measurable, and time-sensitive process objectives for the different components of the flour fortification program, expertise from all of the stakeholders must be tapped at various stages during the planning and implementation of the M&E system. This will lead to time- and situation-specific, realistic objectives about what should be achieved, e.g., the amounts of fortified flour produced and when and where fortified flour or flour products are available.

Step 4: Collect credible data

In order for the results of the M&E system to be accepted by the diverse stakeholders and for the M&E system to be sustainable, reliable and credible data must be collected. Relevant issues that may constrain or expand the range of feasible indicators include the following: the type of data to be collected – e.g., the presence of fortified flour or flour products in the household, the knowledge and attitudes of the consumers, and the nutrient status of the target population (e.g., women of reproductive age); the frequency of data collection; the cost of data collection, analysis, and reporting; how (including where) the data will be collected and who will be responsible for the data collection; the reliability and validity of the data to be collected; and the interpretation of results, including what would trigger an action, what actions would be taken, and by whom. Regular and periodic M&E data collection requires a clear commitment and willingness from all involved parties to share the relevant data and to collect new data or information as required. Once fortification has started, the flour industry must have in place a strong quality assurance/quality control (QA/QC) system, and appropriate food control and regulatory monitoring procedures must be followed to ensure that safe and adequately fortified flour (according to national standards) is continuously produced or imported and marketed in sufficient quantities to meet the dietary and nutritional needs of the populations. Such systems must be in place before population-level M&E activities are implemented. Population-level M&E of flour fortification programs is intended to assess the proportion of the population that has access to and consumes fortified flour or flour products (this is usually referred to as coverage monitoring) and to track whether the nutrient status of the population improves (this is usually referred to as impact monitoring).

Selection of indicators and criteria for success. Indicators are specific factors that can be assessed or measured to describe the characteristics of a flour fortification program and that allow program managers to gauge progress toward achieving the program objectives. The indicators “operationalize” the components of the flour fortification program that are the focus of the M&E system.

Some examples of process and outcome indicators and measures for a flour fortification M&E system are presented in Table 1. The choice of indicators and measures should be based on the characteristics and context of the flour fortification program. Additional information from the regulatory monitoring component of the fortification program will provide the data on the overall quantity of adequately fortified flour and flour products produced locally or imported. This will determine whether sufficient amounts of fortified flour are regularly available to specific populations so that potential nutritional benefit can be expected. The biological indicators for assessing the nutritional impact of the flour fortification program should be sensitive enough to reflect the changes that should occur in the nutrient or health status of the population after a given period of time.

As stated above, improvement in the nutrient and health status of a population depends on the proportion of the population that regularly consumes fortified flour (program coverage or market penetration) over a minimum period of time. For example, experience has shown that substantial change in the folate status of a population is observed relatively soon after sustained, high market penetration of fortified flour, and reductions in the incidence of neural tube defects can be detected after about 1 year (if large enough numbers of births can be monitored for birth defects).14–18 Experiences related to the impact of flour fortification on the iron status of populations is limited, especially because many countries...
Table 1  Examples of some process and outcome indicators for M&E of a flour fortification program based on the logic model in Figure 3.

<table>
<thead>
<tr>
<th>Measure</th>
<th>Indicator</th>
</tr>
</thead>
<tbody>
<tr>
<td>Industry production and distribution activities</td>
<td>Quantity of certified quality fortificant mix procured vs. total quantity of flour milled</td>
</tr>
<tr>
<td></td>
<td>Procurement of equipment for fortification</td>
</tr>
<tr>
<td></td>
<td>Total number of mills that have been licensed to fortify flour</td>
</tr>
<tr>
<td>Regulatory safety and control activities</td>
<td>Legislation and standards for flour fortification with vitamins and minerals in place</td>
</tr>
<tr>
<td></td>
<td>Design of adequate inspection of mills and enforcement of standards and regulations for fortified flour</td>
</tr>
<tr>
<td>Social marketing and communications activities</td>
<td>Proportion of materials (radio or TV spots, posters, counseling sessions, others) produced compared to planned</td>
</tr>
<tr>
<td></td>
<td>Proportion of materials (radio or TV spots, posters, counseling sessions, others) distributed and placed or aired compared to planned</td>
</tr>
<tr>
<td>Producers comply with good manufacturing practices (GMP) and quality requirements of existing regulations</td>
<td>Proportion of mills with QC/QA procedures in place for fortified flour</td>
</tr>
<tr>
<td>Increased production of fortified flour according to specifications</td>
<td>Proportion of Mill Technical audits conducted/planned</td>
</tr>
<tr>
<td></td>
<td>Proportion of fortified/unfortified flour produced</td>
</tr>
<tr>
<td></td>
<td>Tonnage of fortified flour produced</td>
</tr>
<tr>
<td></td>
<td>Percentage of the total flour market (including imports and small scale production) that is fortified</td>
</tr>
<tr>
<td></td>
<td>Proportion of fortification mix usage/planned</td>
</tr>
<tr>
<td></td>
<td>Proportion of flour samples adequately fortified according to standards at mill</td>
</tr>
<tr>
<td>Retailers increase stock and sales and promote fortified flour and byproducts</td>
<td>Proportion of flour samples taken/samples planned at mill</td>
</tr>
<tr>
<td></td>
<td>Proportion of flour samples adequately fortified at wholesale/retail in target regions</td>
</tr>
<tr>
<td>Consumers know to select or demand fortified flour and byproducts</td>
<td>Proportion of flour samples taken/samples planned at wholesale/retail</td>
</tr>
<tr>
<td>Increased awareness, knowledge, and attitudes of producers, retailers, and consumers</td>
<td>Awareness of nutritional benefits of fortified flour and/or byproducts</td>
</tr>
<tr>
<td></td>
<td>Awareness of fortified flour and/or byproducts logo</td>
</tr>
<tr>
<td>Increased purchase of fortified flour and byproducts</td>
<td>Proportion of target audience who are aware of key message</td>
</tr>
<tr>
<td></td>
<td>Proportion of target audience who are aware of presence of fortified flour in markets</td>
</tr>
<tr>
<td>Increased proportion of nonpregnant women of reproductive age regularly consuming fortified flour and byproducts</td>
<td>Proportion of households with flour “labeled” as fortified</td>
</tr>
<tr>
<td></td>
<td>Proportion of households with fortified flour</td>
</tr>
<tr>
<td></td>
<td>Proportion of households with fortified bread</td>
</tr>
<tr>
<td>Decreased prevalence of iron deficiency in women of reproductive age</td>
<td>Proportion of nonpregnant women of reproductive age consuming fortified flour regularly</td>
</tr>
<tr>
<td></td>
<td>Proportion of nonpregnant women of reproductive age consuming fortified bread regularly</td>
</tr>
<tr>
<td>Decreased incidence of neural tube defects (NTD)</td>
<td>Prevalence of anemia in nonpregnant women of reproductive age</td>
</tr>
<tr>
<td>Improved maternal health (millennium development goal)</td>
<td>Prevalence of iron deficiency in nonpregnant women of reproductive age</td>
</tr>
<tr>
<td>Reduced child mortality (millennium development goal)</td>
<td>Prevalence of iron deficiency anemia in nonpregnant women of reproductive age</td>
</tr>
<tr>
<td></td>
<td>Prevalence of low folate levels in nonpregnant women of reproductive age (serum or whole blood)</td>
</tr>
<tr>
<td>Decreased incidence of neural tube defects (NTD)</td>
<td>Number of NTD births/total number of births</td>
</tr>
<tr>
<td>Improved maternal health (millennium development goal)</td>
<td>Maternal mortality ratio</td>
</tr>
<tr>
<td>Reduced child mortality (millennium development goal)</td>
<td>Under-five mortality rate</td>
</tr>
<tr>
<td></td>
<td>Infant mortality rate</td>
</tr>
</tbody>
</table>

that have implemented iron fortification of flour have not systematically assessed the nutritional impact of those programs. Furthermore, the bioavailability of the elemental iron products commonly used in fortified flours is poor. Therefore, a recent expert panel recommended the use of the electrolytic form of elemental iron or, preferably, iron compounds such as ferrous sulphate, ferrous fumarate, or sodium-iron EDTA, in flour fortification, and in some countries these products are beginning to be used.

When a large segment of the population regularly consumes appropriately fortified flour or flour products for at least 6–12 months, the program’s impact could be assessed by comparing current nutrient and health status data with data from the pre-fortification period. We suggest 80% as a minimum population coverage rate from a practical perspective. With this percentage of coverage, it would be relatively easier to identify households (and thus individuals) in a geographic area of interest (country, region) that regularly consume fortified flour products and that would thus be expected to benefit nutritionally. In Guatemala, for example, a study found that the wheat flour-fortification program would likely improve the folate status of women in urban areas, but not among poor, rural, and indigenous women, because only sufficient proportions of the former group regularly consumed sufficient fortified flour. Improvements in the iron status of those urban women are likely to be modest, however, because of the known poor bioavailability of the elemental form of iron used as a fortificant.

The following issues should be considered when selecting indicators for population-level M&E of flour fortification programs:

Selection of target population groups to assess and monitor the impact of flour fortification. Another key consideration for M&E systems that include monitoring of nutrient status is deciding which population group or groups to monitor once the fortified flour and flour products are widely accessible and consumed. The population groups should be selected on the basis of available epidemiologic data that identify the groups in the appropriate geographic areas that have high prevalences of the target nutrient deficiencies and sufficient consumption of fortified flour and flour products to exhibit a substantial change in nutrient status (especially in folate or iron, which are commonly added to fortified flour). In most countries, women of reproductive age fit this profile and should be considered the prime group for monitoring the impact of flour fortification. Thus, women of reproductive age will be used as the target population in the discussions below.

Groups of women could be recruited for assessment through different approaches based on local needs, conditions, and technical, logistical, and funding capacity. Most public nutrition programs around the world are familiar with carrying out population-based nutrition surveys with the use of statistical methods for population-representative sampling. Such nutrition surveys can be carried out as baseline measures and again periodically to enable program managers to track changes in population coverage and program impact. However, such nutrition surveys are relatively costly and many countries may not have the resources to carry them out frequently enough, especially during the early stages of program implementation when adjustments and improvements in the strategy may be most needed to steer the effort toward ensured success.

Thus, an alternative, more affordable approach to monitoring the coverage and impact of flour fortification programs among target populations is based on the following premises: 1) the strength of data collection in a monitoring system is that this is an ongoing activity for tracking trends in particular indicators; and 2) the primary interest is to determine whether changes in the prevalence of key indicators are taking place in the expected direction, rather than to determine the exact prevalence of any specific indicator. Therefore, the monitoring system is not necessarily intended to provide statistically representative data on all indicators of interest or to provide answers to all questions that the system might generate; if needed, special studies can be carried out for such purposes.

Thus, convenience sampling may be based on purposive selection of women of reproductive age in selected communities across a country (or other geographic area), such that the data on household coverage and folate or iron status of women selected will adequately reflect the actual trends in the population as a whole. We refer to this data collection approach as reflective sampling. One example can be recruiting women of reproductive age through sentinel primary health clinics or worksites that employ large numbers of women or recruiting students in sentinel schools to report on their households’ use of fortified flour or flour products.

Existing vs. new data systems. When designing an M&E system, it is usually more economical (and thus preferable) to use existing data systems and indicators as much as possible or to include new indicators in systems that are already operational and time-tested. A new data system should be created only if absolutely necessary. Most M&E systems will require the use of both quantitative and qualitative data. Examples of existing national data sources that can be used to monitor and evaluate flour fortification programs include 1) core health statistics data, such as anemia among nonpregnant women or pregnant women in the first trimester and birth defect...
records from maternity hospitals; 2) on-going household expenditure or consumption surveys; 3) the Demographic and Health Survey (DHS); 4) the Multiple Indicator Cluster Survey (MICS); 5) CDC-supported reproductive health surveys; 6) the World Bank Living Standards Measurement Study (LSMS); 7) household expenditure and income data for transitional economies (HEIDE); 8) ministry of trade and industry surveys; and 9) ministry of agriculture or economy statistics on flour production or importation.

These data sources may or may not already include relevant indicators that could provide data for the M&E system that can be “triangulated” in assessments of program coverage and impact. If not, it is highly recommended that the national flour fortification program managers work with the implementers of these surveys to incorporate a minimal number of key indicators. Other important and potential data sources for a flour fortification M&E system include flour production and distribution figures from flour mills and baking companies or wholesale markets and information obtained through the legal monitoring of the industry by the government food inspection system.

Cross-sectional nutrition surveys. Cross-sectional nutrition and health surveys can provide statistically representative information on the nutritional and health status of the population and on the coverage and implementation status of an intervention such as flour fortification. When performed before or within the first few weeks after the flour fortification program starts, such surveys provide baseline data against which the success of the intervention can be based.

Periodic cross-sectional surveys can answer whether any significant changes have occurred in the magnitude and distribution of selected vitamin and mineral deficiencies and can show trends in the proportion of the population that is aware of the fortification program, has regular access to the fortified products, and regularly consumes the fortified flour and flour products. The basic categories of data collected through such surveys include 1) the demographic characteristics of the survey population; 2) the target group’s knowledge, attitudes, and practices related to the flour fortification program and related interventions (e.g., any vitamin-mineral supplementation efforts, social marketing, or promotion strategies; 3) the target nutrient content of the fortified flour and flour products and related labeling information; and 4) biochemical measurements of relevant vitamin and mineral status. The results and information from such surveys will inform all stakeholders as to whether the objectives of the flour fortification program are being achieved.

When planning a cross-sectional nutrition survey, it is important to consider whether the estimates desired are national or subnational, how the survey subjects will be recruited (e.g., in households, clinics, or schools), whether the survey will be an independent activity or will be incorporated as part of another planned survey, and the sample size. The determination of appropriate sample size for baseline and follow-up must be carefully considered. This is frequently overlooked, and sample size calculations are made without considering that comparisons need to be made before and after the intervention for M&E purposes. The sample size estimates and sampling approach for this before-and-after comparison require several assumptions as described in detail elsewhere.

In many countries, a non-stratified national cluster survey can provide an unbiased overall national estimate of nutrient status (i.e., proportion of the target population who are deficient). If there is a need to determine statistics for a subgroup or strata in a given country because of expected differences in program coverage (i.e. urban and rural or by income) or other factors that might result in significant expected differences in the success of the fortification program by geographic area (i.e. border regions), each area is considered an independent survey universe, and a separate cluster survey is performed in each area (referred to as a stratified cluster survey). This approach provides unbiased estimates for each strata but requires a larger overall sample size and accompanying budget.

Sentinel monitoring. Sentinel monitoring refers to the monitoring of key indicators in the general population or in selected subpopulations. Sentinel sites are useful for tracking trends and capturing changes in the indicators of interest over time; the implementation, progress, and impact of flour fortification programs can be monitored through this approach. Sentinel public health clinics, schools, worksites, or other settings in which large numbers of “average” or “typical” persons can be readily accessed in specific geographic areas or communities, can serve as timely sources of information on many issues, including population coverage, changes in consumption, and impact of population-based flour fortification programs. The data and resulting information from sentinel sites can be sufficient for detecting long-term trends in indicators of interest and for making decisions to strengthen program implementation.

Although data from a sentinel flour fortification monitoring system may not be statistically representative of the situation among the population at large, the data can nevertheless be reflective of the general situation and serve as a continuous “watch” on how the program is evolving and moving toward achieving its objectives. In Morocco, for example, a new Food Fortification Monitoring System is designed to measure the effectiveness of multiple-food vehicle fortification programs. This system
combines industry-based and legal monitoring data and information on the quantity and quality of fortified foods with data on population access, household coverage, and nutritional impact collected through several sentinel primary health centers and schools throughout Morocco. Thirty-eight sentinel health centers in urban and rural communities in 27 provinces and 16 regions of Morocco were selected.

To obtain baseline data on the nutrient status of children aged <5 years and of women of reproductive age, approximately 100 children seen in the sentinel health centers for routine well-child visits and the women of reproductive age accompanying them were assessed. Hemoglobin was measured with use of the HemoCue® (HemoCue, Lake Forest, California, USA) system, and iron, folate, and vitamin A status were determined by measuring serum ferritin, serum folate, and serum retinol, respectively. Brief questionnaires were also administered to the women to assess their knowledge and purchasing habits of the fortified foods. In addition, students in local schools geographically closest to each sentinel health center provided information on the brands of flour, salt, cooking oil, and milk in their homes and whether they were labeled as being fortified or carried the national fortification logo. The students also brought to school samples of flour used for baking bread in their homes. Nurses from the sentinel health centers tested the flour samples for the presence of fortificant iron by using the “spot test” (AACC Spot Test for Iron in Flour, AACC Method 40–40; http://www.aaccnet.org/ApprovedMethods/toc.htm).

The plan for the Food Fortification Monitoring System calls for repeating the questionnaire component and the school-based data collection activities every 6 months. When these data indicate that close to 90% of households have regularly used the specific fortified food vehicles for 6–12 months, the laboratory-based tests for anemia and measurements of indicators for iron, folate, and vitamins A and D will be repeated to assess any changes in the nutrient status of the sentinel children and women.

Another example of a food fortification monitoring system comes from Guatemala, where the Institute of Nutrition of Central America and Panama (INCAP) identified two types of sentinel surveillance sites: in selected geographic areas or communities that would provide information on the health and nutrition status of the population and in selected health facilities.24 The latter system focused on M&E processes involved in the delivery of health care, as well as the interaction between health care providers and users. UNICEF-Guatemala has been supporting the Project Sentinel Schools, whereby accessibility of vitamin A–fortified sugar and iodized salt by the households is monitored by testing samples brought by students. The sentinel schools monitoring system enables authorities to assess the household coverage of adequately fortified salt and sugar in rural communities. Every year, students in >420 rural public schools in Guatemala bring samples of salt and sugar from their homes, and those samples are sent to INCAP for analysis.24,25 The Guatemalan Ministry of Health also carries out regulatory monitoring of the nationally mandated fortified salt and sugar at production sites and in retail outlets.

Important points to be considered very early in the design and planning of a food fortification monitoring system when selecting sentinel data collection sites include the following: 1) the location of the sentinel sites should reflect different settings if the marketing and availability of one or more fortified food vehicles are expected to vary in certain areas (e.g., rural vs. urban) or regions (e.g., north vs. south) of the country; 2) administrators and staff of the sentinel sites should be supportive, and the facility should have the minimum capacity (e.g., infrastructure, personnel, instruments, and tools) to carry out the monitoring activities independently or with available additional support; 3) the sites should serve large enough numbers of the “average” consumer on a regular basis so it is possible to easily recruit the necessary number of subjects in a timely manner; 4) the sites must have a basic capacity to collect and test relevant food and biological samples or to easily and rapidly transfer the same to facilities that can perform the needed tests.

Lot quality-assurance sampling or supervision. The Lot Quality-Assurance Sampling (LQAS) method was developed by industry to assess the quality of batch production to determine at reasonable cost whether a product standard had been reached. LQAS uses binomials to test with a known level of error whether the standard has been reached. During the last 20 years LQAS has been adapted to public health as a practical field method for assessment of preventive and curative health services.26 A recent global review by WHO and the World Bank of more than 800 applications discussed its use for program planning, monitoring, and evaluation of community health programs. It has been used mostly for assessing immunization coverage; knowledge, attitudes, and practices in maternal and child health; family planning and HIV/AIDS programs; and assessing the quality of health workers’ performance.27 LQAS has been used to assess exclusive breastfeeding, complementary feeding, continuing breastfeeding, and Global Acute Malnutrition. LQAS has not been widely used, as yet, to measure food fortification programs, but there are no inherent characteristics of these programs that suggest that LQAS could not be used for this purpose. To use LQAS, health system managers need to identify two thresholds. For example, in
flour fortification we might be interested in the proportion of the service delivery points in a province that sells fortified foods – the upper threshold or coverage target (e.g., 80%) and the lower threshold considered an acceptably low level of coverage (e.g., 50%). It may be useful to think of the LQAS method as a triage system in which 1) successful supervision areas at or above the upper threshold are identified so it is known that they are not the source of programmatic problems, and 2) supervision areas at or below the lower threshold are identified as priorities because they have unacceptably low access to fortified foods. The middle area of the triage system, between the upper and lower thresholds, is an indeterminate zone. In practice, supervision areas (i.e. provinces, counties or communities) with access closer to the upper threshold are more likely to be classified as reaching the target, while areas with coverage closer to the lower threshold are more likely to be prioritized as substandard. In either instance, the error is not serious. If a supervision area, which is slightly below the upper threshold, is judged as having reached it, this error is not pernicious since it is more important for the health system to focus on areas where larger proportions of the population are at risk due to low access to fortified foods. Correspondingly, an area that is slightly above the lower threshold and judged as well below the coverage target does not cause worrisome error since the supervision area has not reached the coverage target and would have to be dealt with sooner or later.

Several characteristics have made LQAS attractive to health system evaluators. First, only a small sample is needed to judge whether a supervision area has reached the predetermined coverage target. With a relatively small sample per area, data collection does not seriously compete with time for provision of health services. Second, the LQAS sampling procedures and analyses are relatively simple. This simplicity is welcome to overworked supervisors and health workers who need management tools that can easily be understood and applied. These two characteristics have made LQAS valuable as a practical management tool for the monitoring and evaluation of community health services. Another attractive feature of LQAS designs is that the data from individual provinces or strata are pooled into an estimate of coverage for an entire program area. In typical applications, when all strata are pooled the result from each province is weighted by the size of its population. To date, the data aggregation across multiple provinces assumes that data has been collected in all provinces.

A decentralized monitoring system using Lot Quality Assurance Supervision (LQAS) is planned to be implemented soon in a country strengthening its monitoring system for flour fortification (Vargas W, 2007 personal communication). This system could provide representative data at the province level and will allow identification of priority counties (municipios) for availability, consumption of fortified foods, and knowledge of the fortification benefits. A sample of 19 women of reproductive age will be sampled in each supervision area (county or municipio) for a total national sample of 4,389 women of reproductive age (15–49 years); this will provide an acceptable error for managerial decision making. At least 92% of the time, supervision areas that have reached the coverage goal (to be determined annually) will be identified correctly. The decentralized monitoring system LQAS has been designed to assess national coverage of fortified foods in households from 32 provinces and 225 counties.

When LQAS data are collected recurrently at multiple time points, they can be used to measure changes in specific parameters over time and in different areas, which could be useful in the M&E of flour fortification programs. LQAS data collections may be particularly useful for identifying areas or population groups in which the consumption of fortified flour and flour products risks falling below a predetermined threshold in countries with successful, ongoing flour-fortification programs.

**Quality and quantity of M&E data.** A good monitoring system produces information that is accurate, reproducible, informative, and timely. Monitoring systems should be periodically reviewed to allow for improvements in data quality and usefulness and process efficiency. Efforts to ensure quality in the collection of data and in the handling and processing of food and biological specimens for testing must be complemented by adequate planning and procedures for computer data entry and analysis. The quality of the data collected can be influenced by the following factors: the ease of use of the data collection forms and instruments, the clarity of each variable and how completely the data are collected and recorded, the ability of data collectors to correctly utilize the data collection tools (e.g., questionnaires and data-collection forms and methods and the instruments used to collect, process, and test specimens), the proportion of nonresponders, the ability of the field staff to correctly record and computerize data, the proportion of nonresponses or erroneous data recorded, and the skills of the analysts to check for and correct data errors.29,30

In addition to the quality of the data collected, the quantity must also be considered. It is essential to understand that the collection of more data does not define a better monitoring system. As mentioned above, the purpose of the collection of each data point (indicator or variable) and how the resulting information will be used should be considered carefully. The motto should be: “There is no need to collect any data that will not be
readily used to guide and improve the program.” It is best to begin with the minimum key data points required that can be easily collected and readily analyzed, interpreted, and reported. Once a monitoring system is well established, the need to incorporate additional data and information can be considered.

Logistics. The cost of equipment, transportation of the data collectors and specimens to different locations, organization of the pilot data-collection exercises, staff training and capacity building for the different components of the system, and other administrative tasks should be addressed early. Those who collect, enter, analyze or interpret the data, or prepare and disseminate the information should be identified and trained and their responsibilities delineated.

Step 5: Justify the conclusions

Data analysis is the process of calculating, tabulating, and classifying the results; interpreting and presenting the information generated in an understandable manner; and making appropriate action recommendations to different stakeholders. The primary purpose of the M&E component of a flour fortification program is to enable the stakeholders to sustain successful aspects of the program and to improve weaker components. Thus, the social and political context of the program and the needs of the stakeholders must be taken into consideration in the analysis and presentation of the results, without compromising the integrity of the M&E system. To express any degree of progress, the results of analysis should be compared with national or previous similar data. For example, the Integrated Nutritional Surveillance System (SIVIN), an ongoing centralized, modular, and integrated information system that has been in place in Nicaragua is able to compare data from various sources to provide a general overview of the performance and effectiveness of ongoing nutrition programs. The results of the indicators obtained in the SIVIN household survey are compared with the same indicators in previous surveys such as the Family Health and Demographic Survey (ENDESA) of 1998 and 2001 and to the National Micronutrient Survey (NMS) in 2000.31 When interpreting the results of M&E, it is important to keep the goals of the flour fortification effort and the audiences of the report in mind. It is advisable to consider the limitations, including possible biases; the validity; the reliability; and the generalizability of the results. Potential alternative explanations should be explored and presented. Comparisons with data obtained from other sources should be made and the meaning of the differences, if any, discussed.

Step 6: Ensure that the lessons learned are shared and used

Once the data are analyzed and interpreted, it is important to make recommendations for action. The recommendations should be tailored to the different stakeholders to make them “doable”, i.e., relevant and useful. The purposes of M&E will shape the scope and kind of sensible recommendations. If M&E efforts have been conducted properly, reasonable conclusions and recommendations can be made to improve the program. One important point raised by some organizations refers to the ownership that decision makers have of the M&E system reports.32 When decision makers have a sense of ownership or involvement in the monitoring process, they are more likely to embrace the results and follow up by taking action.

Dissemination is the process of communicating the lessons learned in a timely, unbiased, and consistent manner. Effective communication will require consideration to timing, style, tone, key message, and format. The results can be presented in written or oral format. If a report is written, it should be easy to understand. Oral presentation to selected audiences should also be considered to increase the likelihood that the results are used and that the recommendations are discussed and followed. Interim reports are useful tools for keeping the audiences up to date on the progress and the ongoing activities.

THE MONITORING AND EVALUATION PLAN FORMAT

An M&E plan for a flour fortification program is the guiding document that describes how the M&E will be carried out. The plan is informed by the six steps described above. The following components of a plan have been adapted from the “Evaluation Plan Template” tool.33 A complete M&E plan should include the following components: Executive Summary; Introduction; Background; Questions; Purpose; Limitations; Audiences; Design and Methods; Logistics; Decisions; Timeline; Monitoring and Evaluation Team; Budget; References and Appendices.

WHEN TO START ASSESSING THE EFFECTIVENESS OF A FLOUR FORTIFICATION PROGRAM

Assessing the effectiveness of a flour fortification program refers to tracking improvements in the nutrient status of target populations over time. Under “ideal” circumstances, the initial or baseline data on the nutrient status of the target population would be collected just
prior to the time fortified flour and flour products are marketed. Follow-up and ongoing periodic data collection would then begin once the population coverage data is thought to be ample, at least theoretically. Consumption of fortified flour and flour products should occur regularly. The periodicity of collecting effectiveness data depends on how quickly a biological or clinical indicator is expected to show an improvement in nutrient status in a population on the basis of the nutrients added to the flour. For example, most countries around the world currently fortify flour with at least iron, or with iron and folic acid,\(^4\) therefore one can possibly expect to document a substantial increase in serum folate levels of the population in approximately 4–months. Furthermore, a 25–50% reduction in the incidence of neural tube defects among newborns can be expected within 12–24 months after large-scale (mandatory) flour fortification.\(^5\)

Some studies from Venezuela\(^34\) and China (Huo J, personal communication) report reductions in the population prevalence of anemia and iron deficiency after \(\geq 18\) months of flour fortification, which suggests that this time period may be sufficient to see the first change in the prevalence of anemia and iron deficiency in countries in which this vehicle is a staple and consumption is sufficient. Very recent observations from household consumption and biological measurements in Central Asian countries are in agreement with the time-lag of 18 months to 2 years before meaningful changes can be observed in hemoglobin and ferritin concentrations among members of the most vulnerable population groups (Tazhibayev S, personal communication).

Considering the minimum fortification of flour with folic acid and iron, we propose that the effectiveness-monitoring data (based on serum folate measurements and/or neural tube defect surveillance to track folate status and serum ferritin and/or hemoglobin measurements to track iron status) be collected, analyzed, and reported as frequently as needed, once the coverage-monitoring data indicate that an ample proportion of the target population (i.e., >80%) is consuming the fortified flour and flour products regularly. If funding is limited, effectiveness data could be collected less frequently. Other relevant questions addressing the effectiveness of the flour fortification program at the population level include the amount of food consumed, the estimation of micronutrient intake from the fortified food, improvement in biochemical indicators, and the prevalence of nutrient deficiencies.

A clear understanding of the stakeholders in flour fortification and their needs, the description and context of the fortification program, the country’s wheat flour and flour products market, and the resources available for the M&E component are critical and should be considered early in the program’s design.

**Acknowledgments**

The authors would like to thank Mary Cogswell, Laurence Grummer-Strawn, Rick Hull, and Cheryl Lucas from the Centers for Disease Control and Prevention and Joseph J. Valdez from the World Bank for the critical comments and thoughtful suggestions during the review of this article.

Disclaimer: The findings and conclusions of this review are those of the authors and do not necessarily represent the views of the Centers for Disease Control and Prevention (CDC).

**REFERENCES**


