Graduate School ETD Form 9
(01/07)

PURDUE UNIVERSITY
GRADUATE SCHOOL
Thesis Acceptance

This is to certify that the thesis prepared

By Nilupa S. Gunaratna

Entitled

Evaluating the Nutritional Impact of Maize Varieties Genetically Improved for Protein Quality

Complies with University regulations and meets the standards of the Graduate School for originality and quality

For the degree of Doctor of Philosophy

Final examining committee members

George P. McCabe, Chair

Kevin V. Pixley

Mary Ellen Bock

Rebecca W. Doerge

Penelope Nestel

Approved by Major Professor(s): George P. McCabe

Approved by Head of Graduate Program: Jun Xie

Date of Graduate Program Head's Approval: April 26, 2007

EVALUATING THE NUTRITIONAL IMPACT OF MAIZE VARIETIES GENETICALLY IMPROVED FOR PROTEIN QUALITY

A Dissertation

Submitted to the Faculty

of

Purdue University

by

Nilupa S. Gunaratna

In Partial Fulfillment of the

Requirements for the Degree

of

Doctor of Philosophy

May 2007

Purdue University

West Lafayette, Indiana

This dissertation is dedicated to my parents:

P. Chandrani and W. Gamini Gunaratna

ACKNOWLEDGMENTS

I would like to recognize my advisor, George McCabe, for all his guidance, encouragement, and patient support during my graduate study. Whether in research, consulting, collaboration, or other work that will have an impact on others, I am glad to have had the opportunity to learn from his example.

I am also grateful to the members of my Ph.D. Advisory Committee for the role they have played in my education. I would like to thank Mary Ellen Bock for her guidance in my graduate studies and in my involvement in the American Statistical Association (ASA). I am thankful to Rebecca Doerge, for first encouraging me to pursue graduate study in statistics and for her constant support and belief in me during that time. I am thankful to Penelope Nestel (University of Southampton School of Medicine, UK) and Kevin Pixley (International Maize and Wheat Improvement Centre (CIMMYT), Mexico) for providing their advice and perspectives on this research, patiently answering all my questions, and helping to secure funding from the HarvestPlus Challenge Program to support this work. This research was supported by HarvestPlus Agreement # 8026.

Many researchers also offered their advice and perspectives. I would like to recognize Hugo De Groote (International Maize and Wheat Improvement Centre (CIMMYT), Kenya) for all his mentorship and support and especially for his constant friendship. I would like to thank George Beaton (Emeritus, University of Toronto,

Canada), Ricardo Bressani (Universidad del Valle de Guatemala, Guatemala), Shibani Ghosh (International Nutrition Foundation), Paulo Evaristo O. Guimarães (Embrapa Milho e Sorgo, Brasil), Bruce Hamaker (Purdue University), Brian Larkins (University of Arizona), Peter Pellett (University of Massachusetts Amherst), Alberto Pradilla (Universidad del Valle de Cali, Colombia), and Nevin Scrimshaw (Emeritus, Massachusetts Institute of Technology) for their contributions. I am further grateful to the many researchers around the world who provided information on past and current nutritional work on *o2* maize and QPM.

I would like to thank Abenaa Akuamoa-Boateng (Ghana Health Service, Ghana), Eveling Ortega Alemán (Universidad Nacional Autónoma de Nicaragua, Nicaragua), and Maria Cristina Dias Paes (Embrapa Milho e Sorgo, Brasil) for sharing information and data on QPM nutritional studies conducted in their countries. My thanks also go to Wayne Haag (Sasakawa Global 2000) and Helena Pachón (International Center for Tropical Agriculture (CIAT), Colombia) for their help in gathering information and data on QPM nutritional studies in Africa and Latin America.

I am thankful to the members of Purdue University's Statistical Consulting Service, Technical Assistance Program, Statistics in the Community (STATCOM), and the STATCOM Network for all the positive experiences I had in those programs. In particular, I would like to thank Regina Becker for her guidance and advice in all those efforts. I would like to express my appreciation to all the Department of Statistics staff, especially Cheryl Crabill, Doug Crabill, Teena Erwin, Norma Lucas, Becca Mitchell, Mary Roe, My Truong, and Darlene Wayman for their constant willingness to help.

Many thanks also go to my fellow students and friends in the Department of Statistics for providing a very enjoyable and supportive environment to learn and work.

Special thanks go to my fiancé, Jason A. Burbank, who has been my friend from childhood and who has supported me and pushed me every step of the way.

Finally, I would like to thank my family and especially my parents, P. Chandrani and W. Gamini Gunaratna, and my sister, Melanie A. Gunaratna, who have always been there for me and who have sacrificed much in their lives to give me the opportunities that I have enjoyed in mine.

TABLE OF CONTENTS

Page

LIST OF TABLES

Table Page

LIST OF FIGURES

Figure Page

ABSTRACT

Gunaratna, Nilupa S. Ph.D., Purdue University, May 2007. Evaluating the Nutritional Impact of Maize Varieties Genetically Improved for Protein Quality. Major Professor: George P. McCabe.

Biofortification, or the genetic improvement of the nutritional quality of food crops, is a promising technology to combat childhood undernutrition in developing countries. Significant progress has been made to develop maize varieties with improved protein quality, collectively known as quality protein maize (QPM). However, debate still continues over whether this agricultural technology will have a significant public health impact. A four-level framework was proposed to evaluate the nutritional impact of QPM and biofortified crops. Using the results of community-level efficacy and effectiveness studies on QPM that have been conducted in several countries in Sub-Saharan Africa, Latin America, and Asia, a meta-analysis was performed to assess the effect of QPM on child growth. The results indicated that consumption of QPM compared with conventional maize varieties leads to an 8% (95% CI: 4-12%) increase in the rate of growth in height and a 9% (95% CI: 4-12%) increase in the rate of growth in weight in infants and toddlers with mild to moderate undernutrition for whom maize is a significant part of the diet. These results were not sensitive to alternative methods of determining the summary effect size and its statistical significance. However, the studies used to

derive these estimates had several methodological limitations, and further community-level nutritional studies were recommended to provide stronger evidence on the efficacy and effectiveness of QPM. Simulation was used to study the potential impact of QPM on nutrient adequacy at the population level. Methods were developed to incorporate potential impact pathways for biofortified crops, allowing quantitative discrimination of scenarios of high and low impact. These methods were then used to examine factors that could modify the impact of a new crop. Simulation results indicate that impact depends on adoption and production patterns, composition of the total diet, variation in food consumption patterns, self-sufficiency in and sources of maize, disease, and other factors, as well as seasonal effects on the above. In particular, they highlight the importance of monitoring total diet and morbidity in targeting and impact assessment. These factors should be considered in the planning and implementation of biofortification programs.

CHAPTER 1. INTRODUCTION

1.1. Quality Protein Maize (QPM)

Maize is a staple food for millions of people in Latin America, Sub-Saharan Africa, and parts of Asia. Like most cereal grains, maize grain is low in both protein quantity and quality and is particularly deficient in lysine and tryptophan, two amino acids that are essential to the diets of humans and monogastric animals (FAO 1992). Efforts to improve the protein quality of maize date back to the 1950s, starting with the evaluation of fortification with essential amino acids (Bressani et al. 1958; Bressani et al. 1963; Scrimshaw et al. 1958) and the identification of genetic variability for lysine and tryptophan content (Bressani et al. 1953; Bressani et al. 1960).

In the early 1960s, E.T. Mertz and colleagues at Purdue University, USA, discovered that the natural *opaque-2* (*o2*) mutation changed the protein composition of the maize endosperm, nearly doubling its lysine and tryptophan content (Mertz et al. 1964). As a result, *o2* maize grain had improved protein quality, while its protein quantity remained the same. A favorable change in the ratio of two other amino acids, leucine and isoleucine, also liberated more tryptophan for the biosynthesis of niacin, another essential nutrient (Vasal 2000).

Plant breeding programs quickly began using the *o2* mutation to develop maize varieties with improved protein quality. It was soon discovered, however, that the *o2*

mutation had several undesirable pleiotropic effects (secondary characteristics) including decreased yield and increased susceptibility to diseases and storage pests. Changes in the kernel characteristics of *o2* maize also made it less appealing to both producers and consumers (Vasal 2000). These drawbacks spurred years of breeding efforts to develop maize varieties that retained the *o2* mutation and the quality protein trait but lacked the accompanying unfavorable characteristics (Krivanek et al. 2007). To differentiate them from the earlier *o2* maize varieties and from "conventional maize" (CM) varieties lacking the quality protein trait altogether, these new varieties were referred to as "quality protein maize" (QPM). All QPM varieties and the earlier *o2* maize varieties were developed through conventional plant breeding and are not genetically modified. No systematic differences have been reported between CM and either QPM or *o2* maize with respect to other nutrients including carbohydrates, fats, fiber, micronutrients, energy, or amino acids other than lysine, tryptophan, and leucine (Bressani 1991; FAO 1992). The goal of this dissertation is to evaluate evidence on the nutritional impact of QPM on humans.

1.2. Evaluating Nutritional Impact

The *o2* mutation was identified at a time when protein deficiency was considered the world's main nutrition problem (Allen 2003). The focus of nutritional interventions subsequently changed, first to the alleviation of energy deficiency in the 1970s and later to the alleviation of micronutrient deficiencies from the 1980s to the present day. Consequently, breeding for improved protein quality in staple cereals grains was largely dropped as a research priority, though some institutions – most notably the International Maize and Wheat Improvement Center (CIMMYT) in Mexico, the University of

KwaZulu-Natal in South Africa, and Crow's Hybrid Seed Company in the United States – continued their QPM breeding programs (Vasal 2000).

In recent years, there has been a renewal of interest in improving the protein quality of cereal crops. Dr. Surinder K. Vasal and Dr. Evangelina Villegas from CIMMYT were awarded the 2000 World Food Prize for their work on developing QPM (World Food Prize Foundation 2007). Improved protein quality is currently an objective in several national and international breeding programs focused on maize as well as other cereals. However, debate persists over whether QPM or other cereals improved for protein quality will have a significant public health impact.

There are two key motivations for evaluating the potential impact of QPM. First, given limited resources, decisions must be made on the priority of improved protein quality as an objective in maize breeding programs. Second, there is growing interest more generally in biofortification, or the genetic improvement of the nutritional quality of food crops, with a focus on available micronutrient levels. Evaluating the nutritional impact of QPM could inform decision-making on QPM breeding and dissemination, while lessons learned from over 40 years of QPM research could also inform future biofortification efforts.

The nutritional impact of QPM, or of any biofortified crop, should be demonstrated at multiple levels (King 2002). In this dissertation, a four-level framework that considers bioavailability, efficacy, effectiveness, and impact in a broader societal context is used to evaluate the existing evidence on the nutritional impact of QPM. This framework could apply generally to other biofortified crops.

At the first level, the change in the amino acid profile of QPM grain must result in increased bioavailability of protein from QPM grain compared with CM grain when consumed by target individuals. Bioavailability refers to the fraction of an ingested nutrient that is utilized for normal physiological functions or storage (King 2002). This increased utilization must then result in improved outcomes among target individuals who consume QPM instead of CM. An outcome is an indicator of a social or health condition that an intervention, program, or technology is accountable for improving. In the case of QPM, the primary target population is children under five years of age, and the primary outcomes of interest are anthropometric measurements of growth.

The effect of QPM on any outcome should be evaluated using efficacy and then effectiveness studies. Efficacy is "the extent to which a specific intervention, procedure, regimen, or service produces a beneficial effect under ideal conditions", while effectiveness is "the extent to which a specific intervention, procedure, regimen, or service, when deployed in the field, does what it is intended to do for a defined population" (Last 1988). As efficacy studies are conducted under well-controlled conditions, they primarily address biological factors relating to the effect of an intervention. Effectiveness studies are likely to involve less control over delivery of the intervention to subjects and subjects' compliance with the intervention, and external confounding factors, both biological and behavioral, are more likely to modify the effect of an intervention in an effectiveness study.

Finally, the impact of QPM should be evaluated in a broader context that, in addition to its nutritional and health effects, also investigates the agricultural, societal, environmental, and economic effects of the technology. A comprehensive impact

assessment would require data on specific areas and numbers of people who suffer from nutrient inadequacies due to their consumption of maize as a staple food and whose nutrient intakes would significantly increase with adoption and consumption of QPM. This assessment would therefore address two important and distinct questions. First, data on the diets of target populations should indicate risk of inadequate intakes of protein, lysine, or tryptophan. The improved nutritional quality of QPM, or of any biofortified crop, is expected to have an impact only by alleviating nutrient inadequacy in a target population. If there is little risk of inadequate nutrient intakes prior to its release, the technology would be expected to have little impact.

This is relevant in the case of QPM as many nutritionists argue that other foods in maize-based diets compensate for the amino acid deficiencies in maize (Rahmanifar and Hamaker 1999). This argument was supported by results from the Nutrition Collaborative Research Support Program (CRSP), which studied nutrient adequacy in selected communities in Egypt, Mexico, and Kenya in the 1980s and concluded that protein intake was unlikely to be a primary limiting factor for the growth and development of young children and little benefit could be expected from increasing the intake of limiting amino acids such as lysine (Beaton et al. 1992). Rahmanifar and Hamaker (1999) have however argued that while this conclusion may hold for populations that are “relatively ‘better-off’”, it may not generalize to all populations, particularly those with limited access to high protein quality foods, and that there exist poor, high maize-consuming populations at risk of protein inadequacy that would benefit from QPM.

Rahmanifar and Hamaker (1999) also pointed out that the decrease in emphasis on protein deficiency as a nutritional problem was accompanied by multiple revisions of human protein and amino acid requirements over the last 40 years. Using recent protein and amino acid requirements to analyze regional and country-specific food availability data, Young and Pellett (1990) concluded that diets in a number of developing countries may be marginal for both lysine and utilizable protein. They recommended that improving dietary protein quality should remain a consideration in the design and implementation of food, nutrition, and agricultural programs and policies, particularly in countries where diets mainly depend on cereals. Further work by Pellett (1996) using regional and country-specific food availability data and dietary surveys from India and Pakistan also suggested that lysine may be lacking in many parts of the world where diets are heavily based on cereals. The United States Institute of Medicine (IOM) has recently said that individuals who consume little or no animal protein would be unlikely to obtain recommended amounts of lysine even if they obtain recommended amounts of total protein unless their diets were usually high in legumes, and even then, lysine intakes could be marginal (IOM 2005). IOM therefore recommends assessing lysine intakes in addition to total protein intakes for individuals who consume proteins with low levels of lysine.

Estimation of protein and amino acid requirements is an ongoing area of research. As protein requirements are related energy requirements, Millward and Jackson (2003) used calculated reference ratios of protein energy to total energy to examine the implications of recent and proposed nutritional requirements in determining the adequacy of protein intakes in developed and developing countries. They concluded that if a diet

were potentially limiting in protein, deficiency was most likely in large, elderly sedentary women, followed by adolescent girls, and least likely in moderately active young children. Calculations indicated significant risk of deficiency with respect to both protein quantity and quality in developed and particularly developing countries. These results would support efforts to increase the availability of high quality protein for at-risk populations. However, the authors also proposed reevaluation of recent protein and amino acid requirements and, in particular, examination of the validity of assumptions used to determine those requirements.

The second question that an impact assessment would answer is whether characteristics of a target region and population would allow sufficiently high adoption and consumption of QPM to have a significant impact on the adequacy of nutrient intakes in that population. QPM, and all biofortified crops, address a nutritional problem by introducing a new agricultural technology. The links between agriculture and nutrition that would allow the release of an agricultural technology to have an impact on the nutrition of a target population must therefore be described. The steps along these potential pathways of impact should be evaluated for factors that could modify the impact of the technology. These could include factors influencing QPM adoption, acceptability of the grain for food preparation and consumption, nutrient losses during processing or preparation, prevalence of disease in the population, and seasonal patterns in food consumption or disease.

1.3. Bioavailability

Several studies have been conducted to evaluate the bioavailability of protein in QPM and *o2* maize. The studies, which began soon after the identification of the *o2* mutation, utilized the nitrogen balance technique in both children and adults to compare protein quality of *o2* maize or QPM grain to CM grain or a reference protein. In nitrogen balance studies, nitrogen intake and losses from the body through the urine and feces are used as a proxy for the corresponding intake and losses of protein, and various indicators of bioavailability are calculated from these values (FAO 1991).

Studies conducted in Guatemala (Bressani et al. 1969; Luna-Jaspe G. et al. 1971) and Peru (Graham et al. 1980; Graham et al. 1989) on young children recovering from protein-energy malnutrition found that nitrogen retention (the proportion of nitrogen intake that is not lost in the urine or feces) in children fed *o2* maize was higher than in children fed CM but lower than in children fed a reference protein (usually casein, a milk protein). This implies that a greater proportion of protein is available from *o2* maize compared with CM for utilization by the body for maintenance and growth.

It was further calculated that a young child could reasonably eat sufficient *o2* maize to achieve nitrogen equilibrium, in which nitrogen intake equals losses, and thereby meet their protein requirements, but they could not do so by consuming CM (Bressani et al. 1969). Similar studies with children recovering from malnutrition were conducted in Colombia to determine the biological value of protein from *o2* maize (Pradilla et al. 1973). Biological value is the proportion of absorbed nitrogen that is retained by the body for maintenance or growth, i.e., of the nitrogen that is not lost in the feces, it is the proportion that is further retained and not excreted in the urine (FAO

1991). These studies found that the biological value of protein in *o2* maize grain was higher than that of CM and close to that of casein (Pradilla et al. 1973), indicating that compared to protein from CM, a greater proportion of protein absorbed from *o2* maize is available for utilization by the body for normal biological processes.

Likewise, studies conducted on adults also found higher nitrogen balance (the difference between nitrogen intake and nitrogen losses) with *o2* maize than with CM (Clark et al. 1977; Kies and Fox 1972), although these results were not always statistically significant (Clark et al. 1977). This implies that there is greater utilization of protein from *o2* maize than from CM to meet adult protein requirements more adequately. Studying adults, Young et al. (1971) further concluded that the biological value of *o2* maize was high and comparable to most sources of animal protein. Large sample sizes are often not feasible for these studies, and the early studies on *o2* maize each typically involved eight or fewer subjects. Overall, however, these studies concluded that the improved quality protein in *o2* maize and QPM is more available than the protein in CM for utilization in humans, and nitrogen balance when consuming *o2* maize or QPM was particularly higher at lower levels of total protein intake (R. Bressani, personal communication).

1.4. Objectives

While the greater bioavailability of protein from QPM is generally accepted and supported by the peer-reviewed literature, there has been more debate over the efficacy, effectiveness, and impact of QPM at the community level. Although some studies have

been conducted to address these issues, very little relevant work has been published, and the question remains whether QPM will have a significant public health impact.

There are three underlying questions that determine the impact of this technology. First, there must be relevant nutrient deficiencies in populations where maize is a staple food, i.e., there must a nutritional need that could be met by QPM. Second, consumption of QPM must result in improved nutritional or health outcomes in target individuals. This is primarily a biological question. Finally, characteristics of target populations must allow sufficiently high adoption and consumption to have impact at the population level. This is primarily a behavioral question.

In Section 1.2, a four-level framework was proposed to evaluate the nutritional impact of crop varieties genetically improved for nutritional quality. In this dissertation, this framework is applied to evaluate evidence on the nutritional impact of QPM. Existing evidence on the bioavailability of protein from QPM grain was evaluated in Section 1.3. In Chapter 2, a meta-analysis of community-level efficacy and effectiveness studies is conducted to assess the effect of QPM on child growth. Recommendations are also made for the design and analysis of community-level studies that would provide more evidence on the nutritional impact of QPM. In Chapter 3, simulation studies are used to investigate the potential impact of QPM on nutrient adequacy at the population level. Diets in Guatemala and Kenya, countries targeted for QPM dissemination, were simulated and used to study the effect of factors such as dietary patterns, adoption and production decisions, health characteristics, and seasonality on nutrient adequacy.

This systematic approach will aid decision-making about QPM breeding and dissemination and identify needs for future research. The objectives of this dissertation are:

- To describe a systematic approach to evaluate the nutritional impact of crop varieties genetically improved for nutritional quality
- To apply this approach to evaluate the nutritional impact of QPM
- To review and evaluate community-level nutritional studies on QPM
- To conduct a meta-analysis to assess the effect of QPM on the growth of young children
- To describe conceptual frameworks and make recommendations for the design and analysis of efficacy and effectiveness studies of QPM
- To use simulation to study the potential impact of QPM at the population level
- To describe the potential modifying role of consumption patterns, adoption and production decisions, disease, and seasonality on the impact of QPM at the population level.

CHAPTER 2. EVALUATION AT THE COMMUNITY LEVEL

2.1. Introduction

Several community-level nutritional studies have been conducted in Latin America, Africa, and Asia to evaluate the efficacy and effectiveness of QPM adoption and consumption on humans. These studies primarily compared the effects of QPM and CM on the growth in height and weight of children less than five years old. In this chapter, these studies were reviewed and their results were used to conduct meta-analyses to assess the effect of QPM on the growth of young children for whom maize is a major part of the diet.

Meta-analysis is a statistical method to integrate the results of studies addressing the same research question (Cooper 1998; Cooper and Hedges 1994; Glass, McGaw, and Smith 1981; Hedges and Olkin 1985; Hunter and Schmidt 2004; Rosenthal 1991). In a meta-analysis, each included study serves as a single independent observation. For each study, an effect size is calculated. An effect size is a statistic that quantifies the treatment effect in each study in a way that is interpretable and comparable across studies. A summary effect size is then calculated to quantify the overall treatment effect across studies.

In this chapter, separate meta-analyses were conducted on the effect of QPM on child height and weight. For each meta-analysis, an effect size was calculated for each

included study to summarize the effect of QPM compared to the effect of CM on children's growth in height or weight in that study. The summary effect size then quantified the overall effect of QPM compared to CM on children's growth in height or weight across such studies. Evaluation of the design of these studies also allowed formulation of recommendations for future studies to evaluate the efficacy and effectiveness of QPM. Together, these studies provide evidence to support decision-making about the development, targeting, and dissemination of QPM.

2.2. Methods

The research conducted in this chapter was approved by the Purdue University Committee on the Use of Human Research Subjects. All data used in these analyses were summary statistics calculated from de-identified existing raw data sets or obtained from reports, conference proceedings, or the researchers who conducted an included study.

2.2.1. Identifying Relevant Studies

Several QPM nutritional studies have been conducted since the 1970s in communities in Latin America, Africa, and Asia where maize is a major part of the diet. Results from some of these studies have been published in technical reports or dissertations, but none of these studies have been reported in the peer-reviewed literature. The studies included in these meta-analyses were therefore identified by researchers at CIMMYT or by colleagues identified by CIMMYT researchers. Information on human

nutritional studies using QPM was also requested in presentations made at scientific meetings and through an announcement published in the Plant Breeding News (PBN-L) electronic newsletter, sponsored by the Food and Agriculture Organization (FAO) of the United Nations and Cornell University, on February 22 and 28, 2005. Researchers who had authored articles on other nutritional aspects of QPM were also contacted by e-mail for information. Information was further requested at the following website, which provided an online bibliography of QPM-related articles: http://www.stat.purdue.edu/~gunaratn/QPM/. Internet searches on PubMed (http://www.pubmed.gov) and using the Google search engine (http://www.google.com) were conducted using the keywords "QPM", "quality protein maize", "maize" and "protein quality", "opaque-2", "maíz de alta calidad proteínica" [Spanish], "maíces de alta calidad proteínica" [Spanish], "maíz" and "calidad proteínica" [Spanish] , and "opaco-2" [Spanish and Portuguese].

Included studies had at least two treatment groups, one of which received *o2* maize or QPM and the other of which received CM. Subjects were children whose heights or weights were recorded at the beginning (baseline) and at the end of the study. Information was available on the study design either through the researcher in charge of the study or through a written report or presentation on the study. Data were also available on the number of subjects and on the mean changes in height or weight in each group, allowing calculation of the effect size described below. Relevant research was considered from any country in any language. Studies were excluded if they lacked a QPM/*o2* maize or CM treatment group, information on the study design could not be

obtained from an original report or researcher, or insufficient data were available to calculate the effect size described below.

2.2.2. Calculating Effect Size

For each included study, effect sizes were calculated separately for two measures: height and weight. For each measure, a study could be included in the meta-analysis if the following could be determined from available information: number of subjects in the QPM or *o2* maize group, number of subjects in the CM group, average change in the measure in the QPM or *o2* maize group between baseline and completion of the study, and average change in the measure in the CM group between baseline and completion of the study.

The following effect size, $\hat{\theta}_i$, was proposed to quantify the effect of QPM relative to the effect of CM in a given study *i*:

$$\hat{\theta}_i = \frac{b_{QPM}}{b_{CM}},$$

where b_{QPM} is the average change in the QPM or *o2* maize group and b_{CM} is the average change in the CM group. The variance of $\hat{\theta}_i$ was estimated using the delta method (Casella and Berger 2002). In general,

$$\operatorname{var}\left(g\left(\hat{b}_1, \hat{b}_2\right)\right) \approx \left(\frac{\partial g}{\partial \hat{b}_1}\right)^2 \operatorname{var}\left(\hat{b}_1\right) + 2\frac{\partial g}{\partial \hat{b}_1}\frac{\partial g}{\partial \hat{b}_2}\operatorname{cov}\left(\hat{b}_1, \hat{b}_2\right) + \left(\frac{\partial g}{\partial \hat{b}_2}\right)^2 \operatorname{var}\left(\hat{b}_2\right).$$

Therefore,

$$\operatorname{var}\left(\hat{\theta}_i\right) \approx \left(\frac{1}{b_{CM}}\right)^2 \operatorname{var}\left(b_{QPM}\right) + 2\left(\frac{1}{b_{CM}}\right)\left(\frac{-b_{QPM}}{b_{CM}^2}\right)\operatorname{cov}\left(b_{QPM}, b_{CM}\right) + \left(\frac{-b_{QPM}}{b_{CM}^2}\right)^2 \operatorname{var}\left(b_{CM}\right).$$

If b_{QPM} and b_{CM} are estimated from independent groups of children, then $\operatorname{cov}\left(b_{QPM}, b_{CM}\right) = 0$. $\operatorname{Var}\left(\hat{\theta}_i\right)$ was calculated for each study for which sufficient information was available.

2.2.3. Calculating Summary Effect Size

The summary effect size, $\hat{\theta}$, which quantifies the overall effect of QPM relative to CM, was calculated as:

$$\hat{\theta} = \sum_i w_i \hat{\theta}_i ,$$

where w_i is the weight corresponding to study i and $\sum_i w_i = 1$. Then,

$$\operatorname{var}\left(\hat{\theta}\right) = \sum_i w_i^2 \operatorname{var}\left(\hat{\theta}_i\right),$$

as the $\hat{\theta}_i$ were derived from independent studies. An asymptotic $100(1-\alpha)\%$ confidence interval (CI) for $\hat{\theta}$ was calculated by:

$$\hat{\theta} \pm z_{100\left(1-\frac{\alpha}{2}\right)} SE\left(\hat{\theta}\right),$$

where $z_{100\left(1-\frac{\alpha}{2}\right)}$ is the $100\left(1-\frac{\alpha}{2}\right)$th percentile of the standard normal distribution and $SE\left(\hat{\theta}\right) = \sqrt{\operatorname{var}\left(\hat{\theta}\right)}$. To calculate a CI for $\hat{\theta}$ in this way, $\operatorname{var}\left(\hat{\theta}_i\right)$ must be known for all i.

Alternatively, a $100(1-\alpha)\%$ bootstrap percentile confidence interval can be calculated for $\hat{\theta}$ as follows (Efron and Tibshirani 1993). If *n* is the number of studies included in the meta-analysis, then a sample of size *n* of included studies is drawn with replacement from the set of included studies. A summary effect size is calculated for this new data set, called a resample. This process is repeated many times, and the summary effect sizes generated for each independent resample form a bootstrap distribution that approximates the sampling distribution of $\hat{\theta}$. The $100\left(\frac{\alpha}{2}\right)$th and $100\left(1-\frac{\alpha}{2}\right)$th percentiles of the bootstrap distribution of a statistic form a $100(1-\alpha)\%$ bootstrap percentile confidence interval for that statistic.

In this analysis, three sets of w_i were considered. The optimal choice for the w_i, i.e., the set of w_i that minimizes the variance of $\hat{\theta}$, is inversely proportional to the variances of the respective $\hat{\theta}_i$. This choice of weights can only be used when $\text{var}\left(\hat{\theta}_i\right)$ is available for all included studies. Another reasonable choice for the w_i is proportional to the sample sizes of the respective studies. Studies with larger sample sizes are expected to provide more precise estimates of the relative effect of QPM to CM. This choice would therefore give larger weight to larger studies, as they provide more information than small studies about the overall effect size. Finally, equal weights were also considered for all included studies.

Summary effect sizes and 95% bootstrap percentile confidence intervals based on 5000 resamples were calculated for height and weight using these three sets of weights. For the optimal set of weights, it was possible to include only those studies for which

$\text{var}\left(\hat{\theta}_i\right)$ could be calculated. Sensitivity of the summary effect size to the contributions of individual studies was explored by calculating summary effect sizes and bootstrap percentile confidence intervals while excluding one study at a time from the analysis.

Using the optimal weights, asymptotic confidence intervals for the summary effect size were also calculated. While the bootstrap percentile confidence intervals are determined by variation in effect sizes among studies, asymptotic confidence intervals are based on variation in effect sizes due to sampling error alone (i.e., within-study variation). Heterogeneity among studies (i.e., between-study variation) was assessed using a statistical test based on the *Q* statistic:

$$Q = \sum_i w_i \left(\hat{\theta}_i - \hat{\theta}\right)^2 ,$$

where $w_i = \dfrac{1}{\text{var}\left(\hat{\theta}_i\right)}$ (Cochran 1954). Under a null hypothesis of no heterogeneity among studies,

$$Q \sim \chi^2_{k-1},$$

where *k* is the number of included studies. Between-study variance was estimated as:

$$\hat{\tau}^2 = \begin{cases} \dfrac{Q-(k-1)}{\sum_i w_i - \left(\sum_i w_i^2 \Big/ \sum_i w_i\right)} & \text{if } Q > k-1 \\ 0 & \text{otherwise} \end{cases}$$

where *Q*, *k*, and w_i were defined as above (DerSimonian and Laird 1986). The optimal weight for study *i* was then calculated as inversely proportional to $\left(\hat{\tau}^2 + \text{var}\left(\hat{\theta}_i\right)\right)$. An

asymptotic confidence interval calculated using this set of weights takes into account both between-study and within-study variation. The corresponding asymptotic confidence intervals were calculated for the analyses of height and weight.

2.3. Community-level Nutritional Studies

2.3.1. Ethiopia

A study was conducted in April 2002-October 2003 in the Eastern Wollega Zone of Ethiopia by the Ethiopian Health and Nutrition Research Institute (Akalu 2005). The study area was divided into four regions. Households in two of the regions were given QPM seed, while households in the other two regions were given CM seed. Farmers in these households were expected to grow the seed and use the harvested grain for household consumption over the following year. In total, 160 children participated in the study, with 80 children in QPM households and 80 in CM households. No data were available to evaluate any systematic differences between the QPM and CM regions.

The heights and weights of the children, most of whom were less than 24 months old, were monitored on a quarterly basis for one year. There was no monitoring or analysis of children's intakes of maize or other foods. This study may not have been blinded, as the farmers grew the maize varieties from which they harvested the grain. It is not known whether the study was approved by an institutional review board (IRB) or whether informed consent was obtained from members of participating households. In the following analysis, this study will be referred to as the "Ethiopia" study.

At baseline, children in the study showed signs of moderate malnutrition, as indicated by low average height-for-age and weight-for-age Z-scores (HAZ and WAZ scores, respectively) (Gibson 2005). HAZ and WAZ scores are calculated by standardizing a child's height or weight, respectively, with respect to the corresponding distribution of a reference population of the same age. Mean HAZ and WAZ scores in the QPM and CM groups at baseline were around -1, indicating that the subjects were on average one standard deviation below the mean of the corresponding reference population with respect to height-for-age and weight-for-age. Mean HAZ and WAZ scores in the QPM and CM groups over time are given in Figure 2.1. Visit 0 denotes the baseline measurement, and visits 1-4 denote the subsequent quarterly measurements. These means were calculated using all data ("All Data") and using data only from the children with no missing data ("Subset"). For HAZ, there were 51 children with complete data in each of the two groups. For WAZ, there were 53 children in the QPM group and 52 children in the CM group with complete data.

The patterns of mean HAZ and WAZ scores over time differed between the full data set and the subset. This is likely a consequence of the fact that only about 64% of the children in this study were measured at each visit during the study. The remaining 36% are represented in some but not all the means in the full data set.

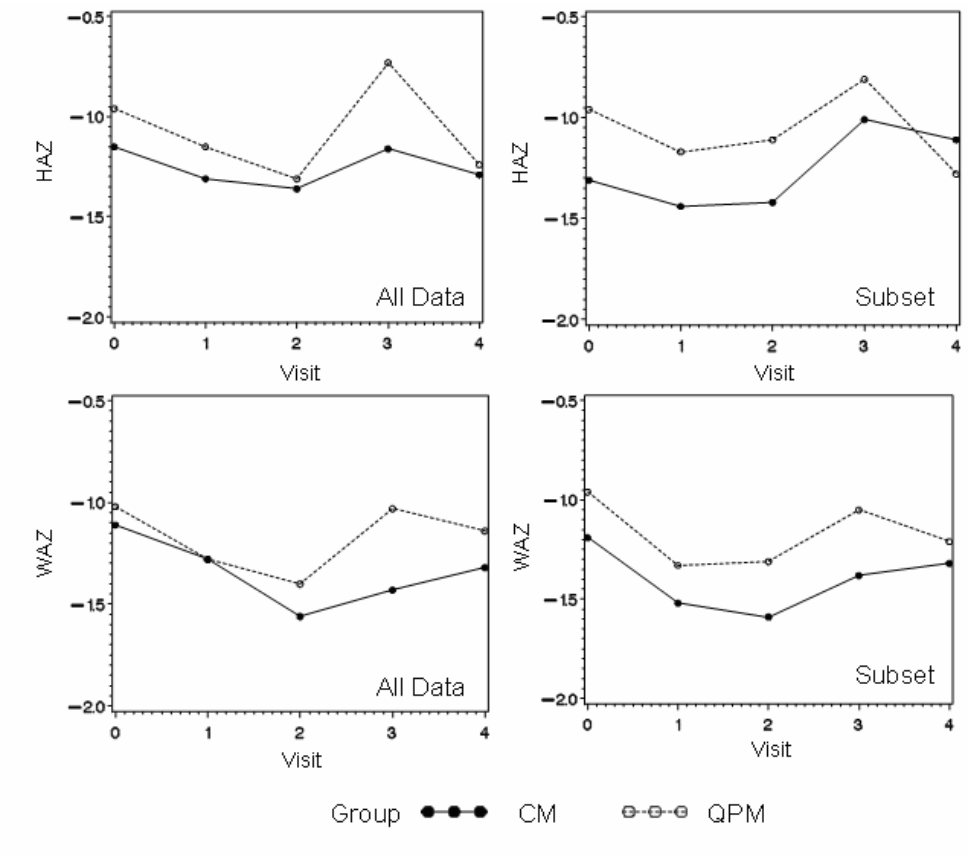

Figure 2.1 Mean HAZ and WAZ scores for the QPM and CM groups in the Ethiopia study. Measurements were taken at baseline and at four quarterly visits during the study year. The graphs on the left were constructed using all available data ("All Data"), while the graphs on the right were constructed using data from children with no missing data ("Subset").

A malaria epidemic struck the study area between the third and fourth visits, negatively impacting subjects' growth. The incidence of malaria among children in QPM households (65%) was higher than among children in CM households (42%). The means computed using the full data set suggest that there may have been some separation between the QPM and CM groups, indicating a favorable effect of QPM consumption, that was later negated by the incidence of malaria. However, no such effect is observed when considering the children for whom complete data are available. In the subset, it

appears that if there were any differences between the CM and QPM groups for HAZ or WAZ, they would have been comparable to the differences that existed at baseline. This example illustrates the value of accounting for missing data in the analysis of data from these studies.

Effect sizes for height and weight for this study were calculated in the following way. First, as it was believed that an atypically high incidence of malaria significantly affected anthropometric measurements in the fourth visit, the third visit was conservatively considered the end of the study. This study was therefore considered to have been 9 months long. Assuming the ages reported by Akalu (2005) were those at the start of the study, the average ages at the start and conclusion of the study were approximately 15 and 24 months. Mean HAZ and WAZ values were used from the subset of children with complete data. If data were missing at random in this study, those mean values would be expected to be comparable to the means that would have been obtained had complete data been available for all children in the study.

The 2000 CDC Growth Charts for young children from the Centers for Disease Control (CDC) were used to convert mean HAZ and WAZ values into mean heights in centimeters and mean weights in kilograms for boys and girls (CDC 2006). Given a roughly equal distribution of sexes in each treatment group, the estimated mean heights and weights were averaged over the sexes to get overall mean heights and weights for each treatment group. The calculated effect sizes were 0.96 for height and 1.11 for weight. Using the full data set, the effect sizes were notably higher at 1.12 for height and 1.31 for weight. These higher values were believed to be artifacts of the missing data structure in the study, and they were not used in the following analysis.

2.3.2. Ghana 1

In the 1990s, a series of four community trials was conducted by the Ghana Health Service – Ashanti in the Ejura-Sekyedumase District of the Ashanti Region in Ghana to evaluate the effect of feeding infants traditional maize porridge made from either QPM or CM (Akuamoa-Boateng 2002). The first study was conducted in May 1993-August 1994. Farmers in nine rural farming communities were randomly assigned seed of either QPM or CM to grow, and the harvested grain was then intended for household consumption over the following year.

Weight was measured monthly and height (length) was measured quarterly on 140 children, aged 4-23 months old, who participated in the study. Statistically significant differences were not observed between the two groups for gains in height or weight (Akuamoa-Boateng 2002). There were significant missing data due to emigration of study participants, and 57 out of 140 children could not be included in the data analysis. Average changes, standard errors, and final sample sizes were reported for each treatment group for both height and weight, allowing effect sizes to be calculated. There was no monitoring or analysis of children's intakes of maize or other foods. As with the Ethiopia study, this study may not have been blinded, as the farmers grew the maize varieties themselves. It is not known whether IRB approval or informed consent was obtained for this study. In the following analysis, this study will be referred to as the "Ghana 1" study.

2.3.3. Ghana 2

The second QPM nutritional study in Ghana was conducted in the Sekyedumase village, Ejura-Sekyedumase District, Ashanti Region, in December 1994-December 1995 (Akuamoa-Boateng 2002). During the 12-month study, 120 children aged 4-15 months were randomly assigned to QPM or CM treatment groups. Their mothers were given dough made of either QPM or CM on a weekly basis to prepare the children's food over the following week. In this and subsequent studies in which dough was provided on a weekly basis, it is not known whether there were problems with rancidity in the absence of refrigeration, which could have affected taste or amount consumed. The children were vaccinated against tuberculosis, diphtheria, tetanus, whooping cough, measles, and polio and de-wormed every six months. Weekly records on morbidity were also kept.

Heights (lengths) and weights were measured quarterly. Statistically significant differences were observed between the two groups for gains in height but not weight. Data for 42 out of the 120 subjects were incomplete due to emigration of the participants from the study area. Those subjects were not included in the data analysis. Average changes, standard errors, and final sample sizes were reported for each treatment group for both height and weight, allowing effect sizes to be calculated. There was no monitoring or analysis of children's intakes of maize or other foods. This study was double-blind, as neither the mothers nor those who distributed the dough knew the identity of the maize that was used. It is not known whether IRB approval or informed consent was obtained for this study. In the following analysis, this study will be referred to as the "Ghana 2" study.

2.3.4. Ghana 3

The third study in Ghana was conducted in four locations in the Ejura-Sekyedumase District, Ashanti Region, during 1998-2000 in two cycles of 12 months each (Akuamoa-Boateng 2002). The mothers of 422 children, aged 4 to 9 months, received dough made of QPM or CM on a weekly basis to prepare their infant's food for the following week. Approximately half the children participated in each study cycle. The children were vaccinated against tuberculosis, diphtheria, tetanus, whooping cough, measles, and polio and de-wormed every six months. They were given a weekly prophylactic dose of chloroquine to reduce the incidence of malaria. Weekly records on morbidity were also kept.

Heights (lengths) and weights of the children were recorded quarterly and monthly, respectively. Analysis of data from 321 children who participated in the entire study found statistically significant differences between the groups for height gain (Akuamoa-Boateng 2002). Differences in weight gain were marginal but not statistically significant at the 0.05 level. Approximately 20% of children in the study were stunted (HAZ < -2) at baseline. As the other Ghana studies were conducted the same district with young children of similar or older ages, comparable or higher levels of undernutrition are expected at baseline in those studies. Three-day dietary assessments were conducted at two times on 118 children in each treatment group. There was no regular monitoring of children's intakes of maize or other foods. This study was double-blind, as neither the mothers nor those who distributed the dough knew the identity of the maize that was used. It is not known whether IRB approval or informed consent was

obtained for this study. In the following analysis, this study will be referred to as the "Ghana 3" study.

Average changes, standard errors, and final sample sizes were reported for each treatment group for both height and weight, allowing effect sizes to be calculated. The principal investigator, Abenaa Akuamoa-Boateng, also provided raw anthropometric data from this study. These data are analyzed below and used to investigate the robustness of the effect size calculated for this study. The raw data set contained the following for each child:

- A record number and child identification number
- Sex
- Age group (4-6 months or 7-9 months at the start of the study)
- Maize group (QPM or CM)
- Height in centimeters at baseline and at the end of the study
- Weight in kilograms at baseline and at the end of the study.

Data were not available on study cycle or location. This structure must therefore be ignored in the following statistical analysis.

As described in the study report (Akuamoa-Boateng 2002), there were records on 422 children, 101 of which had neither a height nor a weight measurement at the end of the study, resulting in a 24% dropout rate. However, at least some of these 101 children were considered dropouts because their mothers failed to pick up the weekly maize dough on a regular basis (A. Akuamoa-Boateng, personal communication). Therefore, some of the reported dropouts were actually subjects that did not comply with the assigned treatment.

Though the dropout rate in this study was relatively high, there was no evidence that it was not random. Specifically, dropping out was not associated with the child's age group, sex, treatment group, baseline height, or baseline weight. Overall, children in the two treatment groups did not significantly differ in baseline height or weight, although among the 4-6 month olds, children in the QPM group weighed less than children in the CM group. The treatment groups did not significantly differ with respect to age distribution or sex ratio. As could be expected, younger children and female children had lower baseline heights and weights.

To study the effect of QPM versus CM consumption on child height, an analysis of variance (ANOVA) model was fitted to the data with change in height (in centimeters) as the response, and age group, sex, maize group, and baseline height as independent variables with fixed effects. Interactions among the independent variables were also considered. Figure 2.2 below suggests a positive effect of QPM on change in height, controlling for baseline height. There were significant effects due to maize group ($F_{1,308} = 34.36$, $p < 0.0001$), age ($F_{1,308} = 12.16$, $p = 0.0006$), and baseline height ($F_{1,308} = 46.91$, $p < 0.0001$). Children aged 4-6 months at the start of the study grew an average of 0.9 cm (95% CI: 0.4-1.4 cm) more during the 12-month study than children aged 7-9 months at the start of the study. Children consuming QPM grew an average of 1.2 cm (95% CI: 0.8-1.6 cm) more than children consuming CM during the same period.

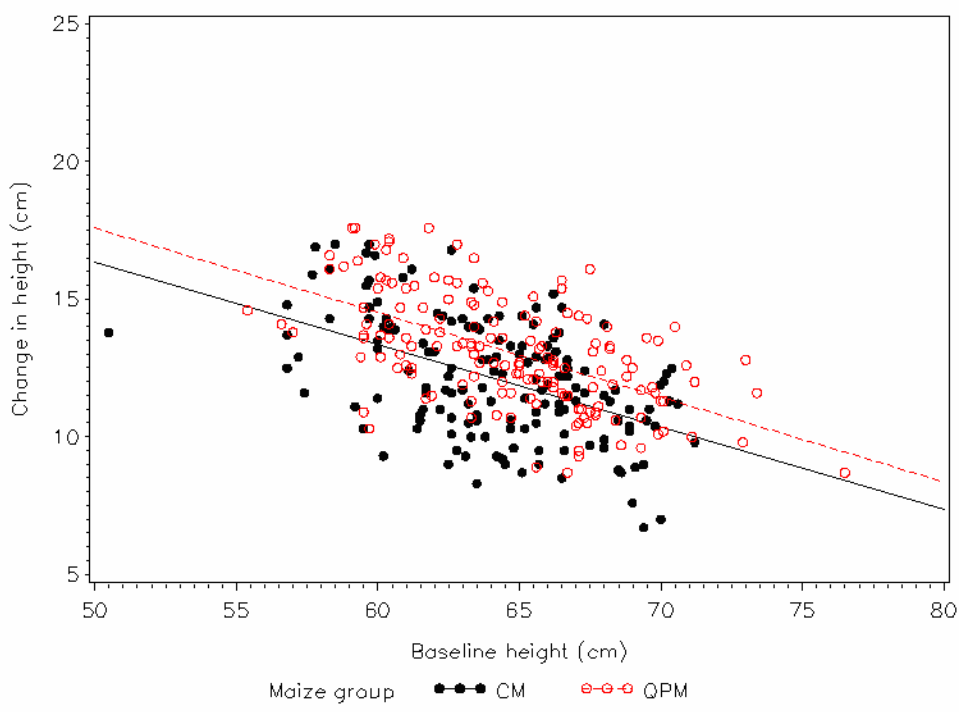

Figure 2.2 Change in height in the QPM and CM groups in the Ghana 3 study.

These results suggest that consumption of QPM over CM had a significant positive effect on the growth in height of children, taking into account age group, sex, and height of the child prior to the start of the study. While it is possible that this significant effect could be due wholly or in part to systematic differences among locations or years with respect to other factors that affect growth in height, such factors may have also had an effect on baseline height. Therefore, controlling for baseline height may have accounted for at least some of these potentially confounding factors.

Similarly, to study the effect of QPM versus CM consumption on child weight, an ANOVA model was fitted to the data with change in weight (kg) as the response, and age group, sex, maize group, and baseline weight as independent variables with fixed effects. Interactions among the independent variables were also considered. Figure 2.3 below

suggests a positive effect of QPM on change in weight, controlling for baseline weight. ANOVA finds significant effects due to maize group ($F_{1,308} = 11.72$, $p = 0.0007$), age ($F_{1,308} = 17.34$, $p < 0.0001$), and baseline weight ($F_{1,308} = 14.07$, $p = 0.0002$). Children aged 4-6 months at the start of the study grew an average of 0.3 kg (95% CI: 0.2- 0.5 kg) more during the 12-month study than children aged 7-9 months at the start of the study. Children consuming QPM grew an average of 0.3 kg (95% CI: 0.1-0.4 kg) more than children consuming CM during the same period.

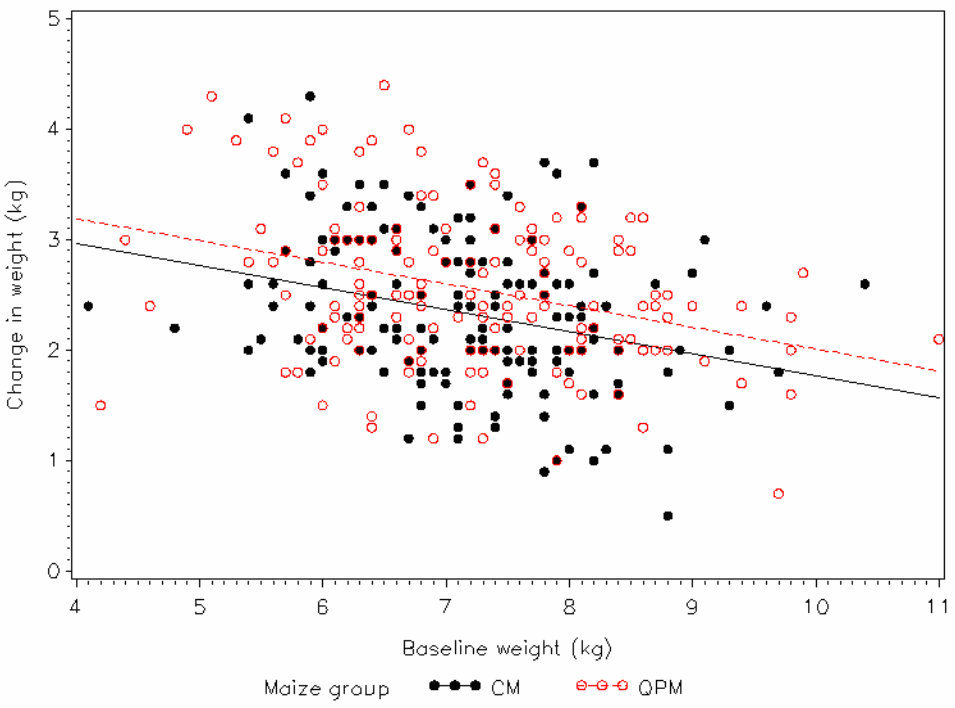

Figure 2.3 Change in weight in the QPM and CM groups in the Ghana 3 study.

This suggests that consumption of QPM over CM had a significant positive effect on growth in weight of children, taking into account age group, sex, and weight of the

child prior to the start of the study. Given the significant effect of maize group on height, this result is not surprising. Again, while it is possible that this significant effect could be due wholly or in part to systematic differences among locations or years with respect to other factors that affect growth in weight, such factors may have also had an effect on baseline weight. Therefore, controlling for baseline weight may have accounted for at least some of these potentially confounding factors.

The average changes in each treatment group were calculated independently and their ratio was taken to find the effect size. However, in the Ghana 3 study, sufficient data were available to adjust the average changes by baseline values. The average changes for each age group and sex were calculated to determine an effect size for each cohort. These alternative calculations of average changes and their resulting effect sizes for height and weight are presented in Tables 2.1 and 2.2, respectively. In these tables, "Independent groups" indicates calculation of average changes separately for each treatment group. The other methods require fitting a linear model with the change in height or weight as the response and age group, sex, and treatment group as independent variables. Estimates are then made for each treatment group overall ("All children") or by age and sex category. These estimates were made with and without adjusting for baseline measurements.

It is observed in Tables 2.1 and 2.2 that the estimated average changes in each treatment group may differ appreciably depending on whether sex, age group, or baseline measurements are taken into account. However, the corresponding effect sizes are robust to these alternative methods of calculation. This supports the choice of the effect size developed for this analysis. If sufficient data were available, it would be desirable to take

into account the additional structure with respect to study cycle and location in this data set as well as in other studies included in this analysis. However, it is possible that the calculated effect size would remain robust to these other factors in this and other studies. To be consistent in the calculation of effect size across the included studies, the effect sizes calculated independently for the two treatment groups without controlling for baseline measurements will be used.

Table 2.1 Effect sizes for height in the Ghana 3 study, calculated independently for the QPM and CM groups ("Independent groups") or by using linear models that account for age group, sex, or baseline measurements.

		Sample size		Change in height (cm)		
Group	Adjustment for baseline	QPM	CM	QPM (SE)	CM (SE)	Effect size
Independent groups	yes	161	156	13.08 (0.13)	12.11 (0.15)	1.08
	no	161	156	13.08 (0.16)	12.11 (0.17)	1.08
All children	yes	161	156	13.12 (0.14)	11.97 (0.14)	1.10
	no	161	156	13.05 (0.15)	11.98 (0.15)	1.09
females, 4-6 months	yes	42	39	13.46 (0.21)	12.32 (0.22)	1.09
	no	42	39	14.10 (0.21)	13.03 (0.21)	1.08
males, 4-6 months	yes	41	50	13.66 (0.20)	12.51 (0.19)	1.09
	no	41	50	13.93 (0.21)	12.86 (0.20)	1.08
females, 7-9 months	yes	38	28	12.59 (0.21)	11.44 (0.22)	1.10
	no	38	28	12.17 (0.21)	11.11 (0.23)	1.10
males, 7-9 months	yes	40	39	12.78 (0.23)	11.63 (0.22)	1.10
	no	40	39	12.00 (0.21)	10.94 (0.21)	1.10

Table 2.2 Effect sizes for weight in the Ghana 3 study, calculated independently for the QPM and CM groups ("Independent groups") or by using linear models that account for age group, sex, or baseline measurements.

		Sample size		Change in weight (kg)		
Group	Adjustment for baseline	QPM	CM	QPM (SE)	CM (SE)	Effect size
Independent groups	yes	160	157	2.56 (0.05)	2.33 (0.05)	1.10
	no	160	157	2.56 (0.06)	2.33 (0.05)	1.10
All children	yes	160	157	2.56 (0.05)	2.31 (0.05)	1.11
	no	160	157	2.55 (0.05)	2.30 (0.05)	1.11
females, 4-6 months	yes	42	39	2.74 (0.08)	2.50 (0.08)	1.10
	no	42	39	2.83 (0.07)	2.58 (0.07)	1.10
males, 4-6 months	yes	41	52	2.70 (0.07)	2.45 (0.07)	1.10
	no	41	52	2.72 (0.07)	2.47 (0.07)	1.10
females, 7-9 months	yes	38	27	2.41 (0.07)	2.16 (0.08)	1.11
	no	38	27	2.39 (0.08)	2.13 (0.08)	1.12
males, 7-9 months	yes	39	39	2.37 (0.08)	2.12 (0.08)	1.12
	no	39	39	2.27 (0.07)	2.02 (0.07)	1.12

2.3.5. Ghana 4

The fourth study in Ghana was conducted in two communities in the Ejura-Sekyedumase District, Ashanti Region, in May-December 2001 (Akuamoa-Boateng 2002). This study evaluated the effects of both QPM and barley malt in the preparation of infant food, with the latter being added to increase the energy density of the food. There were four treatments, in which mothers received QPM or CM dough with barley malt or a placebo (toasted wheat flour) on a weekly basis to prepare infant food. Six hundred children, aged 4 and 6 months, were randomly assigned to one of the treatment groups for the 7-month study.

The children were vaccinated against tuberculosis, diphtheria, tetanus, whooping cough, measles, and polio. They were also given a weekly prophylactic dose of chloroquine to reduce the incidence of malaria and treated for minor ailments as needed. Heights (lengths) and weights of the children were recorded quarterly and monthly, respectively. Analysis of data from 486 children who participated in the entire study found significantly higher height and weight gains with the use of barley malt, regardless of whether QPM or CM was used to make infant food (Akuamoa-Boateng 2002). With the use of barley malt, weight gains were higher in the QPM group than in the CM group.

It is not known whether this study was double-blind with respect to the type of maize provided to the participants. Some consumption studies were conducted, but there was no regular monitoring of children's intakes of maize or other foods. It is not known whether IRB approval or informed consent was obtained for this study. In the following analysis, this study will be referred to as the "Ghana 4" study.

Average changes, standard errors, and final sample sizes were reported for each treatment group for both height and weight (Akuamoa-Boateng 2002). The reported change in the barley malt and placebo groups were weighted by their respective sample sizes and averaged for each type of maize. These averaged changes in the QPM and CM groups were used to calculate effect sizes for height and weight for this study.

2.3.6. India

A study was conducted in the J. J. Colony, Inderpuri, and the Harijan colony of Naraina village in Delhi, India, in 1975-1976 (Singh 1977; Singh et al. 1980). The six-month study was conducted by the Indian Agricultural Research Institute (IARI) and involved 134 children aged 18-30 months. The children were allocated to four groups, where they received daily meals of *o2* maize, CM, or milk or no supplementary meal (control group). Children in the two maize groups were from the J. J. Colony, Inderpuri, children in the milk group were from Naraina village, and children in the control group were from both locations. The researchers took care to balance the four groups with respect to age, sex, and initial weight.

All children were de-wormed before the start of feeding. Daily consumption of the supplemental meal and amounts not consumed were recorded for each child. Dietary surveys of home food consumption were also conducted three times during the study. Anthropometric measurements suggested a positive benefit of *o2* maize relative to CM, but little or no statistical analysis was conducted to establish the significance of these observations. It is not known if there was any blinding with respect to the two maize groups. Blinding would not have been possible with the milk and control treatments. It

is not known whether IRB approval or informed consent was obtained for this study. In the following analysis, this study will be referred to as the "India" study.

For each child, height, weight, and other anthropometric measurements were regressed over time. The estimated slopes corresponded to each child's growth rate during the study. The mean and standard error of these estimated slopes were reported for each treatment group (Singh 1977). These mean slopes were used to calculate effect sizes for this study.

2.3.7. Mexico

A study was conducted in 2001-2002 in four indigenous communities in the Mazateca and Mixe regions of Oaxaca, Mexico, to evaluate the effect of QPM versus CM on the growth of children under five years old (Morales Guerra 2002). This study was conducted by the Colegio de Postgraduados in the State of Mexico. Households were provided with either QPM or CM grain for consumption, and the weights of 67 young children in those households were assessed monthly as a percentage of the median weight of a reference population of the same age (Morales Guerra 2002). This outcome was referred to as physical development ("desarrollo fisico"). The growth reference standards used were the Norma Oficial Mexicana para el Control de la Nutrición, Crecimiento y Desarrollo del Niño y del Adolescente (Secretaría de Salud 1994, as cited in Morales Guerra 2002). At baseline, all children in the study exhibited some degree of undernutrition as measured by physical development, with average physical development among all participants around 80%.

Households in two of the communities received QPM, while households in the other two communities received CM. Distribution of grain to households began in January 2001 but was suspended in February 2001 due to questions regarding the protein quality of the QPM lot. A new lot of QPM was obtained and grain distribution began again in July 2001. Therefore, while the total study duration was 14 months, the treatment lasted only about 8 months. Significant differences were observed between the two treatments in the proportion of children that recovered from malnutrition during the study period (Morales Guerra 2002). Specifically, there was a positive change in physical development in the QPM group but not in the normal maize group. There was no regular monitoring of children's intakes of maize or other foods. The degree of blinding in this study is not known. This study was authorized by the Comité de Ética de la Secretaría de Salud del Estado de Oaxaca (Ethics Committee of the Health Secretariat of the State of Oaxaca). In the following analysis, this study will be referred to as the "Mexico" study.

Raw data were available on the physical development of each child at baseline and at the conclusion of the study (Morales Guerra 2002). Baseline and change in physical development among children in the QPM and CM groups is shown in Figure 2.4. The CM group had some subjects with very low physical development values at baseline, compared to the QPM group. Changes in physical development were both positive and negative in the CM group and did not appear significantly different from normal ($p = 0.2860$ using the Shapiro-Wilk test for normality), while changes in physical development in the QPM group were strictly positive and skewed upward ($p = 0.0018$ using the Shapiro-Wilk test for normality).

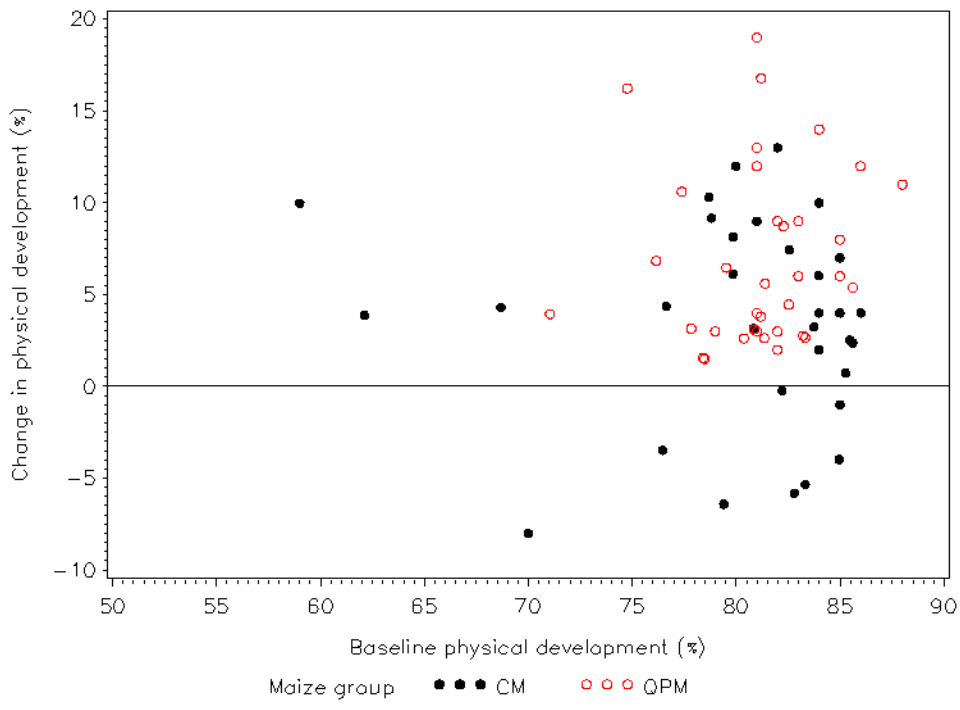

Figure 2.4 Baseline and change in physical development (percentage of median weight of a reference population of the same age) among children in the QPM and CM groups in the Mexico study.

Baseline physical development did not have a significant effect on the change in physical development during this study ($F_{1,64} = 0.03$, p = 0.8735). The type of maize that was consumed did have a significant effect ($F_{1,64} = 7.19$, p = 0.0093). Children who consumed QPM had a 3.4 point greater increase in physical development (95% CI: 0.9-6.0) than children who consumed CM. Exclusion of observations with low baseline values did not significantly change this result.

The analysis above does not take into account that the treatment was randomized at the community rather than the household level, and systematic differences among the four communities could also lead to the significant differences between the treatment groups. Other factors such as amount of maize consumption and socioeconomic

characteristics, which could have an impact on the treatment effect, could be included in the data analysis as one method to address this concern. Also, it should be kept in mind that the application of the treatment began about six months after the baseline measurements used here were taken. There is evidence that physical development increased on average in both treatment groups between these baseline measurements and the start of treatment (Morales Guerra 2002), which could also affect the results of the data analysis.

The ratio of changes in physical development between the QPM and CM groups was used to calculate an effect size for weight for this study. Though this effect size is considered in some of the analyses below, it should be remembered that this statistic was not calculated using actual weights and therefore has a different interpretation from other effect sizes in the meta-analyses. The effect size is 1.97 when calculated using independent groups and 1.98 when a linear model including baseline values was fit to the data. Again, the effect size was robust to the method of calculation. For consistency with the other studies, the effect size calculated using independent groups will be used for subsequent analyses.

2.3.8. Nicaragua

A QPM nutritional study was conducted at the Centro de Desarrollo Infantil Mildred Abaunza in Managua, Nicaragua, in 2005 (Ortega Alemán et al. 2006). During the three and half month-long study, 48 children aged 1-5 years were randomly assigned to consume a maize-based snack made of QPM or CM once a day for five days per week

while they attended the day care center. This was a double-blind study. Informed consent was obtained from participants' guardians at the start of the study.

The children's heights and weights, as well as data on acute respiratory infections and acute diarrheal illnesses, were recorded during the study. One criterion for inclusion was that a child demonstrate undernutrition according two or more anthropometric indicators (height-for-age, weight-for-age, or weight-for-height). Therefore, all study participants were undernourished at baseline. Daily consumption of the snack and amounts not consumed were recorded for each child. A greater proportion of children in the QPM group demonstrated positive changes in grade of malnutrition with respect to weight-for-age and height-for-age but not weight-for-height. They also had fewer cases of acute respiratory and diarrheal infections. In the following analysis, this study will be referred to as the "Nicaragua" study.

Dr. Eveling Ortega Alemán, who with colleagues conducted this study, provided average changes, standard errors, and final sample sizes for each treatment group for both height and weight, allowing effect sizes to be calculated. These effect sizes (1.66 for height and 4.27 for weight) were significantly higher than effect sizes calculated for other included studies and represent estimates that are biologically unlikely (a 66% increase in growth rate for height and a 327% increase in growth rate for weight in the QPM compared to the CM group). It is likely that the small sample size, wide age range, and short duration of this study led to greater noise in the estimation of average changes and consequently in the calculation of the effect size. This is explored further in the analysis below.

2.3.9. Other QPM Nutritional Studies

Other studies have been conducted to evaluate the nutritional impact of QPM or *o2* maize on humans but were excluded from the analyses below because sufficient information and results from the studies could not be obtained or because the studies compared the effects of QPM/*o2* maize to milk-based foods rather than to CM. In 1976-1977, a 17-month study was conducted by the the Instituto de Nutrición de Centro América y Panamá (INCAP) on nine Guatemalan coffee plantations to evaluate the impact of switching from CM to QPM production (Valverde et al. 1983, reviewed in Bressani 1991). This study cited benefits to child growth, but there was evidence suggesting that confounding factors and interventions left the results difficult to interpret (Lauderdale 2000). A series of studies conducted in Brazil in the 1990s had similar problems (Paes and Bicudo 1994; M. C. D. Paes, personal communication). These studies were either discontinued or the data were never analyzed.

In 2000 or shortly before, a six-month study was conducted by Rajendra Agricultural University in Bihar, India (U. Singh, personal communication). Forty children, aged 3-4 years old, who were recovering from malnutrition but healthy, consumed prepared meals made of either QPM or CM, and their height, weight, and middle upper arm circumference (MUAC) were monitored. The results indicated greater increases in height and MUAC in the QPM group (S. Paliwal, personal communication). There was no blinding in this study. More recently, a study similar to the Ethiopia study described above is ongoing in the Eastern Wollega Zone of Ethiopia. The results of this study are expected later this year.

Comparing the effect of QPM to milk, child growth in a highly controlled setting was evaluated by Graham et al. (1990) in Peru. Ten children, aged 13-29 months and recovering from malnutrition, were fed QPM as their sole protein source for 90 days. Their growth was compared to that of ten control children being fed modified cow's milk formula. Growth rates, measured using multiple outcomes, were not significantly different between the two groups. There was no discussion of statistical power in this small, though highly controlled, efficacy study.

2.4. Results and Discussion

2.4.1. Choice of Effect Size and Confidence Interval Method

The chosen effect size was the ratio of the average change in the QPM or *o2* maize group to the average change in the CM group. Both the numerator and denominator of this ratio were average changes over the duration of the study. Both terms could therefore be considered average growth rates in their respective treatment groups, and the effect size could be interpreted as the average growth rate in the QPM/*o2* maize group relative to the average growth rate in the CM group. For example, an effect size of 1.08 could be interpreted as an 8% increase in growth rate in the QPM/*o2* maize group relative to the CM group. An effect size less than 1 indicates that the growth rate was faster in the CM group, while an effect size of 1 indicates that growth rates were the same in the two treatment groups. In addition to relative growth rate, the effect size proposed in this dissertation could be interpreted more generally as a relative rate or a

percent change (percent increase or decrease), making it an useful measure in a broader range of applications.

In meta-analyses of two-group studies to assess a treatment effect, the difference between the mean responses in the two groups, rather than their ratio, is typically used as the effect size (Cooper and Hedges 1994). This mean difference may be standardized using the pooled within-group standard deviation. However, in the studies under consideration here, the mean difference could depend on study duration and age of children in the study. If there is a positive effect of QPM on child growth, studies of longer duration may have larger mean differences between the two treatment groups. Likewise, as growth rates decrease with age, studies with older subjects may have smaller mean differences. In addition to the meaningful interpretation of the ratio of the differences, the proposed effect size removes the effect of child age and study duration in the comparison of QPM/*o2* maize relative to CM. The proposed effect size also did not depend on any within-treatment standard deviations, which was desirable as for several studies, those values were unavailable or were calculated without considering some structure in the study design such as randomization to communities rather than individuals. Another benefit of the proposed effect size was its robustness to alternative methods of calculation, as was seen in the analysis of the Ghana 3 and Mexico studies.

The use of bootstrapping to construct a confidence interval for the summary effect size also offered several advantages. Calculation of asymptotic confidence intervals required estimation of the variances of individual effect sizes. These variances were approximated using the delta method and were based on the estimated standard errors of the average changes in the QPM and CM groups. For several studies, the standard errors

of average changes in the treatment groups were estimated without considering some structure in the respective study designs. This increased the uncertainty associated with the estimated variances of the individual effect sizes, as well as the asymptotic confidence intervals that were calculated for summary effect sizes based on these data. If standard errors of the average changes in the treatments groups were unavailable for a given study, as they were for the Ethiopia study, then that study could not be included in calculating a summary effect size and the associated asymptotic confidence interval. The bootstrap percentile confidence intervals did not depend on the standard errors of average changes in the treatment groups or on the variances of the individual effect sizes and therefore avoided these limitations. Furthermore, bootstrap percentile confidence intervals took into account between-study variation. In these meta-analyses, variation among studies is expected as the included studies represent different populations, ages, and cultures with different diets, child feeding practices, and other factors.

2.4.2. Growth in Height

Table 2.3 describes the characteristics of studies included in the meta-analysis with respect to growth in height. In each study, sample sizes in the treatment groups were comparable or equal. The Ghana 3 and Ghana 4 studies involved significantly greater sample sizes than the other included studies. Children in the included studies were under five years old, with most less than 24 months at the start of the study. This is the age range when diet is most likely to impact child growth (G. Beaton, personal communication).

Table 2.3 Characteristics of included studies with respect to growth in height.

	QPM			CM			Effect size				
Study	**N**	**Change ***	**SE**	**N**	**Change ***	**SE**	**Estimate**	**SE**	**Duration (months)**	**Age at start of study (months)**	**Form of treatment**
Ethiopia	51	8.53	-	51	8.89	-	0.96	-	9	most under 24	seed
Ghana 1	43	10.60	0.65	40	9.91	0.67	1.07	0.10	12	4-23	seed
Ghana 2	39	14.76	0.68	39	12.37	0.68	1.19	0.09	12	4-15	dough
Ghana 3	161	13.08	0.16	156	12.11	0.17	1.08	0.02	12	4-9	dough
Ghana 4	240	7.77	0.23	246	7.28	0.22	1.07	0.05	7	4-6	dough
India	32	0.43	0.03	35	0.37	0.06	1.16	0.19	6	18-30	meal
Nicaragua	24	2.02	1.27	24	1.22	0.66	1.66	1.37	3.5	12-60	meal

* Change refers to the change in height (cm) over the duration of the study, except for India, where it is the change per fortnight (two weeks).

Study participants generally showed signs of mild to moderate undernutrition, as determined by height-for-age, weight-for-age, or weight-for-height, and lived in communities where maize is a dietary staple. Results of a meta-analysis based on these data are therefore expected to apply to young children, particularly infants and toddlers, who show mild to moderate undernutrition and for whom maize is an important part of the diet. Young children in maize-consuming communities who show evidence of undernutrition constitute the primary target population for QPM.

Among the included studies, the maize treatment was delivered to study participants in one of three ways: as a prepared meal given directly to children under supervision; as dough that could be used to prepare infant food, given to mothers; or as seed given to members of a household, which could be used for cultivation, harvest, storage, food preparation, and ultimately consumption by the target child in that household. Studies that provided the treatment in the form of a prepared meal required more intervention and monitoring by the researchers. Consequently, these studies had smaller sample sizes and durations than studies that provided the maize treatment in other ways. There were no discernable relationships between effect size and study duration, sample size (as represented by sample size in the CM group), and form of treatment (Figure 2.5).

Effect sizes for all studies and estimated standard errors of effect sizes for all studies except Ethiopia are also given in Table 2.3. These standard errors were approximated using the delta method and are based on the estimated standard errors of the average changes in the QPM and CM groups. For several studies, the standard errors of average changes in the treatment groups were estimated while ignoring some structure

in the study design (e.g., randomization at the village or community level, rather than at the individual level). As discussed this earlier, this increases the uncertainty associated with the estimated standard errors of those effect sizes.

The effect size for the Nicaragua study was notably higher than the other effect sizes. This study had a small sample size, a wide range of ages among the study participants, and a short duration. Standard errors of average changes in height, relative to the average changes themselves, were high compared with those of other studies. The standard error in the QPM group was twice as high as the standard error in the CM group. In the other studies, standard errors of average changes were comparable between the two treatment groups. Consequently, the estimated standard error of the effect size for the Nicaragua study is also high relative to the effect size and to standard errors of effect sizes for other studies. For these reasons, the Nicaragua study was excluded from the primary meta-analysis for growth in height. Analyses including the study will also be presented for comparison below.

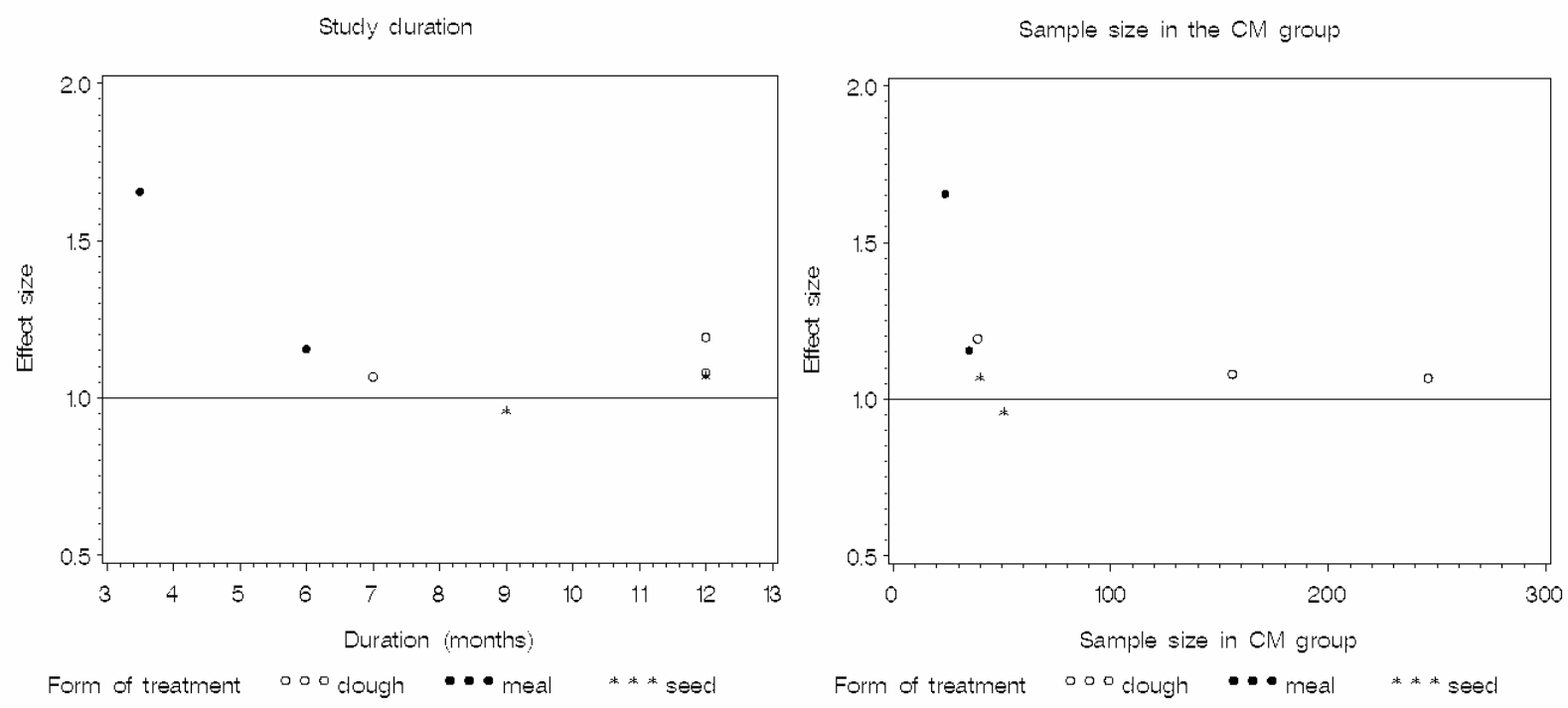

Figure 2.5 Plots of effect sizes for growth in height by study duration, sample size in the CM group, and form of treatment.

Table 2.4 lists the sets of weights used for meta-analyses of the effect of QPM on growth in height. Three sets of weights were used: weights proportional to sample size in the CM group, equal weights, and optimal weights, which were inversely proportional to the variance of the corresponding study's effect size. The primary meta-analysis for growth in height used weights proportional to sample size in the CM group. These weights were preferred over equal weights as studies with larger sample sizes were expected to provide more precise estimates of the effect of QPM relative to CM. These weights were also preferred over the optimal weights as they did not rely on the availability or estimation of standard errors within treatment groups or of individual effect sizes. Weighting by sample size results in greater weights for the Ghana 3 and particularly the Ghana 4 studies. Optimal weights place the majority of the weight on the Ghana 3 study.

Table 2.4 Weights used for meta-analysis of effect of QPM on growth in height. Three sets of weights were used: weights proportional to sample size in the CM group, equal weights, and optimal weights. Weights were calculated with and without the Nicaragua study. Each column of weights sums to 1.

	Main data set			Main data set + Nicaragua		
Study	Sample size	Equal	Optimal *	Sample size	Equal	Optimal *
Ethiopia	0.09	0.17	-	0.09	0.14	-
Ghana 1	0.07	0.17	0.03	0.07	0.14	0.03
Ghana 2	0.07	0.17	0.04	0.07	0.14	0.04
Ghana 3	0.28	0.17	0.76	0.26	0.14	0.76
Ghana 4	0.43	0.17	0.15	0.42	0.14	0.15
India	0.06	0.17	0.01	0.06	0.14	0.01
Nicaragua	-	-	-	0.04	0.14	0.00

* Excludes Ethiopia because within-study variance was not available.

Table 2.5 presents the summary effect sizes, confidence intervals, and p-values for the analysis of growth in height. The three weighting methods were studied with and without the inclusion of the Nicaragua study. For optimal weights, bootstrap and asymptotic confidence intervals were calculated. The summary effect size and bootstrap confidence interval based on the main data set with weights proportional to sample sizes in the CM groups was considered the main result on the effect of QPM on the growth in height of young children. The results indicate that consumption of QPM instead of CM leads to an 8% (95% CI: 4-12%) increase in the rate of growth in height (length) in infants and toddlers with mild to moderate undernutrition for whom maize is a significant part of the diet ($p = 0.0014$). The estimated summary effect size and confidence interval are generally robust to alterative formulation of weights and methods for calculating the confidence interval.

Summary effect sizes changed little with the inclusion of the Nicaragua study except when equal weights were used. This was not surprising as equal weighting gave the greatest weight to the Nicaragua study among the sets of weights that were examined, while the other sets of weights gave small or negligible weight to the Nicaragua study due to its small sample size and high effect size variance. With equal weights and weights proportional to sample size, inclusion of the Nicaragua study skewed the bootstrap confidence intervals upward, though the lower confidence limits were largely unchanged. With optimal weights, inclusion of this study had negligible effect on the point estimates and confidence intervals because the relatively large standard error of the Nicaragua study's effect size resulted in an optimal weight near zero.

Table 2.5 Summary effect sizes, confidence intervals, and p-values for the analysis of growth in height. Three weighting methods were studied with and without inclusion of the Nicaragua study. For optimal weights, bootstrap and asymptotic confidence intervals were calculated. LCL and UCL are the lower and upper confidence limits, respectively.

		Main data set				Main data set + Nicaragua			
Weighting method	CI method	Estimate	LCL	UCL	P-value	Estimate	LCL	UCL	P-value
Sample size	Bootstrap	1.08	1.04	1.12	0.0014	1.10	1.05	1.22	0.0002
Equal	Bootstrap	1.09	1.03	1.15	0.0037	1.17	1.05	1.35	0.0004
Optimal *	Bootstrap	1.08	1.07	1.16	0.0000	1.08	1.07	1.16	0.0000
Optimal *	Asymptotic	1.08	1.05	1.12	0.0000	1.08	1.05	1.12	0.0000

* Excludes Ethiopia because within-study variance was not available.

Asymptotic confidence intervals based on weights inversely proportional to $\text{var}\left(\hat{\theta}_i\right)$ do not take into account between-study variation. Based on the main data set, there is no significant evidence of study heterogeneity with respect to height (Q = 1.96, p = 0.7423). Because k = 5 and $Q < (k-1)$, $\hat{\tau}^2 = 0$, and the asymptotic confidence interval given in Table 2.5 is unchanged. When Nicaragua is included, there is still no significant evidence of study heterogeneity (Q = 2.14, p = 0.8297). Here, k = 6 and $Q < (k-1)$, so $\hat{\tau}^2 = 0$ and the corresponding asymptotic confidence interval in Table 2.5 is unchanged.

Given the main data set, which includes the studies from Ghana, Ethiopia, and India, the sensitivity of the summary effect size to the contribution of each individual study was explored by recalculating the point estimate and 95% bootstrap confidence interval while excluding one study at a time. These results are presented in Table 2.6. Weights were proportional to sample sizes in the CM groups for all calculations. Exclusion of any one study did not significantly change the estimated overall effect of QPM on growth rate in height, regardless of the magnitude of that study's weight in the main analysis, and the positive effect of QPM on relative growth rate in height remains statistically significant with the exclusion of any of these studies.

Table 2.6 Estimates and 95% bootstrap percentile confidence intervals of summary effect sizes for effect of QPM on growth in height, excluding one study in the main data set at a time. Weights were proportional to sample sizes in the CM groups. LCL and UCL are the lower and upper confidence limits, respectively.

Study omitted	Estimate	LCL	UCL
Ethiopia	1.09	1.07	1.15
Ghana 1	1.08	1.04	1.13
Ghana 2	1.07	1.03	1.09
Ghana 3	1.07	1.02	1.15
Ghana 4	1.08	1.02	1.15
India	1.07	1.03	1.11
None	1.08	1.04	1.12

2.4.3. Growth in Weight

Table 2.7 describes the characteristics of studies included in the meta-analysis with respect to growth in weight. In addition to the studies available for the analysis of growth in height, the Mexico study, which reported results relating to weight but not height, is also included in the following analysis. It should be remembered that the calculation of the effect size for Mexico differed from the calculation of other effect sizes as it was based on physical development rather than weight in kilograms. The analyses described below are conducted both with and without the Mexico study, but this study is excluded from the main analysis of the effect of QPM on growth in weight. The effect size and standard error for this study are significantly higher than those of all other included studies except Nicaragua, suggesting a 97% increase in growth rate in weight in the QPM compared to the CM group or, more accurately, a 97% increase in the rate of

Table 2.7 Characteristics of included studies with respect to growth in weight.

	QPM			CM			Effect size				
Study	**N**	**Change ***	**SE**	**N**	**Change ***	**SE**	**Estimate**	**SE**	**Duration (months)**	**Age at start of study (months)**	**Form of treatment**
Ethiopia	53	1.40	-	52	1.26	-	1.11	-	9	most under 24	seed
Ghana 1	43	2.27	0.19	40	2.36	0.19	0.96	0.11	12	4-23	seed
Ghana 2	39	2.92	0.18	39	2.93	0.18	1.00	0.09	12	4-15	dough
Ghana 3	160	2.56	0.06	157	2.33	0.05	1.10	0.04	12	4-9	dough
Ghana 4	240	2.67	0.07	246	2.42	0.06	1.10	0.04	7	4-6	dough
India	32	0.12	0.01	35	0.10	0.01	1.23	0.17	6	18-30	meal
Mexico	35	6.94	0.81	32	3.52	0.99	1.97	0.60	14	most under 60	grain
Nicaragua	24	0.80	0.46	24	0.19	0.45	4.27	10.53	3.5	12-60	meal

* Change refers to the change in weight (kg) over the duration of the study, except for India, where it is the change per fortnight (two weeks), and Mexico, where it is calculated using physical development instead of kg.

improvement in physical development over time. The full duration of the study is listed; however, maize grain was given to subjects for roughly the last eight months of the study.

The estimated effect size for the Nicaragua study is significantly higher than that of all other studies and indicates a biologically unlikely effect of QPM (a 327% increase in growth rate for weight compared to the CM group). The small sample size, wide age range among subjects, and short duration of this study were likely sources of the uncertainty in this estimate, as indicated by its very large standard error. This study was not included in analysis of the effect of QPM on growth in weight. Similar to the earlier analysis of the effect of QPM on growth in height, the main analysis of the effect of QPM on growth in weight includes the studies from Ghana, Ethiopia, and India. In using the standard errors of the effect sizes for these studies in the following analyses, it should be kept in mind that these values were approximated using the delta method and were based on within-group standard errors that did not take into account all structure in the respective data sets.

Again, subjects in these studies were children under five years, with most under 24 months at the beginning of the studies. They showed mild to moderate undernutrition, as indicated by low height-for-age, weight-for-age, or weight-for-height, and lived in communities where maize is a staple food. The results of the meta-analysis described below therefore apply to young children, particularly infants and toddlers, who show signs of mild to moderate undernutrition and live in maize-consuming communities.

In addition to previously described methods of delivering the improved maize to subjects, the Mexico study delivered the intervention to participating children by providing QPM or CM grain to their households, with the intent that the grain is stored

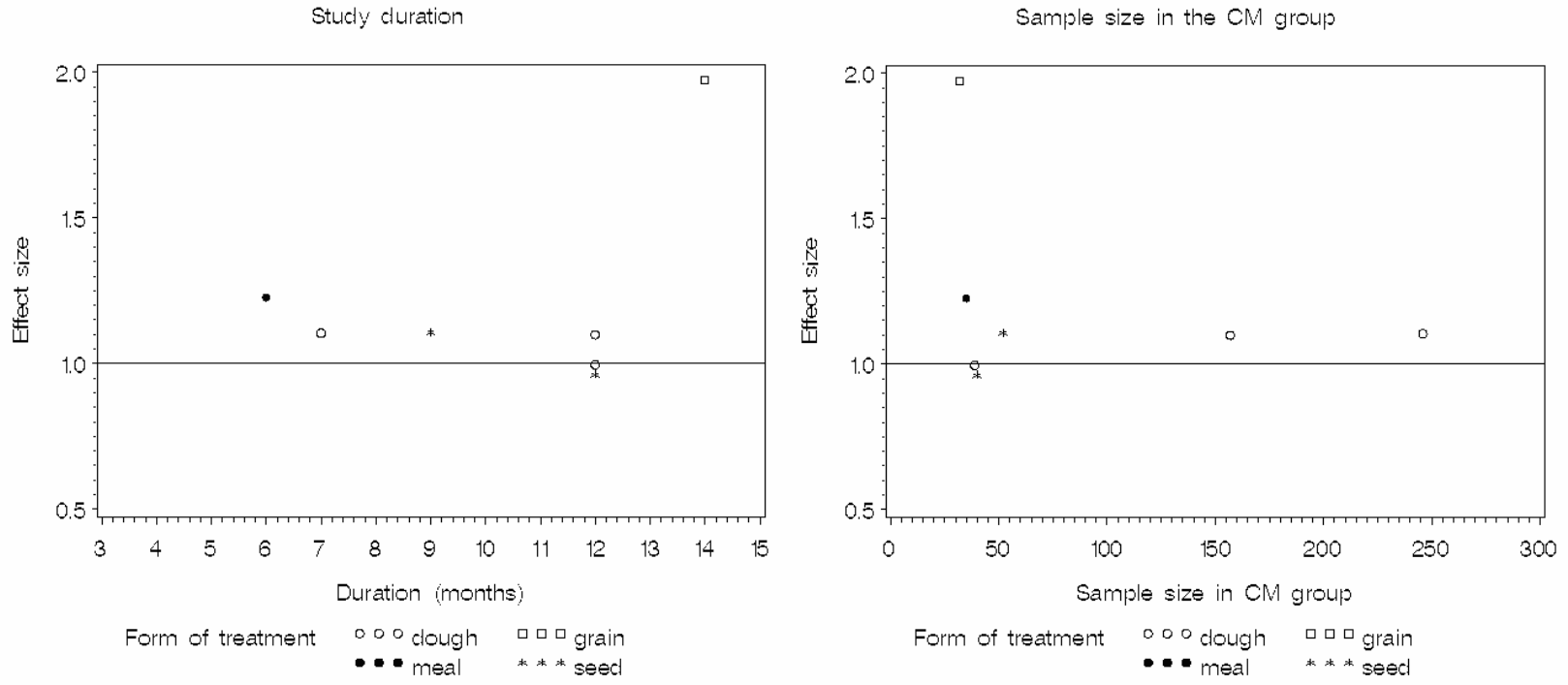

Figure 2.6 Plots of effect sizes for growth in weight by study duration, sample size in the CM group, and form of treatment. The Nicaragua study is not depicted.

and used to make food that is consumed by target children in the household over the duration of the study. The relationships between effect size and study duration, sample size in the CM group, and method to deliver the treatment to subjects is depicted in Figure 2.6. There are no apparent relationships among these variables. The effect size of the Mexico study is clearly larger than all other effect sizes. The Nicaragua study is not included in these or subsequent results.

Table 2.8 lists the sets of weights used for meta-analyses of the effect of QPM on growth in weight. Three sets of weights were used: weights proportional to sample size in the CM group, equal weights, and optimal weights, which were inversely proportional to the variance of the corresponding study's effect size. Weighting by sample size results in greater weights for the Ghana 3 and particularly the Ghana 4 studies. Optimal weights also place the majority of the weight on the Ghana 3 and Ghana 4 studies, but with greater weight on the Ghana 3 study.

Table 2.8 Weights used for meta-analysis of effect of QPM on growth in weight. Three sets of weights were used: weights proportional to sample size in the CM group, equal weights, and optimal weights. Weights were calculated with and without the Mexico study. Each column of weights sums to 1.

	Main data set			Main data set + Mexico		
Study	Sample size	Equal	Optimal *	Sample size	Equal	Optimal *
Ethiopia	0.09	0.17	-	0.09	0.14	-
Ghana 1	0.07	0.17	0.05	0.07	0.14	0.05
Ghana 2	0.07	0.17	0.08	0.06	0.14	0.08
Ghana 3	0.28	0.17	0.48	0.26	0.14	0.48
Ghana 4	0.43	0.17	0.37	0.41	0.14	0.37
India	0.06	0.17	0.02	0.06	0.14	0.02
Mexico	-	-	-	0.05	0.14	0.00

* Excludes Ethiopia because within-study variance was not available.

Table 2.9 presents the summary effect sizes, confidence intervals, and p-values for the analysis of growth in weight. The three weighting methods were studied with and without the inclusion of the Mexico study. For optimal weights, bootstrap and asymptotic confidence intervals were calculated. As in the analysis of growth in height, the summary effect size and bootstrap confidence interval based on the main data set with weights proportional to sample sizes in the CM groups were considered the main result on the effect of QPM on the growth in weight of young children. The results indicate that consumption of QPM instead of CM leads to a 9% (95% CI: 4-12%) increase in the rate of growth in weight in infants and toddlers with mild to moderate undernutrition for whom maize is a significant part of the diet (p = 0.0019). The estimated summary effect size and confidence interval were generally robust to alterative formulation of weights and calculation of the confidence interval.

Inclusion of the Mexico study increases the summary effect size and upper confidence limit when equal weights and weights proportional to sample sizes are used with bootstrapping to generate the confidence interval. Results using optimal weights are negligibly affected as the relatively large standard error of the Mexico study's effect size resulted in an optimal weight near zero.

Based on the main data set, there is no significant evidence of study heterogeneity with respect to weight (Q = 3.37, p = 0.4982). Because $k = 5$ and $Q < (k-1)$, $\hat{\tau}^2 = 0$, and the asymptotic confidence interval given in Table 2.9 is unchanged. When Mexico is

Table 2.9 Summary effect sizes, confidence intervals, and p-values for the analysis of growth in weight. Three weighting methods were studied with and without inclusion of the Mexico study. For optimal weights, bootstrap and asymptotic confidence intervals were calculated. LCL and UCL are the lower and upper confidence limits, respectively.

		Main data set				Main data set + Mexico			
Weighting method	CI method	Estimate	LCL	UCL	P-value	Estimate	LCL	UCL	P-value
Sample size	Bootstrap	1.09	1.04	1.12	0.0019	1.14	1.06	1.32	0.0002
Equal	Bootstrap	1.08	1.02	1.15	0.0073	1.21	1.03	1.48	0.0010
Optimal *	Bootstrap	1.09	1.00	1.11	0.0187	1.09	1.01	1.11	0.0121
Optimal *	Asymptotic	1.09	1.04	1.14	0.0001	1.09	1.04	1.14	0.0001

* Excludes Ethiopia because within-study variance was not available.

included, there is still no significant evidence of study heterogeneity (Q = 5.54, p = 0.3541). As k = 6 and $Q > (k-1)$, it is estimated that $\hat{\tau}^2 = 0.0005$, which results in a summary effect size of 1.09 and a 95% asymptotic confidence interval of (1.03, 1.14). This is only small change from the interval reported in Table 2.9.

Using the main data set, which includes the studies from Ghana, Ethiopia, and India, the sensitivity of the summary effect size to the contribution of each individual study was explored by recalculating the point estimate and 95% bootstrap confidence interval while excluding one study at a time. These results are presented in Table 2.10. Weights were proportional to sample sizes in the CM groups for all calculations. Exclusion of any one study did not significantly change the estimated overall effect of QPM on growth rate in weight, regardless of the magnitude of that study's weight in the main analysis. The positive effect of QPM on relative growth rate for weight remains statistically significant with the exclusion of any of these studies.

Table 2.10 Estimates and 95% bootstrap percentile confidence intervals of summary effect sizes for effect of QPM on growth in weight, excluding one study in the main data set at a time. Weights were proportional to sample sizes in the CM groups. LCL and UCL are the lower and upper confidence limits, respectively.

Study omitted	Estimate	LCL	UCL
Ethiopia	1.09	1.02	1.13
Ghana 1	1.10	1.07	1.14
Ghana 2	1.10	1.06	1.14
Ghana 3	1.09	1.01	1.14
Ghana 4	1.09	1.01	1.14
India	1.09	1.01	1.10
None	1.09	1.04	1.12

2.5. Recommendations for Future Studies

2.5.1. Framework for Impact of QPM

Lauderdale (2000) described two potential pathways for QPM to have nutritional impact, the first through the use of QPM for human food and the second through its use as animal feed. Consumption of QPM, obtained from home production, purchase, or some other source, by target individuals is the most direct pathway to nutritional impact. De Groote et al. (2006) described this impact pathway as crossing multiple levels, in which the new technology reaches the farm household through a network of community institutions and ultimately consumption by individuals decreases a nutrient deficiency, thereby improving nutritional status and health. The discussion below builds on the conceptual framework presented by De Groote et al. (2006) by specifying factors that modify individuals' consumption of QPM and other foods and by incorporating nutritional and health aspects in the conceptual framework.

The impact of QPM on individuals' nutritional status and health can be modified by factors at each level. At the community level, impact may be affected by agricultural extension and seed systems, availability and cost of QPM grain and seed in local markets, and other existing food and public health interventions such as school lunch programs. The purity of a farmer's QPM seed or QPM grain sold in the local market may vary depending on the source, also affecting ultimate impact on nutritional status and health.

At the household level, the decision to adopt and produce QPM will depend on various factors including household and farmer characteristics, availability and cost of seed, access to credit, awareness of QPM and its potential benefits, and acceptability of

QPM for agricultural production, consumption, or sale. The purity of consumed grain will depend on additional factors, such as production conditions, farmers' knowledge of management of pollen contamination, and storage conditions. Families may supplement their household maize production with purchased maize that may be QPM, CM, or some mixture of the two. The primary source of the maize that is consumed by the household may vary by season, particularly if home production is not sufficient to last until the next harvest. Therefore, nutritional impact of QPM may also vary by season.

Food allocation within the household and awareness of the potential benefits of QPM consumption will affect how much QPM is consumed by target individuals. Nutrient losses during food preparation as well as other sources of relevant nutrients in individuals' diets will also affect the potential benefit of QPM. As QPM is expected to provide a nutritional benefit by alleviating a nutrient deficiency, it is expected to have little impact in communities and populations with little risk of inadequate nutrient intakes.

Factors such as household characteristics, families' socioeconomic status, access to clean water and sanitation, education and awareness of disease and disease management, and information and recommendations from external sources on child feeding and care will affect the disease status of target individuals or their caregivers. Disease in turn, through its effects on diet, child feeding practices, appetite, nutrient malabsorption or loss, or nutrient requirements, may affect the potential impact of QPM (Jackson and Calder 2004). In particular, even with apparently adequate consumption that would otherwise lead to positive effects, there may not be a decrease in deficiency due to malabsorption or loss or increased requirements. This highlights the need to

consider the role of morbidity in assessing the impact of QPM. The nutritional and health status of target individuals prior to QPM introduction may further affect the impact of QPM. This framework underlies all community nutritional studies on QPM, with specific studies varying in their starting points along the impact pathway.

2.5.2. Distinction Between Efficacy and Effectiveness

There is a need to demonstrate the efficacy of QPM consumption. Efficacy refers to the biological effect of QPM on outcomes such as child growth or morbidity. The India and Nicaragua studies were efficacy studies, as each involved supervised feeding of maize-based meals to subjects and intakes were carefully monitored and measured during the intervention. While both studies cite benefits from QPM on growth of young children, neither study has yet appeared in the peer-reviewed literature. Graham et al. (1990) published a well-controlled efficacy study comparing consumption of QPM to consumption of a milk-based formula by young children. However, this study did not present typical diets as children exclusively consumed the treatment foods. Also, the primary comparison of interest, consumption of QPM to consumption of CM, was not addressed by this study.

A general framework for efficacy studies is given in Figure 2.7. A dietary intervention is consumed by a target individual, improving the adequacy of the individual's intakes with respect to a nutrient of interest, and this leads to improved outcomes such as increased growth or decreased morbidity. In addition to food intakes, nutrient requirements and any nutrient malabsorption or loss also affect nutrient adequacy. Consumption of the intervention may affect the rest of an individual's diet

(for example, by increasing satiety or changing child feeding practices). Morbidity of the subject can play a significant role by affecting the delivery of the intervention or the subject's compliance with the intervention. It can also affect the subject's diet outside of the intervention, nutrient malabsorption and loss, nutrient requirements, or growth itself. Overall, morbidity can affect consumption qualitatively by a change in the subject's diet, as well as quantitatively by a decrease or increase in the consumption of certain foods. Note that in an efficacy study, the unit of study is a target individual, and while uncontrolled external factors may modify the effect of the intervention on the measured outcomes, this is primarily biological rather than behavioral effect modification (Victora et al. 2004).

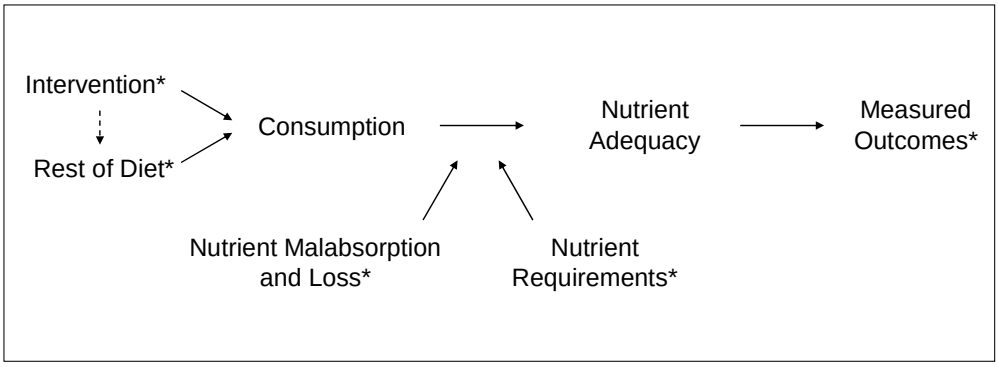

Figure 2.7 General framework for efficacy studies. Items marked with an asterisk are affected by morbidity of the subject.

In contrast, a general framework for effectiveness studies, as it relates to nutritional studies conducted on QPM, is given in Figure 2.8. Maize that is presented to the household as grain or dough must be stored, processed, and prepared by members of

the household that are typically not the target individual. Prepared food is then allocated among members of the household including the target individual. A target individual's dose of the intervention is typically not monitored or measured, presenting a significant source of variation.

Presenting the intervention in the form of seed adds several additional factors to the study. The farmer must completely adopt the variety that is provided, substituting regularly cultivated varieties with the new variety. The quantity of harvested grain depends on many factors including soil fertility, weather conditions, prevalence of crop diseases and pests, agronomic practices, use of inputs, and others, all of which could vary among households. The quality of harvested grain depends on the purity of the seed, farmer's knowledge and management of pollen contamination, other factors such as plot size that may affect pollen contamination, and others. Excesses or shortfalls of grain during the study period may lead to sale or purchase of grain from the local market or sharing of grain among households. Mixing of QPM and CM may also occur during grain storage. These factors have been documented in past QPM studies (see for example the discussion of the Ghana 1 study by Akuamoa-Boateng (2002)).

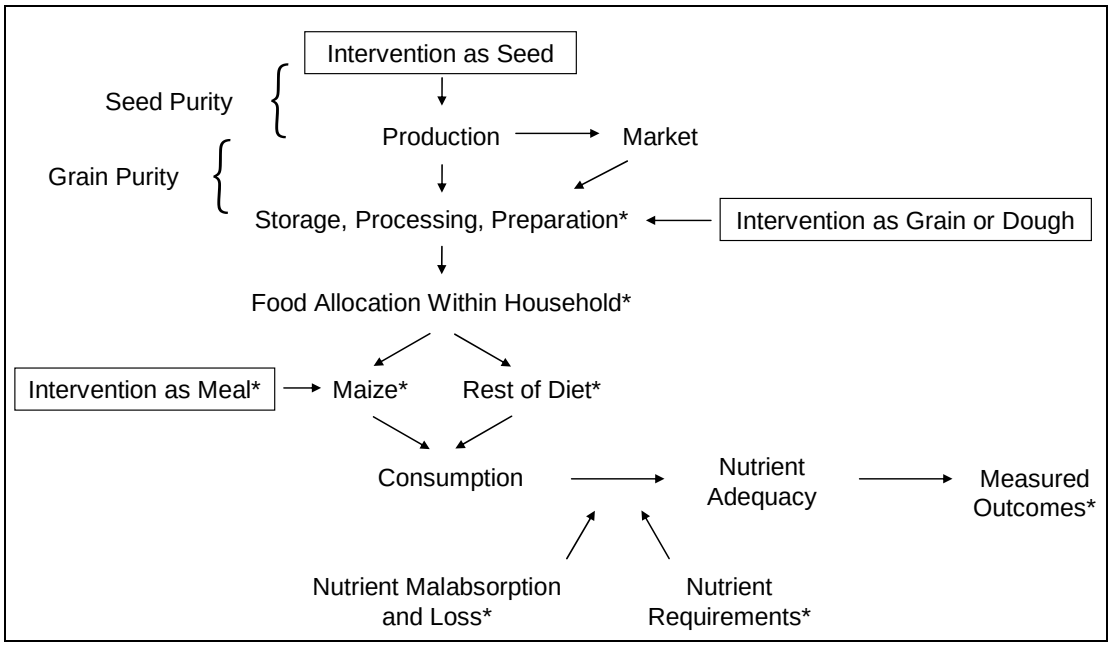

Figure 2.8 General framework for effectiveness studies. Items marked with an asterisk are affected by morbidity of the subject.

Decision-making and behaviors related to agricultural production additionally play a role when the intervention is presented to a household in the form of seed. These decisions and behaviors are made by a household member that is not likely to be the target individual and may not be the person who makes many of the food preparation, allocation, and other household decisions. Furthermore, morbidity of the target individual can affect additional factors in these studies. In these effectiveness studies, the intervention is typically presented to households rather than target individuals. The many additional uncontrolled factors that may modify the effect of the intervention on the measured outcomes would do so through behavioral as well as biological effect modification. In community level studies to evaluate the nutritional impact of QPM or

other biofortified crops, a distinction should be made whether a study evaluates efficacy or effectiveness.

2.5.3. Considerations in Study Design

The goal of efficacy and effectiveness studies is to show an effect of an intervention on outcomes such as child growth among target populations. The main target population for QPM is children under five. However, diet is primarily expected to impact child growth from the age of 6 months to 18-24 months, and older children are not expected to show major responses in growth upon improved nutrition (G. Beaton, personal communication). It is therefore recommended that QPM studies involve participants up to 24 months in age, as any impact from QPM is mostly likely to be observed in this age group.

The primary outcomes in QPM studies have been growth in height, weight, and other anthropometric measurements, though some studies have collected periodic data on child morbidity as well. It is not clear from reports on past QPM studies what measures were taken to ensure data quality. Quality control measures are necessary to collect reliable anthropometric data. Such measures would include repeated measurement of children's heights and weights, as assessed by more than one researcher, and training and standardization of methods among research staff. Growth is also not a very sensitive indicator, as many other factors affect it (for example, disease or parasitic infections). Improving protein quality in the diet would be expected to improve immune function; therefore, immunological indicators could serve as additional measures that would be more sensitive and reliable. There exist indicators, such as delayed cutaneous

hypersensitivity, that allow quantitative assessment of immune function while requiring little access to public health facilities (Gibson 2005; Selmi et al. 2004). The study duration required to observe a measurable treatment effect will likely be shorter with immunological indicators than with anthropometric indicators (N. Scrimshaw, personal communication).

Provision of any treatment, either QPM or CM, delivered as a meal, grain, seed, or other means, will have an effect on the food supply or economic status of the subject or household, as food, money, or other resources will be saved with the provision of the treatment. This could result in changes in consumption patterns or other behaviors of the subject or other members of the household, which in turn could have an effect on measured outcomes. While such changes should affect the QPM and CM groups equally, they will also affect the generalizability (external validity) of the results to situations in which maize is not provided for free. The effect of any intervention on the measured outcomes could be assessed by having a third treatment group in which subjects receive no intervention. However, this may significantly increase the size and cost of the study.

In many QPM nutritional studies, subjects were not individually randomized to treatments. In these studies it is either not clear whether randomization was done or communities, rather than individual subjects, were randomized to treatments. A higher level of randomization was often chosen as it is easier to manage delivery of the treatment to households. Also, if the treatment was provided as seed or grain, randomization within a community could result in sharing of seed, grain, or food among individuals within the community, thereby reducing any observed treatment effect. This was observed in the Ghana 1 study (Akuamoa-Boateng 2002).

Studies that randomized at the community level involved a low total number of communities. The Mexico and Ethiopia studies, for example, each had two communities per treatment group. With a low total number of communities, it is possible that systematic differences among communities between the two treatment groups with respect to factors that also influence the measured outcomes could bias the observed effect from the intervention. For example, differences among communities in soil fertility or prevalence of plant diseases or pests could affect yields if the maize is delivered in the form of seed. Differences with respect to socioeconomic status, disease status, or other factors could lead to systematic dietary or health differences between communities in the two treatment groups. For example, in the Ethiopia study, systematic differences were observed in the incidence of malaria, which impacted child growth more negatively in the QPM group than in the CM group and diminished any observed treatment effect. Alternatively, greater incidence of malaria in the CM group could have biased an observed treatment effect upward. Systematic differences between treatment groups complicate the attribution of a significant difference in outcomes to the treatment effect. The additional structure in the study design that comes from a higher level of randomization is also often ignored in the analysis of data generated from these studies.

When communities rather than individuals are randomized to treatments, the statistical power to detect a treatment effect increases with the number of communities rather than the number of households or target children within those communities. Therefore, these studies should have a sufficient number of communities, and power calculations to determine sample size should take this planned structure into account. However, it is likely that costs would soon become a limiting factor if efforts were made

to expand a study to a larger number of communities. One option would be first to conduct a pilot study with randomization within a community to evaluate the degree to which seed, grain, and food move among households and how this affects food consumption by target individuals. This will vary by culture, and it is possible that randomization within communities would be successful in some areas.

If a pilot study suggests that randomization within a community would lead to poor adherence of participants to assigned treatments and a higher level of randomization is used instead, household or participant characteristics that could influence measured outcomes could be used as covariates when testing for treatment effects. Analysis of such data at baseline would identify systematic differences between treatment groups that could confound the treatment effect in the planned study.

Some QPM studies also accepted more than one subject per household. This clustering of observations at the household level also diminishes the power of a study. Having only one index child per household is recommended to allow independence among study participants. If a household has more than one eligible child, the study participant could be randomly selected from among the eligible children.

Among the QPM nutritional studies, there was not always clear understanding of what constituted blinding. A study is blinded if the identity of a treatment is not known to the subjects or to the researchers administering the treatment. In these studies, the treatment was the improved protein quality in the maize used by the study, and not the maize variety itself. It will likely be evident to both researchers and subjects that a new variety of maize is being evaluated. However, a study may still be blinded if two distinguishable but unrecognized varieties are used. If a commonly recognized variety is

used (e.g., an already released QPM variety or a locally cultivated CM variety), then blinding on the part of the subjects or researchers is unlikely.

Meta-analyses of the QPM nutritional studies described above indicate a significant effect of QPM on the growth of young children, taking into account the limitations in the design and analysis of these studies. Randomization within these studies may have led to systematic differences in measured outcomes between the treatment groups that were not due to the treatment effect. However, any such systematic differences are expected to be equally likely to favor QPM or CM. Therefore, the observed significant effect of QPM is not believed to be an artifact of the randomization in these studies. Publication bias, or the bias that results when studies reporting positive treatment effects are more likely to be published than studies reporting no or negative treatment effects, is not an issue here, as these meta-analyses were conducted on unpublished studies and care was taken to identify all unpublished work and to include all relevant studies for which sufficient information was available.

The only potential source of bias that can be identified that would favor QPM over CM in this analysis is any lack of blinding that may have occurred in these studies. That is, if participants (or their caregivers) were aware of which treatment was nutritionally improved, greater consumption, better compliance, or other changes in behavior among the participants could result in better outcomes in the QPM group. If the researchers were aware of the identity of the treatments, their behaviors may also change in ways that could result in better outcomes in the QPM group. Potential examples include how anthropometric or other data are taken, what information is provided to participants about the intervention, and the degree of interaction between subjects and

researchers. However, there is no evidence that any lack of blinding led to behavioral changes among subjects or researchers in any of the included studies.

2.5.4. Controlling and Measuring Dose

The frameworks for efficacy and effectiveness illustrated that many factors could modify the dose of the treatment received by subjects. This is increasingly true as the intervention occurs at earlier points on the impact pathway. The confounding factors need to be recognized, described, assessed, and monitored in these studies.

The greatest control over the received dose occurs in efficacy studies, in which the treatment food is consumed under supervision. In effectiveness studies, where behavioral factors can play a significant role, researchers have attempted to control the received dose in various ways. Some studies attempted to control at least some potentially confounding factors through stringent inclusion and exclusion criteria for study participants. For example, the Ethiopia study identified potential participating households by the size of the land holding, the family's duration of residence at that site, family size and number of children, presence of a grain buffer, and other factors. While this may reduce noise in the data that are generated, it also reduces the external validity of the research findings.

One way to deal with lack of control in delivering the intervention to subjects or lack of compliance with the intervention by participating subjects is to exclude subjects who did not receive the desired "dose" of the treatment from subsequent analyses. This was done, for example, in the Ghana 3 study. Instead, an "intent-to-treat" approach is recommended, in which all subjects are analyzed in the treatment groups to which they

were assigned, regardless of the actual dose of the intervention that was received. Further analysis in which the actual dose received by individual subjects is included as a covariate in the data analysis may be informative, as a dose-response relationship, in which increasing dose is associated with increasing effect on measured outcomes, would provide further evidence of an overall treatment effect in the study.

In these studies, the received "dose" may have multiple meanings, all of which are relevant. This term could refer to the amount of maize consumed as part of the treatment, the amount of maize consumed in the total diet, the amount of relevant nutrients (e.g., utilizable protein, lysine, or tryptophan) consumed as part of the treatment, or the amount of relevant nutrients consumed in the total diet. In an efficacy study, measuring the dose received from the treatment would involve measuring the amount of the test meal that is consumed and the amino acid content of that meal. In an effectiveness study, the amount of maize consumed, its source (whether from the treatment or some external source), and its amino acid content are of interest. Dietary intakes are often not assessed or analyzed in QPM effectiveness studies, and therefore little may be known about how much of the dietary intervention is actually consumed by the target individual.

It is not clear how many QPM studies have measured the amino acid content of the maize that is consumed. Such measurement is important to confirm the purity and improved protein quality of the maize used in the study. For example, amino acid analysis of the initial QPM lot in the Mexico study found that the lot was not of sufficient quality to use in the study, and the study was suspended until a new lot of QPM could be obtained. It is also useful to verify that the QPM and CM used in a study do not differ

with respect to other nutrients such as energy, carbohydrates, fats, fiber, or micronutrients. This would confirm that differences in outcomes between the two treatment groups are attributable specifically to the improved protein quality of QPM. As QPM or any biofortified crop is expected to have a nutritional impact by alleviating a nutrient deficiency, it is important to measure intake of relevant nutrients from the improved crop and from the rest of diet. In the case of infants and young children, breastfeeding is an important dietary factor and should be taken into consideration. Intakes of other nutrients (or anti-nutrients) related to the nutrients of interest should also be monitored. In the case of QPM, given the interactions between protein and energy, energy intakes should be assessed.

2.5.5. Data Analysis

Analysis of data from efficacy and effectiveness studies should take into account known structure in the data. This includes factors such as study cycle, age group, and community or location. In the Ghana 3 study, sex and age group were omitted from the data analysis because the QPM and CM groups were roughly balanced with respect to number of children in each sex and age group. However, significant age and sex effects are possible even if studies are balanced with respect to these variables. Accounting for these sources of variation reduced the error variation in the study, increasing the power to detect a treatment effect. Consequently, the treatment effect on weight gain was marginal but not statistically significant as reported by Akuamoa-Boateng (2002) but became highly significant when age group and sex were included in the analysis, as described above.

It is particularly important to take into account non-independence of observations in these studies. Non-independence arises from randomization at the community level, having more than one subject per household, and repeated measurements taken on individual subjects over time. Accounting for correlations among subjects within the same household or community and for correlations among observations from the same subject is necessary to conduct the correct test for a treatment effect.

Data at baseline on study participants or their households should be analyzed for any systematic differences between treatment groups. Dietary data on the target population prior to the start of a study would be valuable in assessing the risk of nutrient inadequacy and level of maize consumption prior to the start of a study. As these studies involve significant resources, the value of conducting a study in a location or involving a population with little risk of protein or lysine inadequacy or little maize consumption should be considered.

If data are collected during the study on additional factors that could impact measured outcomes, these factors could be included as covariates in testing for a treatment effect. This would allow investigation of how these factors influence the outcome and how they interact with and modify the effect of the treatment. Inclusion of these factors in the data analysis may also account for some of the uncontrolled error variation in the study and thereby increase power to detect a treatment effect. Among the factors that could be included in this way, the most interesting would be a measure of consumption or dose received from the treatment. This would allow study of a dose-response relationship for the treatment and would provide evidence directly connecting child growth to consumption of QPM.

Finally, data should be reported on participants who were excluded or dropped out from the study and on the amount of missing data in the study results. The missing data structure should be taken into account when analyzing the results of the study. As was observed in the Ethiopia study, the pattern of missing data may lead to incorrect or misleading conclusions about the treatment effect.

2.5.6. Ethical Considerations

Though they involved research on human subjects, very few of these studies were reviewed and approved by an institutional review board (IRB) or obtained informed consent from participating subjects. Documented IRB approval should be obtained within the country where a study will be conducted and through the principal investigator's institutions. Documented informed consent should be obtained from all participating subjects (or their guardians). Both IRB approval and informed consent should be obtained prior to the start of the study. This is desirable for ethical reasons, and it is also a requirement for publication in the peer-reviewed literature. Lack of attention to ethical considerations will continue to be an obstacle to publishing human nutritional research on QPM.

2.6. Conclusions

This is the first review of efficacy and effectiveness studies on crops that have been genetically improved for nutritional quality. Meta-analyses of these studies indicate a positive effect of QPM on growth of young children. Specifically, consumption of

QPM instead of CM varieties leads to an 8% (95% CI: 4-12%) increase in the rate of growth in height and a 9% (95% CI: 4-12%) increase in the rate of growth in weight in infants and toddlers with mild to moderate undernutrition for whom maize is a significant part of the diet.

The comparable results for height and weight suggest that QPM had a positive impact on growth in height, and growth in weight may be a reflection of the concurrent growth in height. This may be due to QPM consumption having an impact on growth in height but not growth in weight given height. Alternatively, low weight-for-height may not have been a significant problem among the children who participated in these studies, and consequently there may have been little impact of QPM on weight-for-height in these populations.

The new effect size proposed for these meta-analyses and the use of bootstrapping to determine the statistical significance of the results allowed several limitations in the design and analysis of available studies to be taken into account. Based on a conceptual framework that could be readily generalized to other biofortified crops, several recommendations were made for the design and analysis of future studies to evaluate the nutritional impact of QPM. The recommendations made here are also directly applicable to future evaluation of other biofortified crops.

CHAPTER 3. SIMULATING IMPACT AT THE POPULATION LEVEL

3.1. Introduction

As the fourth level in the framework to evaluate the impact of a nutritionally improved crop, the potential impact of QPM in a broader societal context is investigated through simulations. The goal of this chapter is to present methods to simulate the impact of biofortification on nutrient inadequacy in target populations. These methods build on guidelines published by the World Health Organization (WHO) and Food and Agriculture Organization (FAO) of the United Nations to design and plan food fortification programs (World Health Organization 2006). The methods described below modify these guidelines to study the potential impact of a biofortified crop and are illustrated through three simulation studies of the potential impact of QPM on protein and lysine inadequacy in target populations.

The WHO guidelines present a four-step method to determine the desired level of food fortification. First, the distribution of usual nutrient intakes is determined in specific population subgroups, based on quantitative dietary intake data. The population subgroups at greatest risk of inadequate nutrient intakes are identified. For most nutrients, prevalence of inadequate intakes in a population is estimated using the estimated average requirement (EAR) cut-point method, where the prevalence of inadequate intakes in a population is estimated by the proportion of the population with

usual nutrient intakes falling below the EAR of that population (Carriquiry 1999). There are three main assumptions to using the EAR cut-point method to determine the prevalence of nutrient inadequacy: intakes and requirements of a nutrient should be independent, the requirement distribution should be approximately symmetric around the EAR, which is its mean, and the variance of the requirement distribution should be smaller than the variance of the usual intake distribution.

Once the population subgroup that is most at risk for inadequate (or excessive) intakes is identified, the usual amounts of the intended food vehicles for fortification that are consumed by that subgroup are measured. Then, given the usual intakes of nutrients and intended food vehicles in that subgroup, the effect of different levels of fortification of the food vehicle is calculated for the subgroup. The results are used to set goals for food fortification in the target population.

The approach taken in this chapter diverges from the WHO guidelines in two significant ways. First, the goal of the studies described below is to simulate the nutritional impact of an agricultural technology. There is a long pathway of impact leading from the introduction of the technology to impact on nutrient inadequacy, and this pathway will be incorporated into the simulation studies. A general framework for the impact of biofortification on nutrient adequacy is given in Figure 3.1. This framework is similar to the framework for impact of QPM described in Chapter 2, reflecting the generalizability of that framework to other nutritionally improved crops. This framework involves different types of decision-making among different members of a household, not all of whom may be targeted for the technology. The types of decision-

making include production decisions such as adoption, household decisions such as allocation of food, and consumption decisions.

Secondly, rather than simulating the (univariate) distribution of usual intakes in a population, the joint distribution of usual nutrient intakes and nutrient requirements is simulated for a target population. It will be seen that the EAR cut-point method cannot be used to assess nutrient adequacy in these applications because its assumptions will be violated. Changes in the joint distribution after release of the technology will then be simulated while taking into account pathways through which the technology could affect intakes or requirements. Statistics will be identified to quantify the technology's effect on the joint distribution and used to study the interactions between modifying factors and the new technology. The primary outcome of interest is nutrient adequacy at the population level. Improved nutrient adequacy is expected to lead to other improved outcomes for target individuals, households, and higher levels.

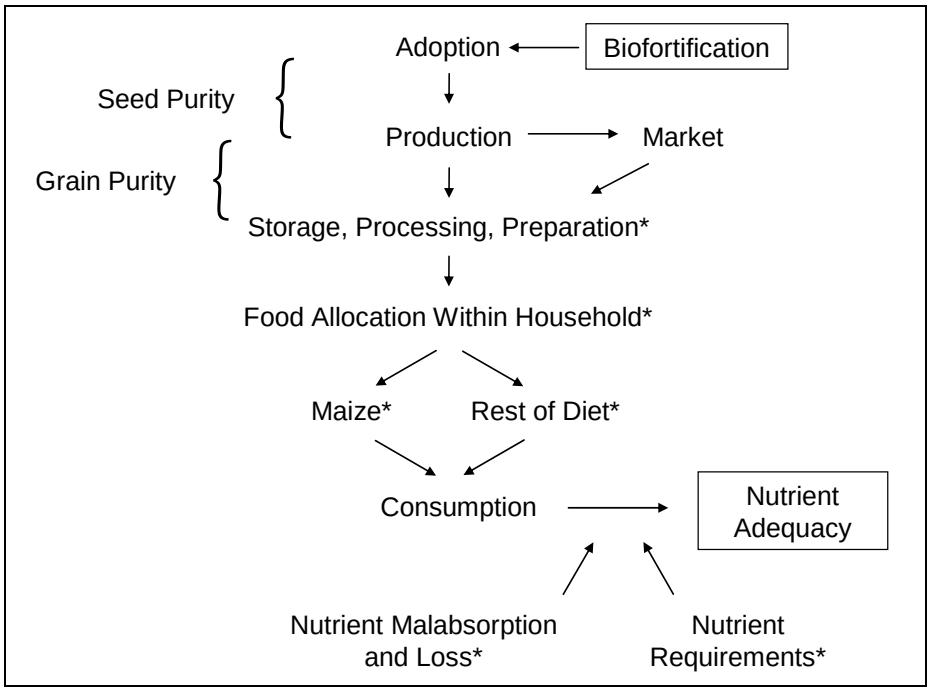

Figure 3.1 General framework for impact of biofortification on nutrient adequacy. Items marked with an asterisk are affected by morbidity of target individuals.

In this chapter, the development of these methods is illustrated by three simulation studies on the potential impact of QPM on protein and lysine adequacy in target populations. The results of these simulation studies illustrate the methods and demonstrate that simulation can be used to identify and study relationships among factors that modify the impact of a biofortified crop. They therefore can be used to identify important considerations in the planning and implementation of biofortification programs. If the same methods were applied to actual rather than simulated diets, these methods could be used make inferences about the potential impact of QPM or other biofortified crops in specific populations.

3.2. Simulation Study 1: Protein Inadequacy and Disease

The first simulation study was based on data from Flores et al. (1966), as reported by Rahmanifar and Hamaker (1999). Data were available on the average utilizable protein intake, age, and weight of toddlers from two rural villages in Guatemala. These data were used to simulate protein requirements in that population according to the 2002 Dietary Reference Intakes of the Food and Nutrition Board (FNB), Institute of Medicine (IOM), based on the average weight of the children and the reference weight for children aged 1-3 years (FNB/IOM 2002). The distribution of average utilizable protein intakes was simulated as a normal distribution with the mean as reported by Rahmanifar and Hamaker (1999) and coefficients of variation (CVs) as suggested by that paper. The simulated intake distribution was independent of the simulated requirement distribution.

Given a joint density of usual intakes and requirements for a nutrient, the prevalence of inadequate intakes in the population can be calculated as the probability that requirement, *R*, exceeds intake, *I*:

$$\Pr(I < R) = \int_0^\infty \int_0^r f_{R,I}(r,i)\,di\,dr,$$

where $f_{R,I}(r,i)$ is the density of the joint distribution of requirements and usual intakes in a population. The above integral of the simulated density was computed by simulating a large number of individuals from that population and computing the proportion whose requirements exceeded their intakes. This yielded an estimate of the prevalence of inadequate intakes in the population as a function of the CV of the usual intake distribution.

Rahmanifar and Hamaker (1999) reported that on average, maize accounted for 58.5% of protein intake in this population. To investigate the potential effect of a disease on adequacy of protein intakes, it was assumed that substitution of CM with QPM in this population would increase utilizable protein intake from maize by 25%. Utilizable protein intakes of simulated individuals were recalculated assuming complete substitution of CM with QPM. The resulting change in utilizable protein intake was different for each individual, as it was a function of initial utilizable protein intake.

The impact of QPM on nutrient inadequacy was quantified by the relative risk (RR) of inadequacy after the introduction of QPM:

$$\text{RR} = \frac{\Pr(\text{inadequate intake with QPM})}{\Pr(\text{inadequate intake without QPM})}.$$

A RR of 1 means that QPM had no effect on nutrient inadequacy, a RR of 0 means that QPM completely eliminated nutrient inadequacy in a population, and an intermediate RR of 0.8, for example, means that introduction of QPM reduced the prevalence of nutrient inadequacy in a population by 20 percent.

The prevalence of inadequate intakes before and after introduction of QPM and the corresponding relative risks are given in Table 3.1 as a function of the CV of the intake distribution and the choice of weight used to simulate the requirement distribution. In general, as the variability of intakes increased in the population, the prevalence of inadequate intakes increased and the relative benefit of QPM decreased (i.e., the relative risk increased). This suggests that variability of nutrient intakes in a population may be a factor that would modify the impact of QPM at the population level. The initial prevalence of inadequacy must be considered when interpreting relative risks, as a high

relative risk may still indicate a substantial decrease in nutrient inadequacy when the initial prevalence of inadequacy was also high. The use of relative risk to quantify the impact of QPM on nutrient inadequacy is investigated further in simulation study 2.

As children in this study on average weighed less than the reference weight for their age and as protein requirements are determined by body weight, protein requirements calculated using the children's average weight were lower than protein requirements calculated using the reference weight. Consequently, intakes appeared more adequate when actual weights were used to simulate requirements in this population. The simulation method used here was able to accommodate either method to calculate requirements in the population.

Table 3.1 Impact of substitution of CM with QPM on prevalence of inadequate intakes in simulation study 1. Simulated intakes and requirements were consistent with data on toddlers from two villages in rural Guatemala, as reported by Rahmanifar and Hamaker (1999). Requirements were simulated using the reported average weight and the reference weight (FNB/IOM 2002) for children aged 1-3 years.

	Reference Weight			Average Weight		
CV of Intake	Inadequate Intakes (%)		RR of Inadequate	Inadequate Intakes (%)		RR of Inadequate
Distribution	Before QPM	After QPM	Intake After QPM	Before QPM	After QPM	Intake After QPM
0.15	30	12	0.40	5	1	0.26
0.20	33	17	0.51	10	4	0.41
0.25	36	21	0.59	14	8	0.53

The prevalence of disease can affect nutrient adequacy in a population by increasing nutrient requirements or decreasing nutrient intakes through qualitative or quantitative changes in individuals' diets. The simulations described above were repeated, allowing for varying levels of disease in the population. Disease was

incorporated into the simulations by allowing each simulated individual to have a certain probability of disease. This probability may be the same for all individuals in a population or it may depend on an individual's characteristics (e.g., whether an individual had an adequate nutrient intake prior to getting a disease). To investigate the potential effect of disease, it was assumed that if an individual we diseased, the individual's protein requirement would increased by 20%. This figure was chosen to illustrate the simulation method, and impact of choosing other values could easily be explored using these methods.

As the probability of disease increased in the population, and requirements increased while intakes remained the same, the prevalence of inadequate intakes prior to QPM substitution naturally increased (Figure 3.2). Meanwhile, the impact of QPM, as measured by relative risk, decreased. However, variability in intakes again played a significant role, both on the magnitude of the effect of disease on the prevalence of inadequate intakes prior to QPM substitution and on the impact of QPM at any given level of disease prevalence.

These relationships were explored further by allowing intakes to decrease if an individual is diseased or by allowing the probability of disease to be higher among individuals who already had inadequate intakes prior to getting a disease. The impact of QPM under these scenarios and with varying probability of disease is depicted in Figure 3.3. The impact of QPM, as measured by relative risk, was reduced when disease was

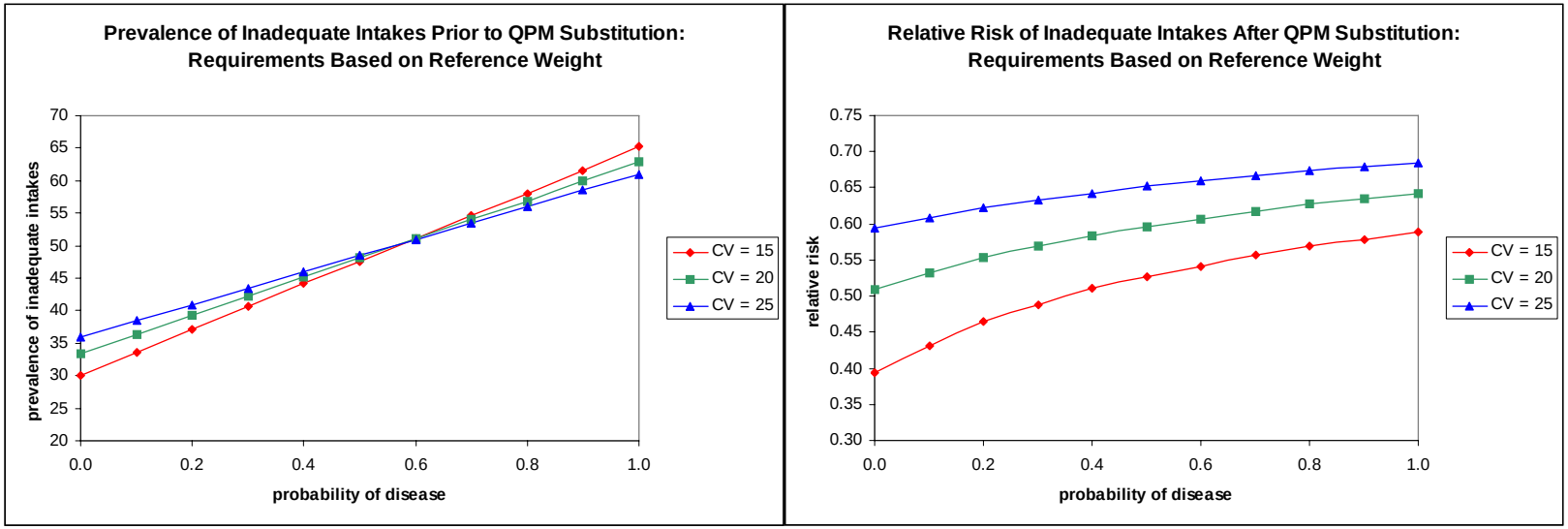

Figure 3.2 Impact of disease prevalence on prevalence of inadequate intakes prior to QPM substitution and relative risk of inadequate intakes after QPM substitution in simulation study 1.

more likely among individuals who initially had inadequate intakes and when disease led to an overall decrease in protein (or food) intake. Again, these results should be interpreted while taking into consideration prevalence of inadequacy prior to QPM substitution.

These results reflected the fact that while nutrient inadequacy may increase with disease, a food-based intervention may have little impact if disease leads to qualitative or quantitative changes in intake. That is, if a child who has fallen ill eats less, is given less food, or is given different foods than what is typically consumed, then an intervention that relies on consumption of a typical diet may not have significant impact. It is interesting to note that if a disease both decreases nutrient intakes and increases nutrient requirements, then intakes and requirements become negatively correlated in populations where that disease is prevalent. This suggests that the EAR cut-point method may not be appropriate to estimate the prevalence of inadequate intakes in a population in this situation. However, intakes and requirements are independent given disease status. Therefore, the EAR cut-point method may be applicable if the prevalence of inadequate intakes is assessed by conditioning on disease status. While there are known relationships between nutrition and disease (Jackson and Calder 2004), disease status is not considered when assessing nutrient adequacy in a population. The methods used in this simulation can be applied to assess nutrient adequacy given disease in populations where dietary and morbidity data are available and the effects of a disease on nutrient requirements are known.

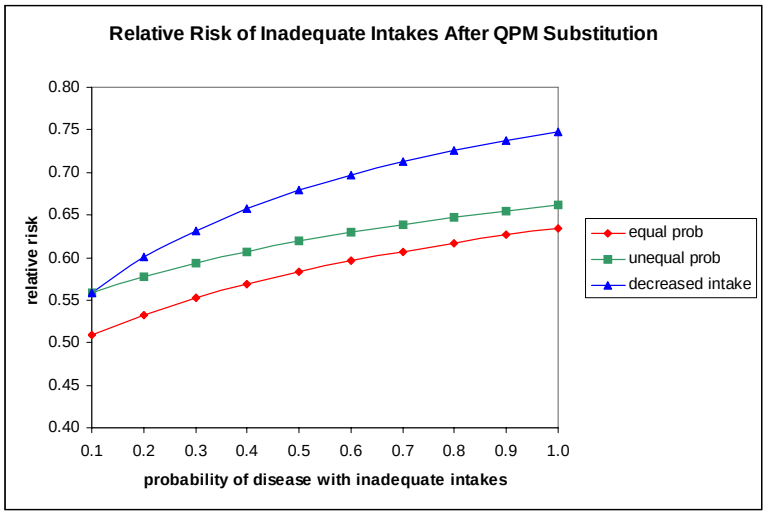

Figure 3.3 Relative risk of inadequate intakes after QPM substitution as a function of disease prevalence. Disease occurred with equal probability among all individuals ("equal prob") or with greater probability among individuals with inadequate intakes prior to disease ("unequal prob"), or incidence of disease led to decreased nutrient intakes ("decreased intake") with equal probability of disease among all individuals.

3.3. Simulation Study 2: Lysine Inadequacy and Seasonality

The second simulation study focused on the potential impact of QPM on lysine inadequacy among adults in Kenya. Simulation study 1 modeled the impact of QPM directly on a population's usual nutrient intake distribution. In this section, the modeling of intakes was improved to study the effects of variability in the consumption of individual foods on the potential impact of QPM. In particular, the effect of seasonal variation in bean consumption on the impact of QPM was explored.

Average daily consumption of maize, other cereals, milk, and beans for selected districts of Kenya was provided by H. De Groote and O. Shadrack (unpublished) (Table 3.2). These values were calculated from food expenditure data collected by the Welfare Monitoring Survey III (WMSIII), conducted in Kenya in 1997 (Ministry of Finance and

Planning 2000a; Ministry of Finance and Planning 2000b). These data provide a sample of likely average consumption values for this QPM target area, and they will be used to illustrate the methods described below. The results that are generated will be useful to study potential seasonal variation in the impact of QPM but should not be used to indicate actual levels of lysine inadequacy in Kenya.

Pellett (1996) developed the following equation to estimate lysine intake from three major food groups: animal source foods, cereals, and pulses including soybeans:

$$\text{Lysine (mg/day)} = 86.3(\text{AP}) + 19.8(\text{CP}) + 63.6(\text{PSP}) + 599,$$

where AP is animal protein (g/day), CP is cereal protein (g/day), and PSP is protein from pulses including soybeans (g/day). Table 3.2 provides food consumption values for maize, other cereals, milk, and beans. The protein content per gram of each of these food items was obtained from the USDA National Nutrient Database for Standard Reference (USDA 2005). Average protein contributions from non-milk animal products, non-bean pulses, and soybeans were obtained from the 1997 Food Balance Sheet for Kenya (FAO 2005). These data were used in the equation provided by Pellett (1996) to calculate the mean daily lysine intake in each district. It was then assumed that the lysine intake distribution in each district was normally distributed with a mean as calculated and a coefficient of variation of 20%. All lysine intakes simulated from these distributions were reduced by 10% to account for digestibility.

Table 3.2 Estimated average consumption of selected foods in districts of Kenya (provided by H. De Groote and O. Shadrack).

		Food consumption (gram/person/day)			
Province	District	Maize	Milk	Beans	Other Cereals
Central	Kiambu	346	115	44	56
Central	Kirinyaga	320	76	95	76
Central	Muranga	322	119	85	55
Central	Nyandarua	342	124	47	39
Central	Nyeri	343	99	86	56
Coast	Kilifi	497	11	24	14
Coast	Kwale	318	19	21	40
Coast	Lamu	271	57	63	73
Coast	Taita Taveta	290	67	46	25
Coast	Tana River	289	148	36	148
Eastern	Embu	194	83	82	39
Eastern	Kitui	297	37	98	43
Eastern	Machakos	368	38	56	16
Eastern	Makueni	306	38	37	24
Eastern	Mbeere	358	39	108	44
Eastern	Meru	228	84	73	53
Eastern	Tharaka	251	57	98	105
Nyanza	Homa-Bay	106	41	10	115
Nyanza	Kisii	289	62	27	45
Nyanza	Kisumu	176	45	18	41
Nyanza	Migori	264	23	26	57
Nyanza	Siaya	206	29	33	80
Rift Valley	Baringo	328	96	50	58
Rift Valley	Bomet	167	174	23	78
Rift Valley	Elgeo-Marakwet	281	69	59	46
Rift Valley	Kajiado	375	244	74	40
Rift Valley	Kericho	233	127	31	46
Rift Valley	Laikipia	418	85	87	33
Rift Valley	Nakuru	326	73	48	37
Rift Valley	Nandi	239	104	30	27
Rift Valley	Narok	419	59	27	24
Rift Valley	Trans-Mara	337	80	45	19
Rift Valley	Trans-Nzoia	281	164	38	17
Rift Valley	Uasin Gishu	272	120	48	29
Rift Valley	West Pokot	234	67	28	30
Western	Bungoma	246	48	38	26
Western	Busia	230	15	14	63
Western	Kakamega	173	55	29	24
Western	Vihiga	347	60	18	5

To study the effects of seasonality in bean consumption, two seasons were simulated: a "post-harvest" season in which bean consumption in each district was assumed to be 25% greater than the value given in Table 3.2 and a "pre-harvest" season in which bean consumption in each district was assumed to be 25% less than the given value. Lysine intake distributions were then calculated for each season and district as described above.

The distribution of lysine requirements was simulated as follows. Reference body weights are 57 kg for women and 70 kg for men, and the EAR for lysine for adults 19 years or older is 31 mg/kg/day (FNB/IOM 2002). The mean of the requirement distribution was therefore taken to be the product of the EAR and the reference weight averaged over the sexes. The requirement distribution was assumed to be normally distributed with a coefficient of variation of 12%. This distribution was simulated independently of the intake distribution. As with other aspects of these simulations, the parameters above were chosen to illustrate the simulation method. Other methods to simulate the requirement distribution could be evaluated using the same methods.

It was assumed that the lysine concentration in maize protein increased from 28 mg/g protein to 45 mg/g protein with the substitution of CM with QPM. While 45 mg/g protein is reasonable for QPM varieties, there are efforts to increase this improvement by selecting for loci that modify lysine and tryptophan levels in the endosperm (Krivanek et al. 2007). The potential impact or benefits of these efforts could also be explored through simulation. Protein content of QPM was assumed to be the same as that of CM, and only complete substitution of CM with QPM was considered. In addition to the relative risk,

the difference in percent inadequacy before and after the introduction of QPM was also calculated as a measure of the impact of QPM on nutrient adequacy in the population.

Relative risk of lysine inadequacy and change in lysine adequacy after QPM substitution are plotted against lysine inadequacy before QPM substitution in Figures 3.4 and 3.5, respectively. The impact of QPM is clearly related to the degree of inadequacy before its introduction. Using relative risk as a measure of impact, it appears that QPM has less of an impact (a higher relative risk) when lysine inadequacy was already high in a population. However, this is misleading as a high relative risk can still mean a significant drop in inadequacy when initial inadequacy is high. This is clarified in Figure 3.5, in which impact of QPM is measured in absolute terms as the change in lysine inadequacy after the introduction of QPM. These results do not indicate any seasonal effect on QPM impact, given the level of lysine inadequacy before QPM introduction. However, there is a seasonal effect on lysine inadequacy before QPM introduction, leading to an indirect effect on the impact of QPM. This is explored further below. The district of Homa-Bay appeared to be an outlier in many of the simulations in the section. Food expenditures were generally low in that district, with very low bean consumption and greater dependence on cereals other than maize.

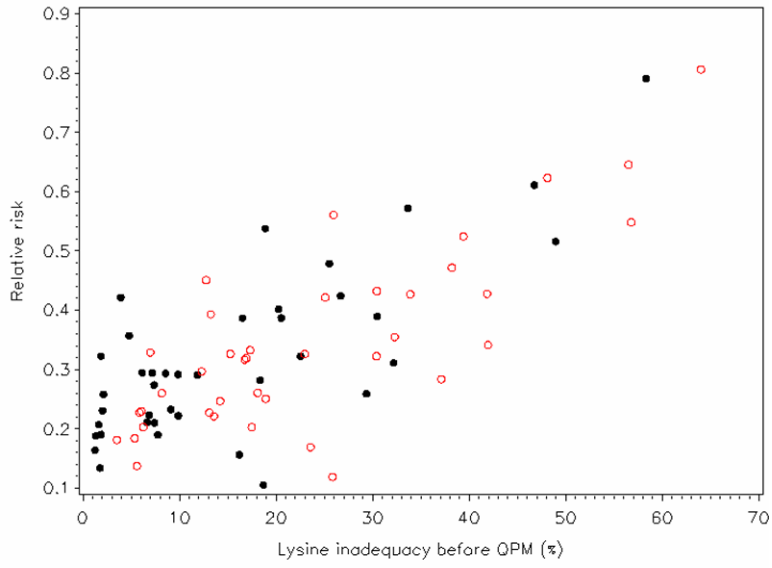

Figure 3.4 Relationship between lysine inadequacy before QPM substitution and relative risk of lysine inadequacy after QPM substitution in simulation study 2. Each district is represented by two points, with a red open circle indicating the pre-harvest season and a black closed circle indicating the post-harvest season.

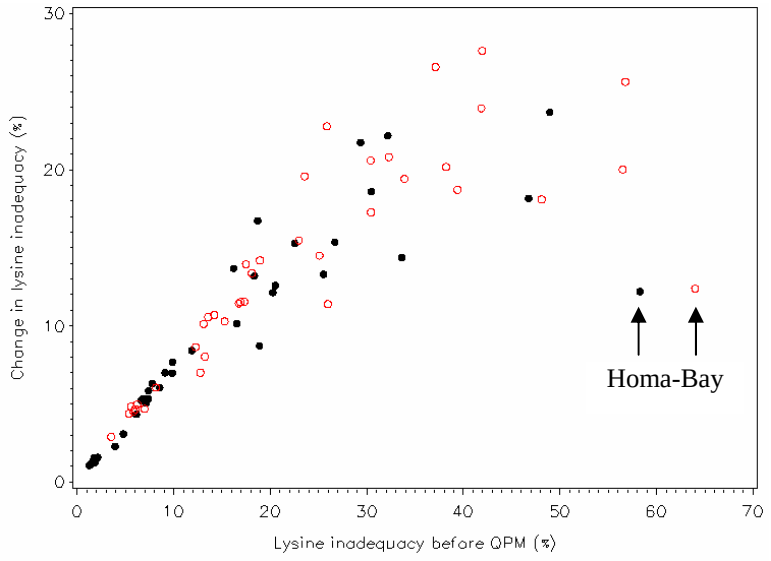

Figure 3.5 Relationship between lysine inadequacy before QPM substitution and change in lysine inadequacy after QPM substitution in simulation study 2. Each district is represented by two points, with a red open circle indicating the pre-harvest season and a black closed circle indicating the post-harvest season.

There is a clear relationship between the relative risks of lysine inadequacy after QPM substitution in the pre- and post-harvest seasons (Figure 3.6). A similar relationship appears when change in inadequacy is used to measure the impact of QPM (Figure 3.7). These relationships indicate that if QPM is likely to have a relatively high impact in one season, it is also likely to have a relatively high impact in the other season. In 37 of the 39 districts, the relative risk was lower in the post-harvest season. However, in absolute terms as the change in the prevalence of lysine inadequacy after QPM substitution, the impact was greater in the pre-harvest season for all districts.

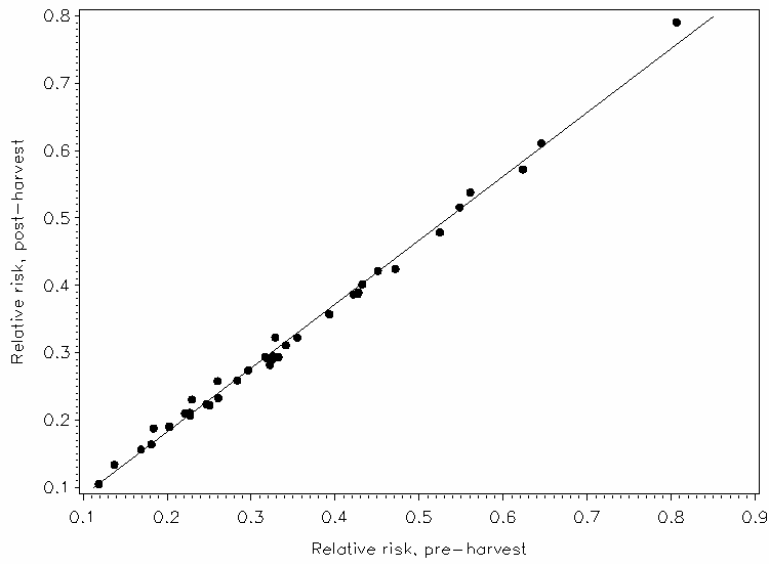

Figure 3.6 Relationship between the post- and pre-harvest seasons for relative risk of lysine inadequacy after QPM substitution in simulation study 2.

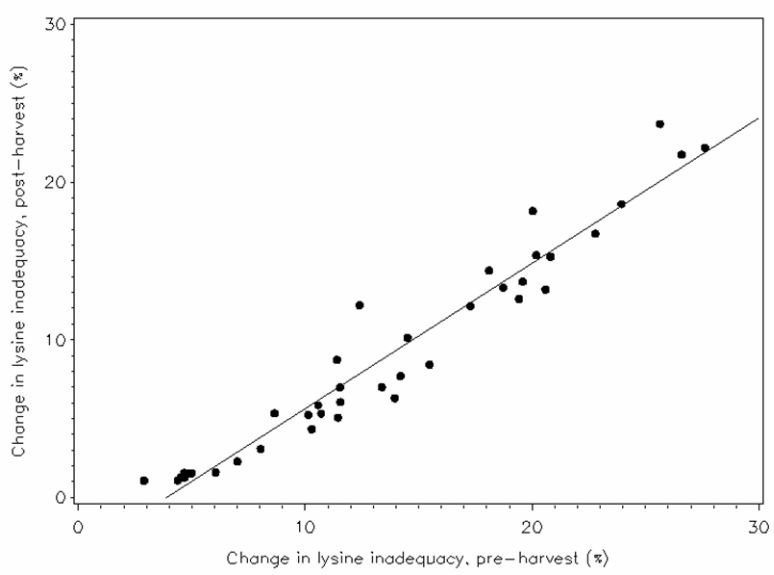

Figure 3.7 Relationship between the post- and pre-harvest seasons for change in lysine inadequacy after QPM substitution in simulation study 2.

In Figure 3.8, the change in lysine inadequacy across seasons is plotted in the presence and absence of QPM. In this Figure, the vertical axis can be interpreted as a measure of seasonal variation in lysine inadequacy. These results indicate that QPM substitution reduces seasonal variation in lysine inadequacy over a range of dietary profiles, as represented by the districts under study. This is a new mechanism for QPM or any biofortified crop to have a nutritional impact. Seasonal variation in inadequacy may be another useful measure of impact for biofortified crops.

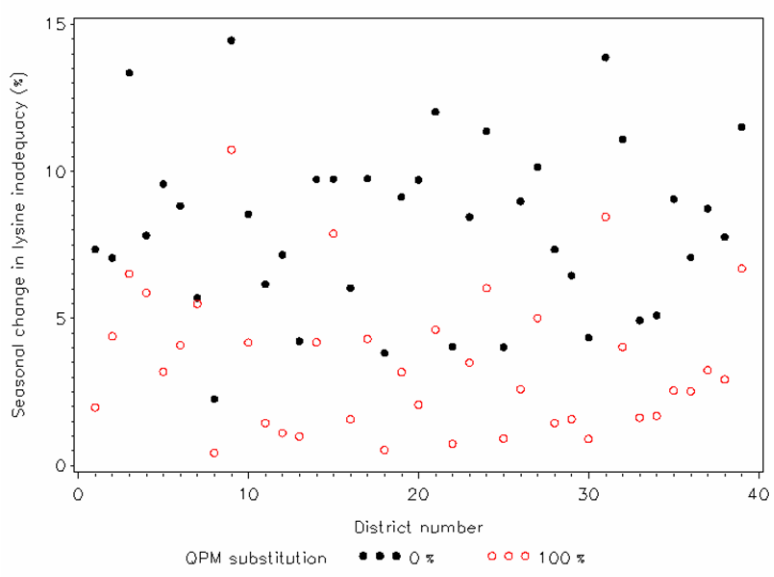

Figure 3.8 Seasonal change in lysine inadequacy before (black) and after (red) QPM substitution in simulation study 2. Values are plotted against a district index.

In Figure 3.9, the change in lysine inadequacy across seasons is plotted against lysine inadequacy in the post-harvest season, with and without complete QPM substitution. Higher levels of lysine inadequacy are associated with greater seasonal variation in inadequacy. Although it appears that QPM introduction decreases seasonal variation in lysine inadequacy, it also decreases lysine inadequacy in general.

These simulations demonstrate that the potential impact of QPM may vary seasonally, and, in particular, it may depend on seasonal variation in the consumption of other foods that are high in quality protein or lysine. It is also clear from these examples that patterns of consumption of other components in the diet can influence the impact of QPM or of any biofortified food. This suggests that monitoring the total diet, and not just the intake of biofortified foods, is important in evaluating the impact of a biofortified crop.

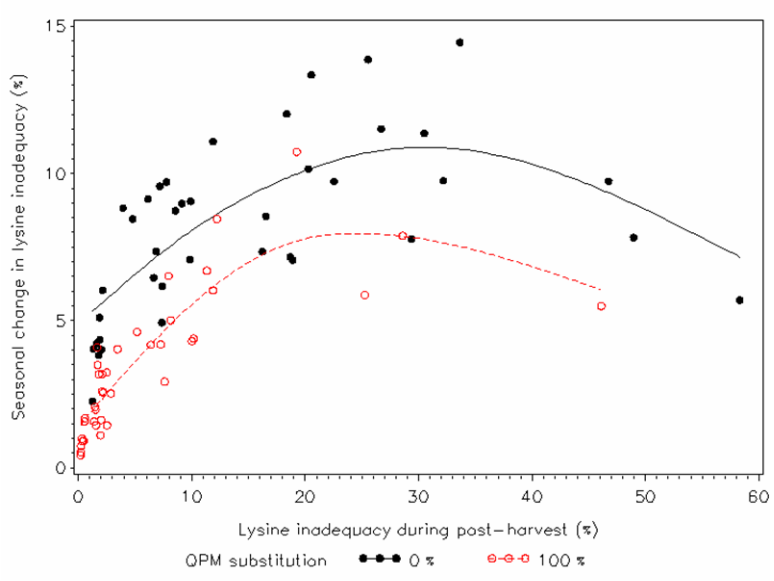

Figure 3.9 Seasonal change in lysine inadequacy before (black) and after (red) QPM substitution in simulation study 2. Values are plotted against lysine inadequacy in the post-harvest season.

3.4. Simulation Study 3: Adoption and Production Patterns

Simulation study 3 was also based on data from the WMSIII. Significant changes were made to the modeling of diets, characteristics of the simulated population, and individuals' source of QPM. To simulate individuals' diets, six dietary components were considered: maize, other cereals, milk, beans, other animal protein, and other pulses and soy protein. Diets within a population were simulated from a six-dimensional multivariate normal distribution. The first four components of the mean of that distribution were drawn from the district-level data given in Table 3.2. The mean values used for other animal protein and other pulses and soy protein were taken from the 1997 Kenyan Food Balance Sheets and were assumed to be the same for all districts (FAO 2005).

Distributions of consumption of individual food items were assumed to have a 20% CV, except for maize and beans, which had a 10% CV. The correlations among levels of consumption of different food items were assumed to be equal to the correlations among the district means given in Table 3.2. Correlations of intakes of other animal protein or other pulses and soy protein with the other dietary components were set to zero. Individuals' lysine intakes were calculated from their simulated diet using the formula proposed by Pellett (1996). This method to determine a distribution for simulating diets was chosen to illustrate the simulation outcome. Diets could be simulated in other ways, or they could be drawn from distributions estimated from dietary intake data of a population of interest.

Separate weight distributions were simulated for men and women, and individuals' simulated weights were used to determine their lysine requirements. The sex ratio in all simulated populations was set to 0.5. Total lysine intakes were reduced by 10% to account for digestibility. The effect of disease was investigated by increasing the lysine requirements of individuals in the bottom 5% of their weight|sex distribution to 44 mg/kg/d. This figure was based on research indicating that lysine requirements of chronically undernourished men are approximately 50% higher than those of well-nourished individuals (Kurpad et al. 2003).

There were three types of people in a simulated population: non-maize producers, maize-producing QPM non-adopters, and maize-producing QPM adopters. Each individual could consume varying amounts of three types of maize grain: grain produced by the QPM adopter, grain produced by the QPM non-adopter, and grain purchased from the local market. Each type of maize grain had some amount of QPM in it. Grain

produced by the QPM adopter was assumed to be high in QPM, although it did not have to be purely QPM; grain produced by the QPM non-adopter was assumed to contain no QPM; and grain purchased on the local market was assumed to have little or no QPM. Every individual in the simulation also had a value indicating their degree of self-sufficiency in maize, i.e., what proportion of their total maize consumption came from home production. This proportion was expected to be relatively high among maize producers and zero among non-producers.

Figure 3.10 illustrates the relationship between initial prevalence of lysine inadequacy and change in lysine inadequacy as a result of QPM introduction. In these simulation results, all individuals were maize producers and there was 100% adoption of QPM. Maize produced by adopters was 90% QPM and 10% CM, thereby allowing impurity in harvested and consumed grain that could have arisen through impure seed, pollen contamination, mixture during storage, or other sources. Maize purchased from the local market was 10% QPM and 90% CM, indicating a situation where QPM was available from the market but appeared in low levels or in mixtures with CM. Each simulated individual was independently assigned a degree of self-sufficiency in maize that ranged between 60 and 100%.

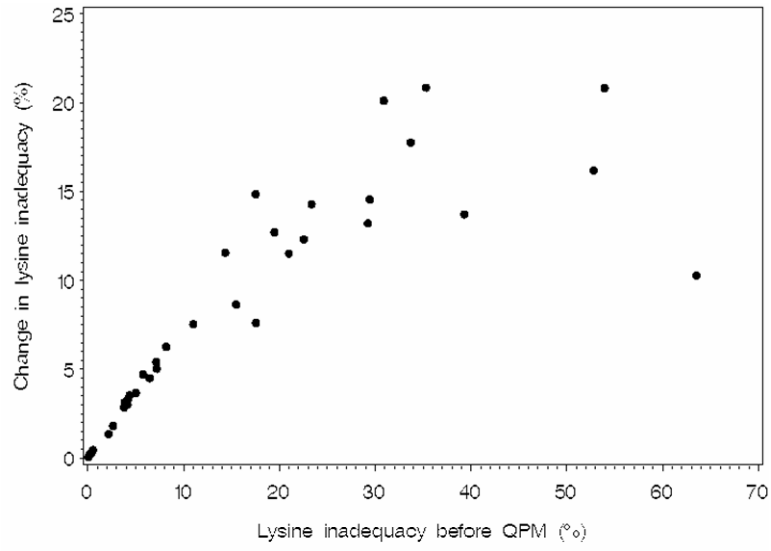

Figure 3.10 Relationship between initial prevalence of lysine inadequacy and change in lysine inadequacy as a result of QPM introduction among maize producers with 100% QPM adoption and other input parameters as described in the text. Each plotted point represents a dietary profile consistent with average diets in a given district of Kenya.

For some dietary profiles, lysine inadequacy before QPM introduction and change in lysine inadequacy after QPM introduction were low, indicating scenarios of little nutritional need or impact. For other dietary profiles, inadequacy before QPM introduction and change in inadequacy after QPM introduction were both high, indicating a nutritional need that could be met by QPM. These results suggest that QPM could have a potential impact in some but not all scenarios, overall diet in a population would be a significant determinant of impact, and targeting of QPM introduction based on dietary data from a population would allow the technology to be used efficiently in meeting nutritional needs without spending resources to develop and disseminate QPM varieties adapted to areas where they would be of little benefit.

Figure 3.11 plots potential impact, as measured by change in lysine inadequacy, versus average maize consumption and average bean consumption, given the same inputs as described above. Though the technology is an improvement in the lysine content of maize, there is no apparent relationship between impact and average amount of maize consumption among the dietary profiles considered. However, there is a clear relationship between impact of QPM and average consumption of beans, which are relatively high in lysine content. This interesting result again highlights the importance of assessing the total diet of a population in targeting and impact assessment.

The results above assume that all individuals in a population are QPM-adopting maize producers. In Figure 3.12, the impact of QPM, as measured by change in lysine inadequacy, is simulated as a function of the proportion of the population engaged in maize production and the QPM adoption rate among those producers. In these simulations, diets were consistent with the average diet in Makueni District, Kenya. If all individuals are maize producers with 100% adoption, lysine inadequacy in this scenario is expected to drop by 12.7% . With a 40% adoption rate, lysine inadequacy would drop by 5.4%, and with 20% adoption, lysine inadequacy would drop by only 2.9%. These values would in general decrease as a smaller proportion of the population engaged in maize production, or comparably, as more households ran out of maize stocks in a given year and individuals effectively became non-maize producers. It is interesting to note in Figure 3.12 that when adoption rates are low (e.g. 10%, a plausible value), impact is greater when a smaller proportion of individuals are engaged in maize production. This

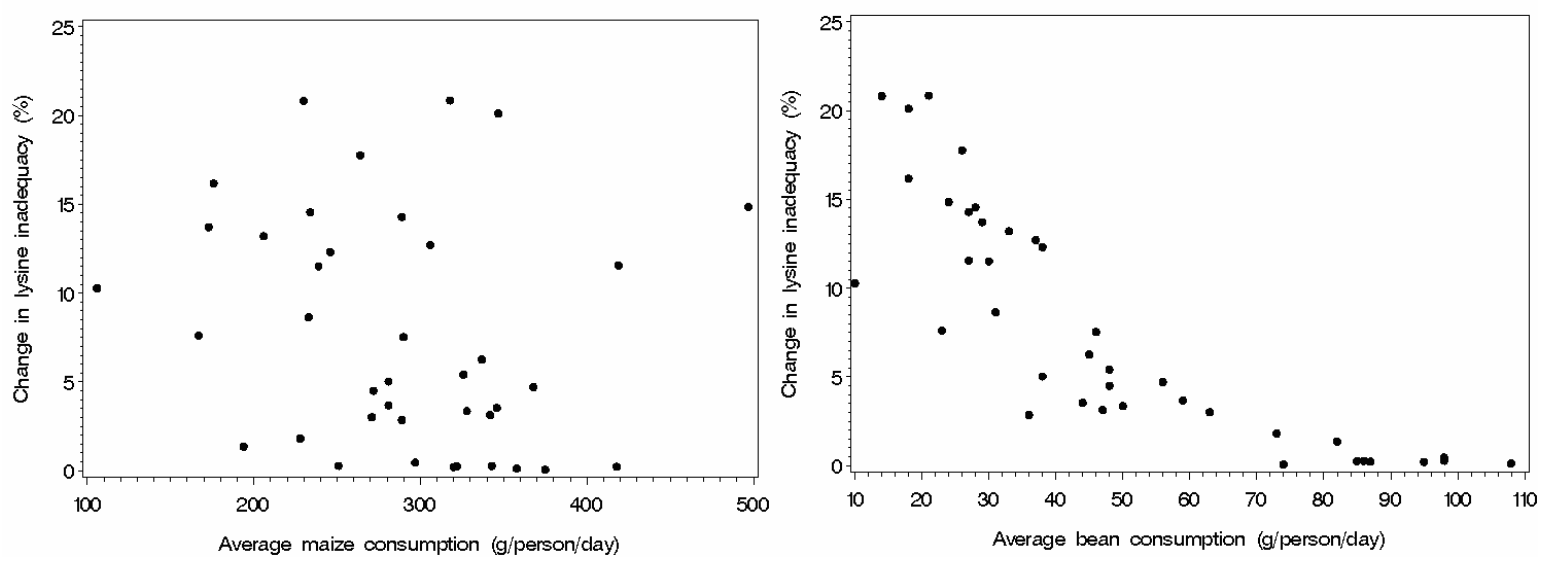

Figure 3.11 Relationship between change in lysine inadequacy as a result of QPM introduction and average maize and bean consumption among maize producers with 100% QPM adoption and other input parameters as described in the text.

result occurs because non-maize producers are entirely dependent on maize from the local market, which contained low levels of QPM. Meanwhile, non-adopting maize producers rely on their own production, which contains no QPM, for the majority of their maize consumption. The additional features of these simulation models presented useful tools to study the roles of and interactions among factors that may modify the effect of QPM. From the simulations in this section, it is clear that patterns of adoption and production affect the potential impact of the improved crop and that these patterns may modify impact in complex ways.

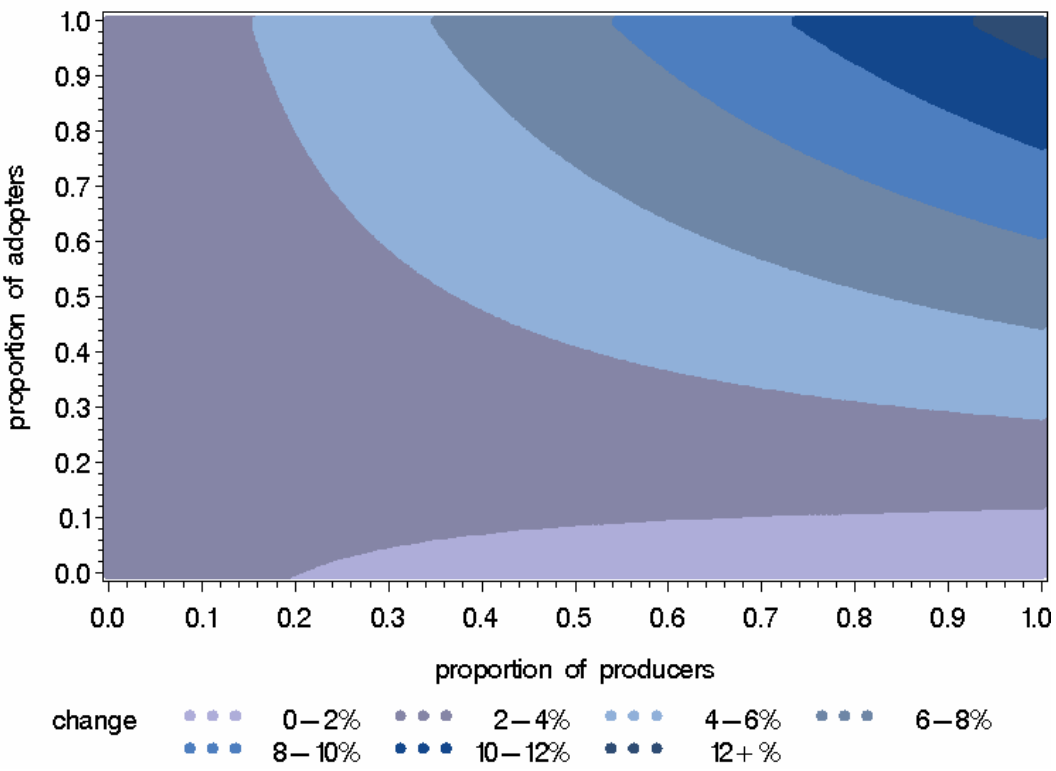

Figure 3.12 Change in lysine inadequacy as a function of the proportion of maize producers in a population and the QPM adoption rate among the producers. Simulated diets were consistent with data on the average diet in Makueni District, Kenya.

Currently, southern Ghana is believed to have the highest rate of QPM adoption, estimated at around 40% (Dankyi et al. 2005). This is a reasonable estimate for an upper bound on adoption rates. When considering release of QPM to a new area, it would be worthwhile to estimate to the potential impact on inadequacy in that area. This would aid decision-making about targeting the technology to areas where it is likely to have an impact and would help to prevent use of resources in directing the technology to areas where it is not needed. In evaluating potential impact, it would be necessary to take into account reasonable expectations of adoption rates.

A similar approach using other nutrients could be used for other biofortified crops. If two biofortified varieties, each improved for a different nutrient, are released in the same area where there is risk of inadequacy for both nutrients, then it is likely that the adoption rate of each individual variety will be relatively low. Consequently, impact on the inadequacy of both nutrients may be relatively low. In such a situation, it would be worthwhile to pyramid both nutritional traits into one improved variety. If lysine or quality protein is one of the traits for which there is a nutritional need, it would make sense to integrate QPM breeding and other biofortification efforts.

3.5. Conclusions

The simulation-based methods described in this chapter are useful tools to study the potential impact of QPM at the population level. These methods can be adapted to study diverse scenarios as well as to study other biofortified crops. Although diets were simulated to illustrate the methods above, simulated diets could be substituted with actual

diets quantified from dietary intake data and used to inferences about specific populations.

These methods take into account the mechanisms through which QPM or another biofortified crop could have a nutritional impact. They allow study of relevant factors that may modify the impact of a new crop at the population level. Potential impact vary by adoption and production patterns, composition of the total diet, variation in food consumption patterns, self-sufficiency in and sources of maize, disease, and other factors, as well as seasonal effects on all of the above. These factors should be considered and monitored in the planning and implementation of biofortification programs.

Different statistics were used to quantify impact of the technology on nutrient adequacy. Relative risk of inadequacy after introduction of a biofortified crop led to incorrect conclusions in some cases, and change in the prevalence of inadequacy appeared to be a more useful indicator to evaluate the impact of a biofortified crop. Variation in nutrient inadequacy also appeared to be a potentially useful indicator, particularly in investigating seasonal effects, as biofortified crops may reduce not only nutrient inadequacy but also variation in nutrient inadequacy.

The simulation-based methods in this chapter significantly build on existing guidelines to simulate impact, in particular by taking into account potential impact pathways. These methods allow quantitative discrimination between scenarios of high and low impact and can be applied generally to biofortified crops. They also allow identification of factors that should be monitored in the planning and implementation of biofortification programs. In particular, they highlight the importance of assessing total

diet and morbidity in a target population, which may not otherwise be assessed when evaluating the impact of an agricultural technology.

CHAPTER 4. CONCLUSIONS

In this dissertation, a four-level framework was proposed to evaluate evidence on the nutritional impact of QPM. A search of the peer-reviewed literature identified several studies that have demonstrated increased bioavailability of protein in QPM and *o2* maize grain, compared to protein in CM grain. However, there were no published studies that demonstrated QPM's efficacy, effectiveness, or impact in a broader context. Unpublished efficacy and effectiveness studies on QPM and *o2* maize were identified and a meta-analysis was conducted to assess the effect of QPM on the growth of young children. The results indicate that consumption of QPM instead of CM leads to an 8% (95% CI: 4-12%) increase in the rate of growth in height and a 9% (95% CI: 4-12%) increase in the rate of growth in weight in infants and toddlers with mild to moderate undernutrition for whom maize is a significant part of the diet.

The studies used to derive these estimates had methodological limitations, several of which were addressed through the effect size developed for the meta-analysis and the use of the bootstrap to assess statistical significance of the results. Further community-level nutritional studies were recommended to provide stronger evidence on the efficacy and effectiveness of QPM. Conceptual frameworks were described and several recommendations were made for the design and analysis of such studies. These frameworks and recommendations are also directly applicable to future evaluation of

other biofortified crops. The first part of this dissertation presented the first systematic review of efficacy and effectiveness studies on crops that have been genetically improved for nutritional quality.

The simulation-based methods described in the second part of this dissertation proved to be useful tools to study the potential impact of QPM at the population level. These methods took into account the mechanisms through which QPM or another biofortified crop could have a nutritional impact and allowed study of factors that may modify the impact of a new crop at the population level. The simulations indicated that impact may vary by adoption and production patterns, composition of the total diet, variation in food consumption patterns, self-sufficiency in and sources of maize, disease, seasonal patterns, and other factors. These factors should be considered and monitored in the planning and implementation of biofortification programs. The methods described here allow quantitative discrimination of scenarios of high and low impact and can be adapted to study diverse scenarios as well as to study the potential impact of other biofortified crops. QPM is a valuable model for biofortification as the data requirements and methodologies to evaluate nutritional impact will be similar for all nutritionally improved crops. Challenges like those encountered with QPM are also likely in breeding, targeting, dissemination, and impact assessment in biofortified crops. The over 40 years of QPM research can significantly inform current efforts in biofortification.

In addition to questions about nutritional impact, other questions have been raised about QPM over the years that have not been completely answered. Criticism of the priority given to QPM in maize breeding persists in part because of the lack of attention to these concerns. Perhaps the largest concern that has been voiced is whether there still

exists a yield penalty from inclusion of the quality protein trait. Recent data comparing yields of QPM and CM varieties, controlling for environment and type of germplasm (e.g., hybrid, open pollinated variety, or synthetic), could be analyzed to address this concern. Generally, QPM varieties are competitive with CM varieties in many tropical environments (Bjarnason and Vasal 1992; Pixley and Bjarnason 1993), though small yield gaps persist in some countries (K. Pixley, personal communication).

While the goal has traditionally been to increase QPM yields to the level of CM yields, QPM yields may have to exceed CM yields to facilitate adoption in at least some areas. Otherwise, farmers may choose to continue cultivation of a familiar variety, rather than adopt a new variety with no visible benefit. This may have happened in Central America, where recently developed QPM hybrids could not displace older CM hybrids in the area (H. Córdova, personal communication). Questions also remain on whether QPM still has greater susceptibility than CM to diseases or storage pests. Differences for agronomic traits other than yield are mostly case-specific, although in general, QPM varieties may still be more susceptible to ear rots and post-harvest insect damage than CM varieties (K. Pixley, personal communication). Again, existing data could be analyzed to address these concerns.

Another concern that has not been adequately addressed is whether pollen contamination could lead to significant loss of the quality protein trait in farmers' fields, particularly given the small plot sizes of the farms that are likely to be targeted with QPM. It is said that pollen contamination and loss of the quality protein trait are not as serious a concern as was originally believed; however, no data have been published to support that statement. Evaluation of the impact of pollen contamination should occur in

typical farmers' fields, ideally under farmer management, rather than in field tests conducted by researchers. Studies of gene flow from transgenic maize could also be used to address this issue. One proposed solution to the problem of pollen contamination has been to push for complete adoption of QPM in a target area. However, this appears neither feasible nor desirable.

There appear to be fewer remaining concerns about the acceptability of QPM grain for the preparation of commonly consumed foods. However, investigations on acceptability of QPM often had methodological limitations as many of these studies evaluated target individuals' opinions of one QPM variety compared to one CM variety. These studies often attributed any difference in acceptability between the two varieties to the quality protein trait and associated these differences with the comparison of QPM and CM in general.

Finally, costs associated with QPM breeding have not always been clear. Extra cycles of breeding are required to introgress the quality protein trait into a new variety as the *o2* allele must be fixed prior to selection for modifying loci (Krivanek et al. 2007). However, conversion of existing varieties to QPM is not the only breeding method that is used, and other methods may require less time, depending on the availability of elite QPM germplasm. Laboratory screening of amino acid content is also costly, though not required for all stages of breeding (De Groote et al. 2006). Although genetic diversity of QPM germplasm has been significantly increased, it is not clear how much diversity exists among released QPM varieties. In Sub-Saharan Africa, for example, the majority of released QPM varieties are derived from maize population Across 8363SR (Krivanek et al. 2007).

Nevertheless, huge advances have been made in QPM genetics and breeding over the last 40 years, and this remains an active and productive area of research. Many maize breeding programs are actively working to develop QPM varieties and to disseminate them widely with the intention of having a positive impact on the nutrition, health, and even survival of undernourished infants and children. However, while significant progress has been made on the breeding, genetics, and biochemical basis of QPM, comparable progress has not been made in evaluating the nutritional impact of QPM for target individuals and populations. It is hoped that the positive results presented here will encourage greater attention and investment in evaluating the nutritional impact of QPM in sound scientific studies.

LIST OF REFERENCES

LIST OF REFERENCES

Akalu, G. (2005) Nutrition research in Ethiopia including QPM village study. Presentation at 2nd Maize HarvestPlus Meeting, Sete Lagoas, MG, Brasil, August 10-12, 2005.

Akuamoa-Boateng, A. (2002) Quality protein maize: Infant feeding trials in Ghana. Ghana Health Service, Ashanti, Ghana.

Allen, L. H. (2003) Interventions for micronutrient deficiency control in developing countries: Past, Present and Future. Journal of Nutrition 133: 3875S-3878S.

Beaton, G. H., Calloway, D. H. and Murphy, S. P. (1992) Estimated protein intakes of toddlers: Predicted prevalence of inadequate intakes in village populations in Egypt, Kenya and Mexico. American Journal of Clinical Nutrition 55: 902-911.

Bjarnason, M. and Vasal, S. K. (1992) Breeding of quality protein maize (QPM). In: Plant Breeding Reviews, Vol. 9 (Janick, J., ed.), pp. 181-216. John Wiley & Sons, Inc., New York, NY.

Bressani, R., Arroyave, G. and Scrimshaw, N. S. (1953) The nutritive value of Central American corns. I. Nitrogen, ether extract, crude fiber and minerals of twenty-four varieties in Guatemala. Food Research 18: 261-267.

Bressani, R., Scrimshaw, N. S., Behar, M. and Viteri, F. (1958) Supplementation of cereal proteins with amino acids. II. Effect of amino acid supplementation of corn-masa at intermediate leves of protein intake on the nitrogen retention of young children. Journal of Nutrition 66: 501-513.

Bressani, R., Elías, L. G., Santos, M., Navarrete, D. and Scrimshaw, N. S. (1960) El contenido de nitrógeno y de aminoácidos esenciales de diversas selecciones de maíz. Archivos Venezolanos de Nutrición 10: 85-100.

Bressani, R., Wilson, D. L., Behar, M., Chung, M. and Scrimshaw, N. S. (1963) Supplementation of cereal proteins with amino acids. V. Effect of supplementing lime-treated corn with different levels of lysine, tryptophan and isoleucine on the nitrogen retention of young children. Journal of Nutrition 80.

Bressani, R., Alvarado, J. and Viteri, F. (1969) Evaluación, en niños, de la calidad de la proteína del maíz opaco-2. Archivos Latinoamericanos de Nutrición 19: 129-140.

Bressani, R. (1991) Protein quality of high-lysine maize for humans. Cereal Foods World 36: 806-811.

Carriquiry, A. L. (1999) Assessing the prevalence of nutrient inadequacy. Public Health Nutrition 2: 23-33.

Casella, G. and Berger, R. L. (2002) Statistical inference, 2nd ed. Duxbury/Thomson Learning, Pacific Grove, CA.

CDC (2006) NCHS - 2000 CDC Growth Charts: United States. http://www.cdc.gov/growthcharts/ (accessed Dec 12, 2006).

Clark, H. E., Glover, D. V., Betz, J. L. and Batley, L. B. (1977) Nitrogen retention of young men who consumed isonitrogenous diets containing normal, opaque-2 or sugary-2-opaque-2 corn. Journal of Nutrition 107: 404-411.

Cochran, W. G. (1954) The combination of estimates from different experiments. Biometrics 10: 101–129.

Cooper, H. (1998) Integrating research: A guide for literature reviews, 3rd ed. Sage, Newbury Park, CA.

Cooper, H., and Hedges, L. V. (1994) The handbook of research synthesis. Russell Sage Foundation, New York, NY.

Dankyi, A. A., Sallah, P. Y. K., Adu-Appiah, A., and Gyamera-Antwi (2005) Determinants of the adoption of quality protein maize, Obatanpa, in southern Ghana - logistic regression analysis. Fifth West and Central Africa Biennial Regional Maize Workshop, IITA-Cotonou, Republic of Benin, May 2-6, 2005.

De Groote, H., Gunaratna, N., Krivanek, A. F. and Friesen, D. (2006) Recent advances in the development, uptake and impact assessment of quality protein maize (QPM). Symposium on “New Technology Development to Reduce Hunger in Sub-Saharan Africa” at the 26th Conference of the International Association of Agricultural Economists (IAAE), Gold Coast, Queensland, Australia, August 12-18, 2006.

DerSimonian, R. and Laird, N. (1986) Meta-analysis in clinical trials. Controlled Clinical Trials 7: 177–188.

Efron, B. and Tibshirani, R. J. (1993) An introduction to the bootstrap. Chapman & Hall, New York, NY.

FAO (1991) Protein quality evaluation. FAO Food and Nutrition Paper 51. Food and Agriculture Organization (FAO) of the United Nations, Rome.

FAO (1992) Maize in human nutrition. FAO Food and Nutrition Series No. 25. Food and Agriculture Organization (FAO) of the United Nations, Rome.

FAO (2005) Kenya Food Balance Sheet (Year 1997). http://faostat.fao.org/ (accessed September 18, 2005).

Flores, M., Flores, Z., and Lara, M. Y. (1966) Food intake of Guatemalan Indian children, ages 1 to 5. Journal of the American Dietetic Association 48: 480-487.

Food and Nutrition Board (FNB), Institute of Medicine (IOM). (2002) Dietary Reference Intakes for Energy, Carbohydrate, Fiber, Fat, Fatty Acids, Cholesterol, Protein and Amino Acids (Macronutrients). National Academy Press, Washington, D.C.

Gibson, R. S. (2005) Principles of nutritional assessment, 2nd ed. Oxford University Press, Oxford.

Glass, G. V., McGaw, B., and Smith, M. L. (1981) Meta-analysis in social research. Sage, Newbury Park, CA.

Graham, G. G., Glover, D. V., Lopez de Romaña, G., Morales, E. and MacLean Jr., W. C. (1980) Nutritional value of normal, opaque-2, and sugary-2 opaque-2 maize hybrids for infants and children. I. Digestibility and utilization. Journal of Nutrition 110: 1061-1069.

Graham, G. G., Lembcke, J., Lancho, E. and Morales, E. (1989) Quality protein maize: Digestibility and utilization by recovering malnourished infants. Pediatrics 83: 416-421.

Hedges, L. V. and Olkin, I. (1985) Statistical methods for meta-analysis. Academic Press, Orlando, FL.

Hunter, J. E., and Schmidt, F. L. (2004) Methods of meta-analysis: Correcting error and bias in research findings, 2nd ed. Sage, Newbury Park, CA.

Jackson, A. A. and Calder, P. C. (2004) Severe undernutrition and immunity. In: Handbook of Nutrition and Immunity (Gershwin, M. E., Nestel, P. & Keen, C. L., eds.), pp. 71-92. Humana Press, Totowa, NJ.

Kies, C. and Fox, H. M. (1972) Protein nutritional value of opaque-2 corn grain for human adults. Journal of Nutrition 102: 757-765.

King, J. C. (2002) Evaluating the impact of plant biofortification on human nutrition. Journal of Nutrition 132: 511S-513S.

Krivanek, A. F., De Groote, H., Gunaratna, N. S., Diallo, A. O. and Friesen, D. (2007) Breeding and disseminating quality protein maize (QPM) for Africa. African Journal of Biotechnology 6: 312-324.

Kurpad, A. V., Regan, M. M., Raj, T., Vasudevan, J., Kuriyan, R., Gnanou, J. and Young, V. R. (2003) Lysine requirements of chronically undernourished adult Indian men, measured by a 24-h indicator amino acid oxidation and balance technique. American Journal of Clinical Nutrition 77: 101-108.

Last, J. M. (1988) A dictionary of epidemiology, 2nd ed. Oxford University Press, Oxford.

Lauderdale, J. (2000) Issues regarding targeting and adoption of quality protein maize (QPM). CIMMYT Economics Working Paper 00-02. CIMMYT, Mexico D.F.

Luna-Jaspe G., H., Mora Parra, J. O., Rozo Bernal, C. and Pérez de Serrano, S. (1971) Comparación de la retención de nitrógeno en niños alimentados con maíz de gene opaco-2 y leche de vaca. I. Resultados con baja ingesta de proteína. Archivos Latinoamericanos de Nutrición 21: 437-447.

Mertz, E. T., Bates, L. S. and Nelson, O. E. (1964) Mutant gene that changes protein composition and increases lysine content of maize endosperm. Science 145: 279-280.

Millward, D. J. and Jackson, A. A. (2003) Protein/energy ratios of current diets in developed and developing countries compared with a safe protein/energy ratio: Implications for recommended protein and amino acid intakes. Public Health Nutrition 7: 387-405.

Ministry of Finance and Planning (2000a) Second Report on Poverty in Kenya, Volume I: Incidence and Depth of Poverty. Central Bureau of Statistics (CBS), Ministry of Finance and Planning, Government of Kenya, Nairobi, Kenya.

Ministry of Finance and Planning (2000b) Second Report on Poverty in Kenya, Volume II: Poverty and Social Indicators. Ministry of Finance and Planning, Government of Kenya, Nairobi, Kenya.

Morales Guerra, M. (2002) Efecto del consumo de maíz de alta calidad proteínica en niño(a)s de familias indígenas de las regiones Mazateca y Mixe del Estado de Oaxaca: Una estrategia agronómica de desarrollo entre campesinos que practican agricultura de subsistencia. Colegio de Postgraduados, Montecillo, Texcoco, Edo. de Mexico, Mexico.

Ortega Alemán, E. C., Coulson Romero, A. J. and Ordóñez Argueta, L. I. (2006) Efectos de la ingesta de maíz de alta calidad de proteínas versus maíz normal en el crecimiento y desarrollo físico de niños de 1 a 5 años de edad, Centro de Desarrollo Infantil Mildred Abaunza, Septiembre-Diciembre 2005. Universidad Nacional Autónoma de Nicaragua, Managua, Nicaragua.

Paes, M. C. D. and Bicudo, M. H. (1994) Nutritional Perspectives of Quality Protein Maize. Proceedings of the International Symposium on Quality Protein Maize, EMBRAPA/CNPMS, Sete Lagoas, MG, Brasil.

Pellett, P.L. (1996) World essential amino acid supply with special attention to South-East Asia. Food and Nutrition Bulletin 17: 204-234.

Pixley, K. V. and Bjarnason, M. S. (1993) Combining ability for yield and protein quality among modified-endosperm opaque-2 tropical maize inbreds. Crop Science 33: 1229-1234.

Pradilla, A., Linares, F., Francis, C. A. and Fajardo, L. (1973) El maíz de alta lisina en nutrición humana. Simposio Sobre Desarrollo y Utilización de Maíces de Alto Valor Nutritivo, Colegio de Postgraduados, ENA, Chapingo, México.

Rahmanifar, A. and Hamaker, B. R. (1999) Potential nutritional contribution of quality protein maize: A close-up on children in poor communities. Ecology of Food and Nutrition 38: 165-182.

Rosenthal, R. (1991) Meta-analytic procedures for social research, revised ed. Sage, Newbury Park, CA.

Scrimshaw, N. S., Bressani, R., Behar, M. and Viteri, F. (1958) Supplementation of cereal proteins with amino acids. I. Effect of amino acid supplementation of corn-masa at the higer levels of protein intake on the nitrogen retention of young children. Journal of Nutrition 66: 458-499.

Secretaría de Salud. (1994) Norma Oficial Mexicana para el Control de la Nutrición, Crecimiento y Desarrollo del Niño y del Adolescente. Diario Oficial de la Federación, México.

Selmi, C., Invernizzi, P., Zuin, M., Ansari, A. A. and Gershwin, M. E. (2004) Evaluation of the immune function in the nutritionally at-risk patient. In: Handbook of Nutrition and Immunity (Gershwin, M. E., Nestel, P. & Keen, C. L., eds.), pp. 1-18. Humana Press, Totowa, NJ.

Singh, J. (1977) Studies on assessing the nutritive value of opaque-2 maize: Technical report of the project. Indian Agricultural Research Institute, New Delhi, India.

Singh, J., Koshy, S., Agrawal, K. N., Lodha, M. L., Singh, N. N. and Sethi, A. S. (1980) Relative efficacy of opaque-2 maize in the growth of preschool children. Indian Journal of Nutrition and Dietetics 17: 326-334.

USDA (2005) USDA National Nutrient Database for Standard Reference. United States Department of Agriculture (USDA) – Agricultural Research Service, http://www.ars.usda.gov/main/site_main.htm?modecode=12354500 (accessed September 18, 2005).

Valverde, V., Delgado, H., Belizan, J. M., Martorell, R., Mejía-Pivaral, V., Bressani, R., Elías, L. G., Molina, M. and Klein, R. E. (1983) The Patalul project: Production, storage, acceptance and nutritional impact of opaque-2 corns in Guatemala. Instituto de Nutrición de Centro América y Panamá (INCAP), Guatemala.

Vasal, S. K. (2000) The quality protein maize story. Food and Nutrition Bulletin 21: 445-450.

Victora, C. G., Habicht, J.-P. and Bryce, J. (2004) Evidence-based public health: Moving beyond randomized trials. American Journal of Public Health 94: 400-405.

World Food Prize Foundation (2007) 2000 World Food Prize Laureates Dr. Surinder Vasal and Dr. Evangelina Villegas. http://www.worldfoodprize.org/laureates/Past/2000.htm (accessed March 10, 2007).

World Health Organization (2006) Guidelines on food fortification with micronutrients. World Health Organization and Food and Agriculture Organization of the United Nations, Geneva.

Young, V. R., and Pellett, P. L. (1990) Current concepts concerning indispensable amino acid needs in adults and their implications for international nutrition planning. Food and Nutrition Bulletin 12: 289-300.

VITA

VITA

Nilupa S. Gunaratna was born in Colombo, Sri Lanka. She received her B.A. in the College Scholar Program at Cornell University in 1999 with an emphasis on plant genetics. Nilupa began her graduate study in Statistics at Purdue University in 2001 as a National Science Foundation (NSF) Vertical Integration of Research and Education (VIGRE) Fellow. In 2002, she received a M.S. in Agronomy from Purdue University with a specialization in Plant Genetics and Breeding. In 2003, she received a M.S. in Statistics from Purdue University with a specialization in Applied Statistics. She conducted her dissertation research in Statistics under the direction of Professor George McCabe. Her research interests include the linkages between agriculture, nutrition, and health, and statistical issues and methods to evaluate new technologies in these areas.

Nilupa also served as a statistical consultant for the Statistical Consulting Service and the Technical Assistance Program at Purdue University. She was active in the development of the Statistics in the Community (STATCOM) program at Purdue and served in various leadership roles including Director in 2004-2005. In 2006-2007, she led the American Statistical Association's (ASA's) "Statistics in the Community (STATCOM)" Strategic Initiative, an effort to promote and support the development of pro bono statistical consulting programs at colleges, universities, and ASA organizations and to establish a national STATCOM Network linking such programs.