

10-15-2017

Evaluation of Programs to Improve Complementary Feeding in Infants and Young Children

Edward A. Frongillo Jr.
efrongil@mailbox.sc.edu

Follow this and additional works at: https://scholarcommons.sc.edu/sph_health_promotion_education_behavior_facpub



Part of the [Public Health Education and Promotion Commons](#)

Publication Info

Published in *Maternal & Child Nutrition*, Volume 13, Issue S2, 2017, pages e12436-.

This Article is brought to you by the Health Promotion, Education, and Behavior at Scholar Commons. It has been accepted for inclusion in Faculty Publications by an authorized administrator of Scholar Commons. For more information, please contact digres@mailbox.sc.edu.

SUPPLEMENT ARTICLE

Evaluation of programs to improve complementary feeding in infants and young children

Edward A. Frongillo 

Department of Health Promotion, Education, and Behavior Arnold School of Public Health, University of South Carolina, Columbia, South Carolina, USA

Correspondence

Edward A. Frongillo, Department of Health Promotion, Education, and Behavior, 915 Greene Street, Room 529, Discovery I, University of South Carolina, Columbia, South Carolina 29208-4005, USA.

Email: efrongil@mailbox.sc.edu

Abstract

Evaluation of complementary feeding programs is needed to enhance knowledge on what works, to document responsible use of resources, and for advocacy. Evaluation is done during program conceptualization and design, implementation, and determination of effectiveness. This paper explains the role of evaluation in the advancement of complementary feeding programs, presenting concepts and methods and illustrating them through examples. Planning and investments for evaluations should occur from the beginning of the project life cycle. Essential to evaluation is articulation of a program theory on how change would occur and what program actions are required for change. Analysis of program impact pathways makes explicit the dynamic connections in the program theory and accounts for contextual factors that could influence program effectiveness. Evaluating implementation functioning is done through addressing questions about needs, coverage, provision, and utilization using information obtained from process evaluation, operations research, and monitoring. Evaluating effectiveness is done through assessing impact, efficiency, coverage, process, and causality. Plausibility designs ask whether the program seemed to have an effect above and beyond external influences, often using a nonrandomized control group and baseline and end line measures. Probability designs ask whether there was an effect using a randomized control group. Evaluations may not be able to use randomization, particularly for programs implemented at a large scale. Plausibility designs, innovative designs, or innovative combinations of designs sometimes are best able to provide useful information. Further work is needed to develop practical designs for evaluation of large-scale country programs on complementary feeding.

KEYWORDS

child feeding, children, complementary feeding, evaluation, nutritional interventions, program evaluation

1 | INTRODUCTION

Provision of timely, safe, and nutritionally rich foods in sufficient quantity, in addition to breastmilk from 6 to 23 months of age, is important for child growth and development (Bhutta et al., 2013). In low- and middle-income countries, complementary feeding practices are often inadequate in timing of introduction and nutrient quality of complementary foods (Arabi, Frongillo, Avula, & Mangasaryan, 2012); inadequate complementary feeding is one of several contributors to high prevalence of child undernutrition. Although there is some evidence that intervention to improve complementary feeding can be efficacious, evidence for effectiveness of large-scale programs to improve complementary feeding is limited (Bhutta et al., 2013).

Learning how to improve complementary feeding practices at scale is a high priority.

Evaluation of complementary feeding programs is needed to enhance knowledge on what works, to document responsible use of resources invested, and for advocacy. Evaluation is

the systematic application of social research procedures for assessing the conceptualization, design, implementation, and utility of social intervention programs ... Evaluators use social research methodologies to judge and improve the ways in which human service policies and programs are conducted, from the earliest stages of defining and designing

programs through their development and implementation
(Rossi & Freeman, 1989).

Evaluations document whether and to what extent goals and targets of the program were met and identify lessons learned and effective approaches for future planning (UNICEF, 2016). To be feasible and useful, evaluations—including their design, tools, and processes—must align with program design, information needs, and the time and resources available (UNICEF, 2016).

Evaluation is done to learn and to influence decisions. Decisions that need to be made regarding interventions and programs concern: (a) management and administration, for example, appropriateness of program changes, improvement of delivery, and accountability to funders; (b) planning and policy, for example, testing of innovation, expansion or curtailing of programs, and advocacy; and (c) testing of scientific hypothesis or professional practice (Rossi & Freeman, 1989). Evaluation is done throughout the three periods in the life cycle of a program: (a) during conceptualization and design of the program to confirm paths through which program should achieve impact, (b) during implementation to understand and improve functioning, and (c) after period of implementation to determine effectiveness (Figure 1; Rossi & Freeman, 1989).

Despite a rich literature on the importance of and methods for evaluation of programs, in practice, evaluations often are carried out in a limited way or not at all. Given the emergent opportunity to improve complementary feeding globally, commitment to and knowledge about conducting evaluation of large-scale programmatic efforts is particularly important. This paper aims to explain the role of evaluation in generating information for each of the three periods of the program life cycle with a focus on complementary feeding, and to illustrate each using the innovative, mixed methods from the “Alive & Thrive” project.

Alive & Thrive was funded by the Bill and Melinda Gates Foundation with the primary aim to learn “how to strengthen delivery systems to programmatically achieve scale in distinct country contexts and how to motivate behavior change in different populations, from decision makers to mothers” (Piwoz, Baker, & Frongillo, 2013). Alive & Thrive designed and implemented scaled-up programs in infant and young child feeding (IYCF) in Bangladesh, Ethiopia, and Vietnam beginning in late 2008 with the development of program model for each country (Baker, Sanghvi, Hajeerbhoy, Martin, & Lapping, 2013). Alive & Thrive also funded a set of small grants for intervention projects to improve IYCF in multiple countries. Examples for this paper are taken from the main project, primarily Bangladesh and Vietnam, and the small-grant project in Mexico.

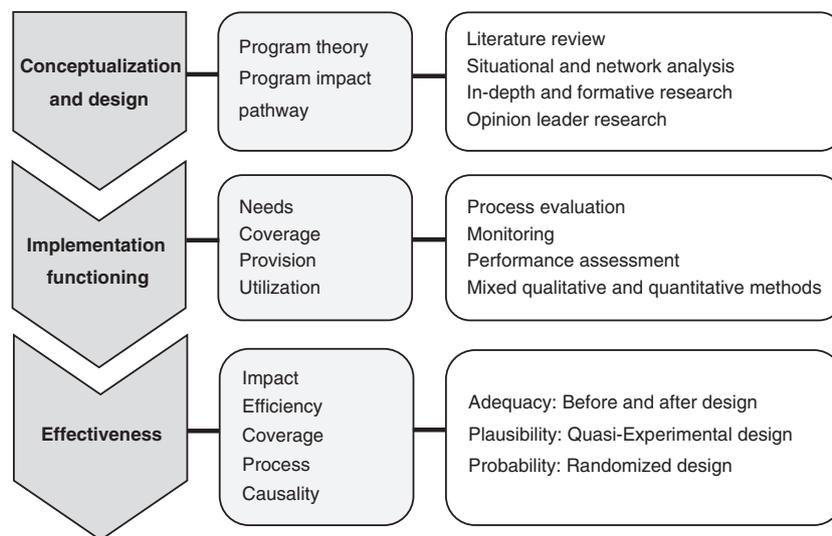


FIGURE 1 Evaluation through the program life cycle

Key messages

- Evaluation of complementary feeding programs is needed to enhance knowledge on what works, to document responsible use of resources, and for advocacy.
- Evaluation should be conducted from the beginning of and throughout program conceptualization and design, implementation, and determination of effectiveness.
- Essential to evaluation is articulation of a program theory on how change would occur and what program actions are required for change.
- Evaluations sometimes cannot use randomization, particularly for programs implemented at large scale, and practical designs that account for the complexity of systems are needed for evaluation of large-scale country programs on complementary feeding.

2 | CONCEPTUALIZATION AND DESIGN

Essential to the success of any evaluation is the articulation of a program theory, which specifies “what must be done to achieve the desirable goals, what other impact may also be anticipated, and how these goals and impacts would be generated” (Chen, 2015). The purpose of a program theory is to analyze and make explicit the “assumptions by stakeholders about what action is required to solve a social problem and why the problem will respond to this action” (Chen, 2015). These assumptions are about how change would occur (i.e., the causal processes through which a program is supposed to work) and about what program actions are required for change through these processes (Chen, 2015).

Although logic models are often used to depict programs, program theory is not the same as a logic model. Program theory systematically lays out the assumptions about change and actions underlying a program and the plausible pathways through which the program has impact. In contrast, a logic model describes program components—for example, listing inputs, activities, outputs, and outcomes—as generic categories. A logic model typically is not detailed enough to represent program theory. Instead, an analysis of program impact pathways can be used. Unlike a static logic model, an analysis of program impact pathways makes explicit the dynamic connections of program inputs from delivery through recipient utilization of those inputs and impact while accounting for factors in that context that could influence program effectiveness (Avula et al., 2013; Kim, Habicht, Menon, & Stoltzfus, 2011). The analysis of program impact pathways assists during implementation by helping to identify specific corrective or additional actions that emerge as essential to successful implementation and ultimately impact. It also assists during evaluation of effectiveness by providing an explicit picture of how change was supposed to have occurred, guiding what data to collect, impact analyses, and their interpretation.

To contribute to developing the program model, systematically obtained information is needed from assessments to understand the context, opportunities and challenges, and lessons learned from past programmatic experience. Methods often used for this purpose are literature review, analysis of existing data, qualitative data collection, and consultations with stakeholders.

Alive & Thrive articulated a socioecological model for change consisting of four levels: (a) individual; (b) interpersonal; (c) community: organization, service providers, and products; and (d) enabling environment: policy and legislation, politics and conflict, economics, religion, technology, and natural environment (Baker et al., 2013). The development of the IYCF programs in Bangladesh, Ethiopia, and Vietnam reflected ideas from multiple behavioral models, including those related to stages of behavioral change, reasoned actions (e.g., intentions and norms), interpersonal interactions, self-efficacy, and learning from role models, and also ideas from community-level models (e.g., diffusion of innovation through social networks). Based on these models, to develop strategies tailored to the country contexts and learn about what could affect implementation and adoption, Alive & Thrive conducted many studies, including a desk review of the literature, stakeholder meetings, situational analysis, formative research, media audits, opinion leader research, network analysis, surveys of

the knowledge and practices of frontline workers, market analysis of commercial complementary foods, assessment of private-sector interest in and capacity to produce a fortified product, and assessment of models for IYCF counseling services. Constraints on feeding practices were identified, and strategies to address them were incorporated into a theory of change. The theory of change articulated what impact would be expected to be achieved if the constraints were addressed with specified strategies. The tailored strategies were then developed in detail and documented in analyses of program impact pathways (Avula et al., 2013; Nguyen, Menon, et al., 2014).

Monterrosa et al. (2013) developed an intervention in Mexico for improving IYCF that used a communication strategy in which scripted messages were delivered through nurses and radio, aiming to change beliefs, attitudes, social norms, intentions, and behaviors related to breastfeeding, dietary diversity, and food consistency. The strategy was based on the results of a prior ethnographic study that examined maternal knowledge and developed an emic knowledge framework to help explain and interpret maternal complementary feeding behaviors (Monterrosa, Pelto, Frongillo, & Rasmussen, 2012). This in-depth qualitative study identified eight concepts held by mothers and showed that maternal feeding decisions were mostly determined by the highly organized knowledge and beliefs of mothers. From this foundation, the intervention was developed using the theory of planned behavior and a social marketing approach, resulting in five scripted messages to be delivered that targeted beliefs and attitudes underlying the key behaviors to be improved. The five messages were about breastfeeding, food consistency, flesh foods, vegetables, and feeding again if food was rejected. Focus-group discussions were used to develop and refine messages and study materials.

3 | IMPLEMENTATION FUNCTIONING

Evaluating implementation functioning is done through addressing four categories of questions: needs, coverage, provision, and utilization (Habicht, Victora, & Vaughan, 1999). The question for needs is, are the needs of targeted recipients understood and addressed by the program? The question for coverage is, are the target recipients being reached? The four questions for provision are as follows: are the services available and accessible? Are the services in line with design specifications? Is their quality adequate? What resources are being expended? The question for utilization is, are the services being used?

Information to address these questions can be obtained through process evaluation, operations research, and monitoring of program processes and performance. Possible designs for assessing implementation functioning are discussed in Section 4. Measures for implementation functioning, both what is being accomplished and how, can be chosen or created to address the questions regarding needs, coverage, provision, and utilization at three levels: systems, program implementers, and program recipients. At the system level, we can measure staff and their roles and the infrastructure meant to support them, including location and operation of facilities. For program implementers, we can measure contacts with recipients, provision of services, distribution of goods and materials, knowledge and skills, drive, attitude, autonomy, respect in community, workload, remuneration, training, tools, support,

and supervision. For program recipients, analogously, we can measure exposure to and use of services, attendance, knowledge and skills, drive, attitude, autonomy, influence of authority figures, workload, poverty, illness, mental health, training, tools, and support. Further information about procedures and measures for assessing implementation functioning is provided in Saunders (2016).

Theory-driven methods for process evaluation were used in Alive & Thrive to generate learning on processes and pathways to program impact (Rawat et al., 2013). The data collection was linked closely with detailed program impact pathway models, using mixed methods and multiple data sources, and with program implementation timelines, engaging with program implementation and management teams. In Bangladesh, an analysis of program impact pathways identified what was important for implementation (e.g., the role of paid and volunteer staff) and utilization (e.g., resource and time constraints that require complementary interventions; Avula et al., 2013). Mixed qualitative (i.e., interviews and observations) and quantitative methods were used to examine the content of training materials; IYCF knowledge; communication with mothers; and what influenced promotion, trial, and adoption of IYCF practices (Avula et al., 2013). In Vietnam, similar methods were used to examine the pathways through which the social-franchise model was intended to improve IYCF practices (Nguyen, Menon, et al., 2014). Six components were assessed: franchise management, training and IYCF knowledge of health providers, service delivery, program exposure and utilization, maternal behavioral determinants (i.e., knowledge, beliefs, and intentions) toward optimal IYCF practices, and IYCF practices (Nguyen, Menon, et al., 2014). Process data collected 12 months after the launch of the first franchises were used to examine the quality of facilities, service delivery, and client perceptions and use (Nguyen, Kim, et al., 2014). Quantitative data documented, for example, the coverage reported by mothers for exposure to interpersonal counseling and mass media (Nguyen et al., 2016).

For the Mexico intervention, nurses delivered each of the five messages to each mother once, and the same messages aired seven times each day on three radio stations for 21 days in the intervention communities; the control communities were in a different state from the intervention communities and were not exposed to the scripted messages (Monterrosa et al., 2013).

4 | EFFECTIVENESS

Evaluating effectiveness is done through assessing impact, efficiency, coverage, process, and causality (Table 1). Impact asks what has happened and addresses what were the benefits and harm. Efficiency asks what it cost and addresses whether the benefits and harm are worth the costs incurred. Coverage asks who was reached and addresses who received the actions (and benefits and harm). Process asks how did it happen and addresses what factors were in place and actions that occurred. Causality asks why did it happen and addresses why the factors and actions together resulted in the benefits and costs.

Three categories of designs can be used for evaluation of effectiveness: adequacy, plausibility, and probability (Table 2; Habicht et al., 1999). Adequacy designs ask: did the expected changes occur? Adequacy designs often use two sets of measurements, before and after program implementation, but no control group. Plausibility designs ask, did the program seem to have an effect above and beyond external influences? Plausibility designs are often quasi-experimental such as using a historical control or a nonrandom comparison group with two or more sets of measurements before and after program implementation. Probability designs ask, was there an effect? Probability designs use a randomized comparison group, often with sets of measurements before and after program implementation.

Adequacy designs do not attempt through the features of the design to directly attribute observed changes to the program, and instead aim to assess whether the observed changes are consistent with what was expected in magnitude and direction (Habicht et al., 1999). Adequacy designs can demonstrate that a program was feasible to implement and capable of generating changes in line with expectations. Sometimes strong arguments can be marshaled to support attribution based on program theory and the ruling out of competing explanations. Often demonstration of adequacy of a program is followed by a second, larger study using a plausibility or probability design that is better able to control for confounding (i.e., alternative explanatory factors) and make attribution of effects to the program.

Plausibility designs control for confounding using nonrandomized control groups; these designs are also called quasi-experimental (Shadish, Cook, & Cambell, 2002). The controls may be historical (i.e., retrospective), concurrent, or possibly prospective. Many different

TABLE 1 Assessment questions for evaluation of effectiveness

Assesses	Asks	Addresses
Impact	What has happened?	What were the benefits and harm?
Efficiency	What did it cost?	Are the benefits and harm worth the costs incurred?
Coverage	Who was reached?	Who received the actions (and benefits and harm)?
Process	How did it happen?	What were the factors in place and actions that occurred?
Causality	Why did it happen?	Why did the factors and actions together result in the benefits and costs?

TABLE 2 Designs for evaluation of effectiveness

Type	Question	Features
Adequacy	Did the expected changes occur?	Two (i.e., before and after) measurements required, often no control group
Plausibility	Seem to have effect above and beyond external influences?	Quasi-experimental such as historical or nonrandom comparison group, two or more measurements
Probability	Was there an effect?	Randomized comparison group, two measurements

plausibility designs exist (Habicht et al., 1999) that are cross-sectional, longitudinal such as regression discontinuity and interrupted time series (e.g., Grijalva et al., 2007), longitudinal control (e.g., Monterrosa et al., 2013), or case control (e.g., Edwards, Frongillo, Roe, & Rauschenbach, 1993). Frequency, individual, or propensity score matching can be used in the design to control for some potential confounders by balancing them across groups, and analytic methods such as multiple regression also are used to control for confounding by statistically holding potential confounders constant (Gertler, Martinez, Premand, Rawlings, & Vermeersch, 2011).

A longitudinal-control design, also called a nonequivalent groups design, is perhaps the most commonly used design, including for use to evaluate complementary feeding programs. In this design, measures are taken before and after the implementation of the program in both the program group and a nonrandomized control group. Analysis of the difference in the changes over time—called difference in differences—is made using methods for repeated measures (Gertler et al., 2011). Often sampling for this design (and others) is done using clusters, so statistical methods for accounting for the clusters such as mixed models (Goldstein, 2011), generalized estimating equations, or sandwich estimators are needed. The design might be longitudinal at the level of the cluster only or longitudinal at the level of the individual or household. The ability to make plausible causal inference (e.g., to avoid selection bias) rests with the similarity of the program and control groups at the baseline and in what happened over time other than that resulting from the program.

Probability designs use randomization at the cluster or individual level to strengthen causal inference by producing an estimate of the probability that differences between program and control groups were due to chance (Habicht et al., 1999). Randomization is important for two reasons. First, the assignment to program or control by investigators establishes the causal direction of relationships with outcomes. Second, randomization helps ensure that the two groups are equivalent on all factors other than the program assignment, whether measured or not. There are a variety of probability designs that can be used, with the most common for program evaluation being longitudinal control in which the randomized program and control groups are assessed before and after participants are exposed to the program. Difference-in-differences or related methods are used for analysis. Probability designs are used both for efficacy studies in which the exposure to the program is maximized, potentially sacrificing generalizing to real-life situations, and for effectiveness studies in which the program is implemented as it would be in real-life practice, enhancing generalizability (Habicht et al., 1999).

In addition to the design, measures for effectiveness are needed to assess impact, efficiency, coverage, process, and causality. For programs intended to improve complementary feeding, measures of complementary feeding behaviors are important to assess as primary outcomes. Indicators have been developed based on recommended feeding behaviors (World Health Organization, 2008). These indicators, which were developed for estimating prevalence, may not be ideal for evaluation; for example, using the number of food groups provided may be more informative than using the indicator of providing four or more food groups. The type of measure chosen will inform the analytical method to be used.

Measures of closely related behavioral determinants (e.g., beliefs, attitudes, norms, and intentions) and consequences (e.g., growth status) are also useful. Because programs that are intended to have impact on one domain may have impact (positive or negative) on others—for example, a program to improve complementary feeding may also improve language and motor development—it is important to use a broad set of health, nutrition, and development measures to assess impact of programs on children (Frongillo, Tofail, Hamadani, Warren, & Mehrin, 2014). Furthermore, because we want to understand what were the factors in place and actions that occurred (i.e., process) and why did the factors and actions together result in the benefits and costs (i.e., causality), measures of immediate and underlying determinants (Black et al., 2013) as well as outcomes should be made. Finally, information on costs and coverage are needed to determine cost-effectiveness and cost-benefit and to inform future program planning.

Validity of measures may be questionable when standardized or tested methods are not used and when data are self-reported. Self-reported measures are subject to recall and social desirability bias, a challenge because objective measures of feeding behaviors are difficult to obtain. One study has demonstrated the validity of self-report of exclusive breastfeeding in comparison to an objective measure (Moore et al., 2007), but similar research has not been done for complementary feeding measures. Socially desirable responsiveness can be measured so that it can be adjusted in analyses.

In both Bangladesh and Vietnam, Alive & Thrive used a probability design with randomization at the level of clusters (Menon, Rawat, & Ruel, 2013). In each country, repeated cross-sectional surveys were done with 4 years between baseline and end line. The designs were longitudinal at the cluster level rather than the individual level so that changes in children of the same age range could be examined over time. In Bangladesh and Vietnam, two program packages were compared (Menon et al., 2013). The intensive package consisted of intensive interpersonal counseling on IYCF practice, mass media, and community mobilization. The nonintensive package consisted of usual counseling along with mass media and less intense community mobilization. The differential effects of the two program packages were examined with difference-in-differences analyses using fixed-effects regression models accounting for clustering for specified outcomes including breastfeeding, complementary feeding, growth status, and child development. A measure of socially desirable responsiveness was made to quantify and adjust for potential bias. Behavioral determinants and underlying factors on mothers and households were also measured to assess for secular changes over time.

The intervention study in Mexico used a longitudinal-control plausibility design (Monterrosa et al., 2013). The intervention communities were in one state and the control communities in another adjoining state to be able to separate exposure to the radio messages. Measures were taken at baseline and end line on beliefs, attitudes, norms, intentions, and feeding behaviors. The differences between intervention and control communities were examined with difference-in-differences analyses using fixed-effects regression models accounting for clustering.

5 | CONCLUSIONS

Evaluations to provide information to learn and make decisions to improve programs are essential and should always be an integral part of programming. Evaluations should be designed based on the information needed and resources available. Investments in evaluation are worthwhile only if the data are used to improve program, which requires that the information is provided in a timely and understandable way to those who can act on it.

Evaluations need to be based on a program theory that articulates how the program will lead to intermediate and final changes. From the beginning, this program theory, often depicted as a program impact pathway or similar conceptual framework, must show a clear theory of change for how the program will achieve improvements in complementary feeding within the expected timeframe.

To be useful, evaluations need to identify what information will be useful and to collect data on the related measures for implementation functioning and effectiveness. Measures should be made along the paths in the program impact pathway and should include measures of program coverage and quality, outcomes, immediate and underlying determinants and potential confounding factors. The choice of measures should also consider timing because programs may reach and benefit children, but may not discernibly improve growth and development outcomes in a short time. Measures also need to consider the age of the child.

Although many scientists consider randomized designs to be superior for establishing causality, randomized designs sometimes are not the best choice for evaluations in practice (Habicht et al., 1999; Hébert et al., 2016). Evaluations may not be able to use randomization, particularly when programs are implemented at scale and are established as social programs, which participants are entitled to receive. Plausibility designs, innovative designs, innovative combinations of designs, and designs based on complex adaptive systems (Paina & Peters, 2012) sometimes may be best able to provide useful information for programs in many situations.

Further research is needed to develop alternative measures of complementary feeding that minimize or eliminate potential recall and social desirability bias and to develop practical evaluation designs for evaluation of large-scale country programs on complementary feeding. The process of scaling up into existing or modified systems is complex, dynamic, and unpredictable; the framework of complex adaptive systems may be helpful in reflecting these features, providing opportunities for understanding how scaling up best can occur (Paina & Peters, 2012; Pérez-Escamilla & Hall, 2016). Therefore, research is needed on how features of complex adaptive systems such as path dependence, emergent behavior, networks, feedback loops, and phase transitions can be incorporated into planning, implementation, monitoring, and evaluation of complementary feeding programs.

ACKNOWLEDGEMENTS

This paper is based on a presentation given at "First Foods: A global meeting to accelerate progress on complementary feeding in young children November 17–19, 2015, Mumbai, India." France Bégin, Laurence Grummer-Strawn, and Maria Elena D. Jefferds provided suggestions for the presentation. Sulochana Basnet, Jessica Escobar,

Shibani Kulkarni, Nazartun Monalisa, Phuong Nguyen, Mai Hoa Nguyen, and Andrea Warren provided suggestions to improve the manuscript.

SOURCE OF FUNDING

None.

CONFLICTS OF INTEREST

The author declares that he has no conflicts of interest.

CONTRIBUTIONS

EAF developed, drafted, and revised this article, and approved the final version.

REFERENCES

- Arabi, M., Frongillo, E. A., Avula, R., & Mangasaryan, N. (2012). Infant and young child feeding in 28 developing countries in relation to the Human Development Index. *Child Development, 83*, 32–45.
- Avula, R., Menon, P., Saha, K. K., Bhuiyan, M. I., Chowdhury, A. S., Siraj, S., ... Frongillo, E. A. (2013). A program impact pathway analysis identifies critical steps in the implementation and utilization of a behavior change communication intervention promoting infant and child feeding practices in Bangladesh. *Journal of Nutrition, 143*, 2029–2037.
- Baker, J., Sanghvi, T., Hajeebhoy, N., Martin, L., & Lapping, K. (2013). Using an evidence-based approach to design large-scale programs to improve infant and young child feeding. *Food and Nutrition Bulletin, 34*, S146–S155.
- Bhutta, Z. A., Das, J. K., Rizvi, A., Gaffey, M. F., Walker, N., Horton, S., ... Lancet Nutrition Interventions Review Group, Maternal and Child Nutrition Study Group (2013). Evidence-based interventions for improvement of maternal and child nutrition: What can be done and at what cost? *Lancet, 382*, 452–477.
- Black, R. E., Victora, C. G., Walker, S. P., Bhutta, Z. A., Christian, P., de Onis, M., ... the Maternal and Child Nutrition Study Group (2013). Maternal and child undernutrition and overweight in low-income and middle-income countries. *The Lancet, 382*, 427–451.
- Chen, H. T. (2015). *Practical program evaluation: Theory-driven evaluation and the integrated evaluation perspective* (2nd ed.). Thousand Oak, CA: Sage Publications.
- Edwards, D. L., Frongillo, E. A., Roe, D. A., & Rauschenbach, B. S. (1993). Home-delivered meals benefit the diabetic elderly. *Journal of the American Dietetic Association, 93*, 585–587.
- Frongillo, E. A., Tofail, F., Hamadani, J. D., Warren, A. M., & Mehrin, S. F. (2014). Measures and indicators for assessing impact of interventions integrating nutrition, health, and early childhood development. *Annals of the New York Academy of Sciences, 1308*, 68–88.
- Gertler, P. J., Martinez, S., Premand, P., Rawlings, L. B., & Vermeersch, C. M. J. (2011). *Impact evaluation in practice*. Washington, DC: Available from: http://siteresources.worldbank.org/EXTHDOFFICE/Resources/5485726-1295455628620/Impact_Evaluation_in_Practice.pdfThe World Bank.
- Goldstein, H. (2011). *Multilevel statistical modelling* (4th ed.). New York: John Wiley and Sons.
- Grijalva, C. G., Nuorti, J. P., Arbogast, P. G., Martin, S. W., Edwards, K. M., & Griffin, M. R. (2007). Decline in pneumonia admissions after routine childhood immunisation with pneumococcal conjugate vaccine in the USA: A time-series analysis. *Lancet, 369*(9568), 1179–1186.
- Habicht, J. P., Victora, C. G., & Vaughan, J. P. (1999). Evaluation designs for adequacy, plausibility and probability of public health programme performance and impact. *International Journal of Epidemiology, 28*, 10–18.

- Hébert, J. R., Frongillo, E. A., Adams, S. A., Turner-McGrievy, G. M., Hurley, T. G., Miller, D. R., & Ockene, I. S. (2016). Perspective: Randomized controlled trials are not a panacea for diet-related research. *Advances in Nutrition*, 7, 423–432.
- Kim SS, Habicht JP, Menon P, Stoltzfus RJ. (2011). How do programs work to improve child nutrition? Program impact pathways of three nongovernmental organization intervention projects in the Peruvian highlands. International Food Policy Research Institute Discussion Paper 01105. Washington, DC: IFPRI. Available from: <http://www.ifpri.org/sites/default/files/publications/ifpridp01105.pdf>
- Menon, P., Rawat, R., & Ruel, M. (2013). Bringing rigor to evaluations of large-scale programs to improve infant and young child feeding and nutrition: The evaluation designs for the Alive & Thrive initiative. *Food and Nutrition Bulletin*, 34, S195–S211.
- Monterrosa, E. C., Pelto, G. H., Frongillo, E. A., & Rasmussen, K. M. (2012). Constructing maternal knowledge frameworks: How mothers conceptualize complementary feeding. *Appetite*, 59, 377–384.
- Monterrosa, E. C., Frongillo, E. A., González de Cossío, T., Bonvecchio, A., Villanueva, M. A., Thrasher, J. F., & Rivera, J. A. (2013). Scripted messages delivered by nurses and radio changed beliefs, attitudes, intentions, and behaviors regarding infant and young child feeding in Mexico. *Journal of Nutrition*, 143, 915–922.
- Moore, S. E., Prentice, A. M., Coward, W. A., Wright, A., Frongillo, E. A., Fulford, A. J. C., ... Kabir, I. (2007). Use of stable isotope techniques to validate infant feeding practices reported by Bangladeshi women receiving breastfeeding counseling. *American Journal of Clinical Nutrition*, 85, 1075–1082.
- Nguyen, P. H., Menon, P., Keithly, S. C., Kim, S. S., Hajebehoy, N., Tran, L. M., ... Rawat, R. (2014). Program impact pathway analysis of a social franchise model shows potential to improve infant and young child feeding practices in Vietnam. *Journal of Nutrition*, 144, 1627–1636.
- Nguyen, P. H., Kim, S. S., Keithly, S. C., Hajebehoy, N., Tran, L. M., Ruel, M. T., ... Menon, P. (2014). Incorporating elements of social franchising in government health services improves the quality of infant and young child feeding counselling services at commune health centres in Vietnam. *Health Policy and Planning*, 29, 1008–1020.
- Nguyen, P. H., Kim, S. S., Nguyen, T., Tran, L. M., Hajebehoy, N., Frongillo, E. A., ... Menon, P. (2016). Supply- and demand-side factors influencing utilization of infant and young child feeding counselling services in Viet Nam. *PLoS One*. doi:10.1371/journal.pone.0151358
- Paina, L., & Peters, D. H. (2012). Understanding pathways for scaling up health services through the lens of complex adaptive systems. *Health Policy and Planning*, 27, 365–373.
- Pérez-Escamilla, R., & Hall, M. V. (2016). Scaling up breastfeeding programmes in a complex adaptive world. *Maternal & Child Nutrition*, 12, 375–380.
- Piwoz, E., Baker, J., & Frongillo, E. A. (2013). Documenting large-scale programs to improve infant and young child feeding is key to facilitating progress in child nutrition. *Food and Nutrition Bulletin*, 34, S143–S145.
- Rawat, R., Nguyen, P. H., Ali, D., Saha, K., Alayon, S., Kim, S. S., ... Menon, P. (2013). Learning how programs achieve their impact: Embedding theory-driven process evaluation and other program learning mechanisms in Alive & Thrive. *Food and Nutrition Bulletin*, 34, S212–S225.
- Rossi, P. H., & Freeman, H. E. (1989). *Evaluation: A systematic approach* (4th ed.). Newbury Park: Sage Publications.
- Saunders, R. P. (2016). *Implementation monitoring & process evaluation*. Thousand Oaks, CA: Sage Publications.
- Shadish, W. R., Cook, T. D., & Campbell, D. T. (2002). *Experimental and quasi-experimental designs for generalized causal inference*. Boston: Houghton Mifflin.
- UNICEF. First Foods: A global meeting to accelerate progress on complementary feeding in young children, November 17–19, 2015, Mumbai. Summary of global presentations and recommendations. New York: UNICEF, 2016.
- World Health Organization (2008). *Indicators for assessing infant and young child feeding practices. Part 2: Measurements*. Geneva: World Health Organization. Available from: http://whqlibdoc.who.int/publications/2010/9789241599290_eng.pdf

How to cite this article: Frongillo EA. Evaluation of programs to improve complementary feeding in infants and young children. *Matern Child Nutr*. 2017;13(S2):e12436. <https://doi.org/10.1111/mcn.12436>