Commentary

Food Fortification With Folic Acid for Prevention of Spina Bifida and Anencephaly: The Need for a Paradigm Shift in Evidence Evaluation for Policy-Making

Homero Martinez*, Helena Pachón, Vijaya Kancherla, and Godfrey P. Oakley, Jr.

*Correspondence to Dr. Homero Martinez, Global Technical Services Unit, Nutrition International, 180 Elgin Street, Suite 1000, Ottawa, ON K2P 2K3, Canada (e-mail: hmartinez@nutritionintl.org).

Initially submitted November 18, 2020; accepted for publication March 9, 2021.

Context-specific evidence evaluation is advocated in modern epidemiology to support public health policy decisions, avoiding excessive reliance on experimental study designs. Here we present the rationale for a paradigm shift in evaluation of the evidence derived from independent studies, as well as systematic reviews and meta-analyses of observational studies, applying Hill’s criteria (including coherence, plausibility, temporality, consistency, magnitude of effect, and dose-response) to evaluate food fortification as an effective public health intervention against folic acid–preventable (FAP) spina bifida and anencephaly (SBA). A critical appraisal of evidence published between 1983 and 2020 supports the conclusion that food fortification with folic acid prevents FAP SBA. Policy-makers should be confident that with mandatory legislation, effective implementation, and periodic evaluation, food fortification assures that women of reproductive age will safely receive daily folic acid to significantly reduce the risk of FAP SBA. Current evidence should suffice to generate the political will to implement programs that will save thousands of lives each year in over 100 countries.

anencephaly; folic acid; food fortification; low- and middle-income countries; neural tube defects; spina bifida

Abbreviations: FAP, folic acid–preventable; NTD, neural tube defect; RCT, randomized controlled trial; SBA, spina bifida and anencephaly.

Editor’s note: The opinions expressed in this article are those of the authors and do not necessarily reflect the views of the American Journal of Epidemiology.

Spina bifida and anencephaly are major neural tube defects (NTDs) that affect pregnancies worldwide. NTDs are a leading cause of stillbirths and under-5 (under age 5 years) child mortality in high-, middle-, and low-income countries (1). Based on scientific evidence from randomized controlled trials (RCTs) (2–4), nonrandomized intervention trials (5, 6), and observational studies (7, 8), the US Public Health Service recommended in 1992 that all women capable of becoming pregnant consume 400 μg of folic acid (vitamin B9) per day to lower the risk of NTDs, specifically folic acid–preventable (FAP) spina bifida and anencephaly (SBA) (9). While this primary-prevention recommendation may be met through timely consumption of daily supplements, a public health approach based on large-scale fortification of staple foods with folic acid has been determined to be more effective (10).

Policies of mandatory food fortification with folic acid started in the United States, Canada, Costa Rica, and Oman in the late 1990s (11–15), followed by South Africa, Chile, and Australia in the 2000s (16–18). Currently, about 80 countries have mandatory folic acid food fortification programs (19). It is estimated that effective mandatory folic acid food fortification programs prevented 18% of all potential FAP SBA cases worldwide in 2017 (20) and 22% of cases worldwide in 2019 (21). Surveillance studies in countries without mandatory policies for folic acid fortification report an average NTD prevalence range reflecting 10–20 cases per 100,000 births in high-income countries and 40–130 cases per 100,000 births in low- and middle-income countries, while the lowest prevalence is reported in countries with effective fortification—consistently about 5 cases per
100,000 births (22, 23). From a public health perspective, this excessively high prevalence of FAP SBA in countries without folic acid fortification of food staples, especially low- and middle-income countries, is equivalent to a highly preventable epidemic.

Despite multiple studies with findings that support the effectiveness of folic acid fortification in improving folate status among women of reproductive age (24–26) and reducing the prevalence of NTDs (27), some policy-makers in countries without folic acid fortification programs are skeptical of implementing fortification. The skepticism arises from a concern that the evidence does not come from RCTs, the type of study considered to provide the highest-quality evidence for evaluation of causation and efficacy (28). However, since previous RCTs have unequivocally established that folic acid prevents NTDs (2–4), it would be unethical to conduct new ones. A new RCT would require folic acid to be withheld from one group of reproductive-age women involved in the study and, as a result, would put those women at risk of an FAP SBA pregnancy.

In this commentary, we suggest a paradigm shift away from traditional evidence evaluation schema, which are more applicable to clinical decisions but less relevant in guiding public health decisions, to support the protective effect of large-scale folic acid food fortification in the prevention of FAP SBA.

**THE NEED FOR A PARADIGM SHIFT**

Traditionally, evidence-based medicine has long relied on RCTs as the preferred study design for informing clinical decisions. RCTs control for potential bias and confounding when assessing impact (efficacy) better than studies with nonexperimental designs (29). Given that different study designs support statistical interpretations ranging from association to causal inferences, clinicians have proposed an evidence-based pyramid that places study designs along a continuum: Weaker evidence is placed lower on the hierarchy (the base of the pyramid), and stronger evidence is placed higher (coming closer to the top) (30). Systematic reviews and meta-analyses are assumed to provide the highest level of evidence (31).

The Grading of Recommendations, Assessment, Development, and Evaluation (GRADE) Working Group developed an approach with which to guide the rating of the quality of evidence from systematic reviews to inform guidelines for both clinical and public health questions (32). An important aspect of the Working Group’s approach is the recognition that study design alone may not account for the risk of bias. Instead, they recommend focusing on the quality of data collection, the magnitude of the effect, the dose-response gradient, and control of potential biases in order to assess the strength of the contribution made by a study. The World Health Organization and the Cochrane Collaboration have both adopted this approach to support their recommendations (32). In addition, depending on the kind of decision to be made (i.e., medical treatment, clinical prognosis, or decisions about public health effectiveness), the pyramid of evidence may vary, giving greater weight to specific study designs (30). Therefore, a retrospective cohort study or case-control study initially evaluated as providing a “low quality” rating may be upgraded if, for example, the magnitude of the treatment effect is very large and if there is evidence of a dose-response relationship or adequate control for plausible biases (33, 34). Murad et al. (35) have further proposed that it is important to evaluate the quality of the evidence contributing to systematic reviews and meta-analyses before placing these designs at the top of the evidence pyramid.

For several years, different authors in the fields of epidemiology and public health have proposed that, when evaluating the evidence available for guiding public health decision-making, there is a need to move beyond the evidence-based pyramid (36–40). Sir Austin Bradford Hill, one of the pioneers of modern epidemiology, noted that the evaluation of data for making causal inferences should be rigorous. Hill proposed a number of assessment criteria that have been used extensively in establishing causality from observational studies (41). In a recent evaluation of Hill’s criteria, Lucas and McMichael supported the value of evidence from observational studies; they stated that most epidemiologic research is nonexperimental and conducted in an “inherently ‘noisy’ environment in free-living populations” (42, p. 792). They underlined the challenges to controlling for potential confounding variables in settings which usually have less opportunity to do so than RCTs (42). Stoltzfus (38) has highlighted several limitations of RCT designs in providing information for guiding public health nutrition interventions, supporting the need to consider mechanistic theories when evaluating the strength of the evidence used to inform public health guidance. Similarly, Vandenbroucke et al. pushed for using a pluralistic approach to evaluate causal evidence, underscoring that “the important causal questions are asked not within studies, but between them” (39, p. 1785). Commenting on a recent series of papers on the future of epidemiology, Diez Roux (36) concluded that observational studies now stand to provide the evidence needed in population health due to lack of feasibility and other inherent limitations of RCTs. In implementing evidence-based policies, an important but often overlooked consideration is the likelihood of harm from not assuming causality and making no change in current policy. As Frieden points out, “waiting for more data is often an implicit decision not to act” (37, p. 472).

**SIR BRADFORD HILL’S CRITERIA FOR INFERRING CAUSALITY**

In evaluating the evidence related to the protective effect of folic acid fortification on FAP SBA, it is important to address 6 of Hill’s causal criteria, including coherence, plausibility, temporality, consistency, magnitude of effect, and dose-response. Drawing both from single-site observational studies and from systematic reviews and meta-analyses, the effectiveness of folic acid fortification in reducing the risk of FAP SBA is clear.

The reduction in FAP SBA in response to an improvement in red blood cell folate levels postfortification is both coherent and plausible with our current knowledge of the
biological mechanisms related to folate metabolism and its role in DNA synthesis and methylation processes that help regulate cell synthesis and growth, which are critical during the period when the neural tube is closing (43). Observational studies carried out in different countries (Cameroon, Chile, Fiji, Tanzania) have documented improvements in blood folate levels and reductions in NTDs (17, 44–46).

Evidence from pre- and postfortification studies (Canada, Chile, Costa Rica, Iran, South Africa, United States) also shows a temporal relationship between improvements in population folate status following implementation of fortification programs and subsequent reductions in the prevalence of FAP SBA (14, 47–52).

The magnitude of the protective effect of folic acid fortification is more pronounced in regions with high prefortification prevalence of NTDs. For example, in Canada, where the prevalence of NTDs prefortification was known to be higher in eastern provinces and lower in western provinces, reductions in NTDs postfortification were higher in the former (38.0 cases/10,000 births) and lower in the latter (2.1 cases/10,000 births) (53). Similar results have been found elsewhere (49, 54). On average, a meta-analysis by Keats et al. (26) documented a protective effect against NTDs (pooled odds ratio = 0.59, 95% confidence interval: 0.49, 0.70). Theoretical models predict that the lowest prevalence of NTDs observed in populations where maternal red blood cell folate concentration is above the optimal cutoff of 906 nmol/L recommended by the World Health Organization (55) to achieve the maximum protection against NTDs is approximately 5–6 cases per 10,000 births (56), as has actually been found in different countries (14, 49, 54).

Studies carried out in various countries have shown a consistent (i.e., repeatedly observed in different studies, in different settings, and at different times) protective effect against FAP SBA through folic acid fortification, including studies conducted in Argentina, Australia, Brazil, Canada, Chile, Costa Rica, Iran, Jordan, Oman, Peru, Saudi Arabia, South Africa, and the United States (24, 57). The systematic review and meta-analysis by Keats et al. (26), which included 16 national fortification programs, found that maize and/or wheat-flour fortification with adequate levels of folic acid over a period of 1–5 years resulted in significant increases in serum folate levels among women of reproductive age, as well as a significant reduction in the prevalence of folate deficiency (relative risk = 0.20, 95% confidence interval: 0.15, 0.25).

Research has documented a dose-response relationship between 1) folate/folic acid consumption and serum and red blood cell folate levels (17, 18, 45, 48, 58–65); 2) maternal red blood cell folate concentrations and percent reduction in the prevalence of FAP SBA (10); and 3) voluntary (partial) fortification implementation and mandatory (full) implementation and reduction in FAP SBA prevalence, demonstrated both in single-site studies (52–54, 66) and in a systematic review and meta-analysis of 179 studies (27). Atta et al. (27) found that the pooled total prevalence of spina bifida in countries with mandatory folic acid fortification was 1.5 times lower than that in countries with voluntary fortification or no fortification.

Evidence presented in this commentary on folic acid’s protective role in prevention of FAP SBA is informed by experimental, quasi-experimental, and observational study designs in populations worldwide. Thus, the overall target validity (a joint measure of both internal and external validity) of the effectiveness of fortification is high and should encourage policy-makers to implement this intervention (67).

CONCLUSION

In summary, as a team of nutrition researchers and epidemiologists with multiple years of experience in food fortification and FAP SBA, we advocate a shift towards context-specific evidence evaluation and avoidance of excessive reliance on hierarchical models of causal evaluation based predominantly on RCTs, which can be less than optimal when guiding public health interventions for preventing FAP SBA. In this commentary, we have aimed to convey to public health advocates and nutrition policymakers, especially in low- and middle-income countries, that there are robust and consistent data from both individual and pooled studies that support folic acid fortification as an effective public health intervention for reducing the occurrence and recurrence of FAP SBA (20). Policy-makers should be confident that with mandatory legislation and effective implementation allowing periodic evaluation, food fortification assures that women of reproductive age will receive the daily folic acid needed for healthy pregnancy and significantly reduces the risk of FAP SBA. Skepticism and inaction hinder political will, cost thousands of lives each year globally, cause unnecessary suffering to families, and place an avoidable burden on health-care systems. A change in paradigm can counter skepticism and clarify that large-scale food fortification—a proven and highly cost-effective intervention already recognized as one of the greatest public health achievements of the past century—is the most effective measure in accelerating the global prevention of FAP SBA (68–70).

ACKNOWLEDGMENTS

Author affiliations: Global Technical Services Unit, Nutrition International, Ottawa, Ontario, Canada (Homero Martinez); Hospital Infantil de México Federico Gómez, Mexico City, Mexico (Homero Martinez); Hubert Department of Global Health, Rollins School of Public Health, Emory University, Atlanta, Georgia, United States (Helena Pachón); Food Fortification Initiative, Atlanta, Georgia, United States (Helena Pachón); and Center for Spina Bifida Prevention, Department of Epidemiology, Rollins School of Public Health, Emory University, Atlanta, Georgia, United States (Vijaya Kancherla, Godfrey P. Oakley, Jr.).

We thank Jessie Genoway for editing the final draft of this article. She received no compensation.

Conflict of interest: none declared.
REFERENCES


38. Stoltzfus RJ. How can the scientific community support the generation of the evidence needed to improve the quality of


