Special Article

EV McCollum Award Lecture, 2003

Long-latency deficiency disease: insights from calcium and vitamin D[1–4]

Robert P Heaney

ABSTRACT

Nutrient intake recommendations and national nutritional policies have focused primarily on prevention of short-latency deficiency diseases. Most nutrient intake recommendations today are based on prevention of the index disease only. However, inadequate intakes of many nutrients are now recognized as contributing to several of the major chronic diseases that affect the populations of the industrialized nations. Often taking many years to manifest themselves, these disease outcomes should be thought of as long-latency deficiency diseases. Sometimes they come about by the same pathophysiologic mechanism that produces the index disease, but sometimes the mechanisms are completely different. Well-documented examples of both short- and long-latency deficiency states involving calcium and vitamin D are described briefly. Then, the insights derived from these nutrients are tentatively applied to folic acid. Discerning the full role of nutrition in long-latency, multifactorial disorders is probably the principal challenge facing nutritional science today. The first component of this challenge is to recognize that inadequate intakes of specific nutrients may produce more than one disease, may produce diseases by more than one mechanism, and may require several years for the consequent morbidity to be sufficiently evident to be clinically recognizable as "disease." Because the intakes required to prevent many of the long-latency disorders are higher than those required to prevent the respective index diseases, recommendations based solely on preventing the index diseases are no longer biologically defensible. *Am J Clin Nutr* 2003;78:912–9.

KEY WORDS Long-latency deficiency disease, short-latency deficiency disease, index disease, calcium, vitamin D, folic acid, nutrient intakes, scientific and policy challenges

INTRODUCTION

At its birth roughly a century ago, nutritional science had to overcome the prevailing view that all disease was caused by external invaders, either bacterial or toxic. The idea that not eating something could make one sick was inconceivable. We owe much to pioneers like EV McCollum, who were convinced, and ultimately convinced the scientific and medical communities, that foods contained substances that the body needed for health and that not getting enough of them caused explicit disease. Unfortunately, the medical community's approach to nutrition is still strongly influenced by the early external invader model. Some of the most prominent contemporary medical nutritional efforts, such as those involving cholesterol, saturated fat, and salt, typify that paradigm. Additionally, scarcely a day passes without finding a report in the general news media of a study linking this or that cancer (or other dire outcome) with consumption of some nutrient. Clearly, the toxicity model continues to capture the attention of the medical community.

The early triumphs of nutritional science related largely to short-latency deficiency diseases such as beriberi, pellagra, rickets, scurvy, etc. It could hardly have been otherwise inasmuch as the connection between cause (withdrawal of a key nutrient) and effect (disease) would not have been perceptible had the disease latency period been protracted.

Although some of the short-latency disorders such as rickets recur sporadically today, such diseases are largely a feature of medicine's past. National nutritional policies, sporadic food fortification, and nationwide food distribution systems have combined to reduce the incidence of nutritional diseases with short latency periods, at least in the industrialized nations, to such an extent that most physicians simply do not experience them anymore. Only in the context of pregnancy and well-baby care is there much medical emphasis on the positive value of eating right, mainly because the time horizon is short enough to make the benefit perceptible to the practitioner. By contrast, the long-latency disorders that afflict the human race today, such as cancer, cardiovascular disease, and central nervous system degeneration, constitute a field that, from the standpoint of clinical nutrition, is left largely to nutritional quacks and charlatans. Discerning the extent to which nutrition may play a role in such disorders, positive or negative, is probably the principal challenge facing nutritional science today.

Sufficient work has been done in the past several years with 2 bone-related nutrients, calcium and vitamin D, to provide some possibly useful insights that may be transferable to other nutrients and deficiencies. In the present article, I review them briefly and suggest some parallels in the burgeoning field of

[1] From Creighton University, Omaha.

[2] Presented at the American Society for Clinical Nutrition 43rd Annual Meeting, April 12, 2003, San Diego.

[3] Presentation of the EV McCollum Award supported by Wyeth Nutrition.

[4] Address reprint requests to RP Heaney, Creighton University, 601 North 30th Street, Suite 4841, Omaha, NE 68131. E-mail: rheaney@creighton.edu.

Received May 28, 2003.

Accepted for publication June 23, 2003.

folic acid research. In this analysis, I distinguish first between short- and long-latency deficiency and second between deficiencies brought about through the same mechanism that produces the index disease and those that operate entirely through other mechanisms. Then, I conclude with a quick glance at the scientific and policy challenges presented by long-latency deficiency diseases.

THE APPROACH OF NUTRITIONAL SCIENCE TO DEFICIENCY DISEASE

A key feature of the usual approach of nutritional science to deficiency disease needs to be delineated. In addition to an understandable bias toward deficiency diseases with short latency periods, there has been a tendency to link individual nutrients to single diseases and, by extension, to single disease mechanisms. Thus, thiamine was linked to beriberi, niacin to pellagra, vitamin D to rickets, iodine to goiter, etc. Although it was soon recognized that most of the nutrients concerned played key roles in the metabolism of many tissues, the index disease for each nutrient was implicitly or explicitly considered to reflect the most vulnerable process or most sensitive tissue. This may explain why nutrient intake recommendations today are still largely pegged to the prevention of the index diseases for specific nutrients. The presumption has been that if the intake of the nutrient is sufficient to prevent the expression of the index disease, then the intake is, ipso facto, adequate for the functioning of the total organism. But this approach overlooks 2 additional possibilities. First, there may be long-term consequences of lesser degrees of deficiency that nevertheless operate through a mechanism similar to that inducing the index disease. Second, entirely different metabolic mechanisms may be involved. If the disease latency period were prolonged, then these mechanisms would probably be missed; or if they were recognized, they would probably be attributed to nonnutritional causes. Calcium and vitamin D each illustrate both possibilities.

CALCIUM

The index disease for calcium is osteoporosis. It was slow to be accepted as such, perhaps largely because it is a disease with an inherently long latency period. Nevertheless, it is clear that bone serves as the body's nutrient reserve for calcium. In fact, calcium is unique among the nutrients because its reserve has acquired an important role in its own right, quite distinct from the function of calcium ions in inter- and intracellular signaling. We walk on our calcium reserve.

It is self-evident that if calcium intake is inadequate during growth, the full genetic program for skeletal mass cannot be realized; and after growth, if calcium intake is inadequate to offset obligatory losses, acquired skeletal mass cannot be maintained. The only uncertainty has centered on where the boundary between adequate and inadequate might lie. This boundary has been well defined in experimental animals for years (eg, *see* references 1 and 2), and the relative slowness of the acceptance of a causal role for calcium in human osteoporosis is due in part to the multifactorial complexity of that disorder and to the fact that the skeletal mass accumulated by populations with different diets seems to differ far less than does the calcium content of their diets. Adaptive mechanisms are evidently at work, helping populations adjust to widely varying intakes. As recently as 1989, the *British Medical Journal* could carry a two-part review that concluded that low calcium intake played no role in the development of osteoporosis (3, 4).

Nevertheless, it was clear from animal data that the principle of inadequate intake leading to reduced bone mass was operative. For example, diets containing $<$115 mg Ca/100 kcal resulted in reduced bony accumulation during growth in rodents (1), and in its publication on primate nutrition, the National Academy of Sciences noted explicitly that diets containing $\leq$50 mg Ca/100 kcal resulted in osteoporosis and recommended an intake $\approx$3 times that value for maintenance. [Anomalously, the Food and Nutrition Board, which is ultimately a creature of the National Academy of Sciences, placed the adult human calcium recommended dietary allowance at only 40 mg Ca/100 kcal as recently as 1989 (5). Although the 1997 recommendations (6) are slightly higher, we still recommend substantially lower calcium intakes for ourselves than we actually feed the chimpanzees in our vivaria.]

A further factor confusing the matter in humans has been the fact that the only outcome of importance with respect to osteoporosis is fragility fractures, for which good epidemiologic data have not been widely available. Variation in fractures of the proximal femur, which are perhaps the most serious of the osteoporotic fractures and the ones most likely to be identifiable in epidemiologic studies, is caused more by differences in hip geometry and patterns of falling than by bone massiveness, and hence, at a population level, this fracture has been a poor surrogate for osteoporosis when seeking nutritional linkages. In addition, retrospective reconstruction of nutrient intakes over several decades is notoriously inaccurate. Thus, both the dependent variable (fragility fractures) and the independent variable (calcium intake) have been hard to quantify in observational studies testing the hypothesis relating them.

Despite these early obstacles, there is now a consensus that calcium intake is important for the bone health of humans, just as it is for that of our laboratory and household companion animals. The foregoing considerations were cited mainly because they illustrate some of the pitfalls along the road to finding a connection between a nutrient and a long-latency disorder.

In addition to the more or less commonsense involvement of calcium intake with bone health, low calcium intakes have been implicated over the last several years in a surprising variety of nonskeletal disorders. In my own attempts to make sense of these disparate calcium connections, I suggested a classification scheme that identifies just 3 separate mechanisms by which low calcium intake produces disease (7). The first mechanism, which involves the structural role played by the size of the nutrient reserve, is intuitively straightforward and is the basis for current intake recommendations. The second mechanism is a consequence of the fact that intestinal calcium absorption is highly inefficient, and thus unabsorbed calcium is naturally present in the intestinal lumen. Over the course of human evolution, this unabsorbed calcium has come to serve useful functions precisely because it is not absorbed. Finally, when the adaptive mechanisms that the body uses to compensate for low calcium intake are deployed continuously, they may have consequences for the function of tissues that are unrelated to the calcium economy proper. (It is worth noting in passing that all 3 disease mechanisms are themselves unrelated to the fundamental role of calcium as second messenger in intracellular signaling. That role is so essential to cells that they

maintain their own intracellular calcium reserves. As a result, this function is essentially never compromised by nutritional inadequacy. All of the now recognized manifestations of calcium deficiency are expressed at a higher level, ie, the integrative functioning of complex organisms.)

Each of these mechanisms has been explored at greater length elsewhere (7). In the present article, I touch on them only in enough detail to illustrate how quite distinct mechanisms can contribute to very different disease syndromes, each of which ought to be recognized as manifestations of nutritional deficiency.

Disease latency for osteoporosis is long because, even with a zero calcium intake, bone cannot be torn down fast enough to produce clinically appreciable depletions in less than several years. Apart from intrinsic bone-wasting disease or severe paralysis of one sort or another, excretion of calcium over net absorbed intake in an amount >100 mg/d would be rare. Such a rate of loss would deplete an adult skeleton by ≈3%/y. In testing isolated specimens in a biomaterials laboratory, even a 3% depletion would have a measurable effect on bone strength, but clinically the depletion would not be apparent until it had gone on for ≥10 y. An even slower rate of loss, more typical of developing osteoporosis, may require ≈30 y to produce depletion severe enough to result in manifest fragility.

Perhaps the clearest demonstrations of the first role of calcium (ie, the skeleton as calcium nutrient reserve) are the studies of Chapuy et al (8) and Dawson-Hughes et al (9), in which calcium and vitamin D supplements given to older subjects not only protected against age-related bone loss, but reduced nonvertebral fracture risk 35–55%. [Although the bone mass effect of the supplementation probably explained some of the fracture reduction, a reduction in the bone resorptive process itself, which was brought about by a decrease in the homeostatic activity involved in depletion of the reserves, probably also contributed to the reduced fragility (10).] This bone-sparing effect tends to be more evident in older persons than in younger persons because of an apparent decrease with age in the ability to adapt to low calcium intakes. (In a sense, therefore, the calcium intake requirement for skeletal maintenance in the elderly uncovers what is probably the optimal intake at all ages, because this requirement precludes the need for compensatory mechanisms in younger persons, who, although more able to bring such mechanisms into play, may nevertheless pay a nonskeletal price for the protection they afford—*see* below.)

The second basic mechanism whereby calcium intake may influence health is the functionality of calcium within the gut lumen. Unabsorbed calcium binds with substances such as food oxalate and unabsorbed fatty and bile acids. In the first instance, it reduces the absorption of oxalate and hence decreases its renal excretory load, thereby reducing the risk of calcium oxalate stones. (Although it is true that urinary calcium increases with dietary calcium, it is also true that urinary oxalate is a much more powerful risk factor for stones than is urinary calcium, and thus any reduction in the renal oxalate burden has a disproportionately larger effect on the risk than does the modest increase in urinary calcium.) In the second instance, the complexing of unabsorbed fatty acids and bile acids by calcium reduces their cancer-promoting activity in the colon mucosa. Intestinal oxalate binding produces an immediate benefit, and in that sense calcium oxalate kidney stones can be considered a short-latency calcium deficiency disease. By contrast, the anti–cancer-promoting effect is a long-latency phenomenon that is expressed in years and perhaps even decades.

Both of these luminal effects illustrate another important feature of certain deficiency diseases that may not have been as fully evident with, for example, beri-beri or pellagra. As far as is now recognized, the propensity to stone formation is not itself a nutritional problem. Instead, it involves a defect in the excretion of solution stabilizers by the kidney. Low calcium intake simply uncovers the propensity to stone formation by allowing a solute excretory load that is greater than impaired urine can keep in solution or suspension. In other words, oxalate stone–formers represent a population segment that is unusually sensitive to low calcium intake. Similarly, the prevention of colon cancer by high calcium intakes occurs only in persons with a predisposition to mucosal neoplasia.

Once considered, these first 2 mechanisms, the size of the reserve and the functionality of unabsorbed calcium in the intestinal lumen, are conceptually straightforward. The third mechanism is only beginning to be unraveled, and in contrast with the other 2, is decidedly indirect. The high parathyroid hormone (PTH) secretion that is the normal accompaniment of low calcium intake has been known for some time to be associated with increased intracellular calcium ion concentrations (11). Research indicates that calcitriol, the synthesis of which is stimulated by PTH, is mainly responsible. Calcitriol binds to membrane receptors in various tissues and effectively opens calcium channels (12). This nongenomic effect occurs at physiologic concentrations of calcitriol and can be blocked by calcium channel blockers. The increase in intracellular calcium ion concentration has been recognized for some time in hypertension (13) and has been shown more recently in adipocytes, both human and murine (12). The increased intracellular calcium ion concentration, in turn, results in increased muscle tone in smooth muscle cells and switches adipocytes from a predominantly lipolytic mode to a predominantly liposynthetic mode (12). The net effect is that, as calcium intake decreases, intracellular calcium concentrations tend to increase, thus producing a group of disorders that Fujita and Palmieri have characterized as "calcium paradox disease" (14). Disease expression with this mechanism requires ≥3 conditions: low calcium intake, impaired ability to pump Ca^{2+} out of the cells concerned, and an environmental trigger (eg, abundant energy supplies for obesity and, perhaps, high salt intakes for hypertension).

Both obesity and hypertension, as well as the other disorders that have been associated with low calcium intake, are distinctly multifactorial, and each represents a more or less stereotyped morbid outcome with many different mechanisms for reaching it. Low calcium intake is only one of those. For example, in the population samples that we analyzed in our laboratory at Creighton University, we estimated that ≈3% of the variation in adult body weight can be attributed to calcium intake (15, 16). Despite the smallness of the effect in our data and in other reports, there has been substantial consistency across the various studies published to date (17). The situation for hypertension is similar—a small portion of the population variability is attributable to calcium intake, but effects are fairly consistent across many studies.

However, as Rose pointed out several years ago, small shifts in the mean of a population usually produce large changes in

the proportion of persons in the tails of the distribution who may be above or below a specified level (18). We showed in our own data, for example, that although calcium intakes in the range of currently recommended intakes reduce variability in weight gain only slightly, they shift the mean of the distribution sufficiently to reduce to zero the average midlife weight gain in women and thereby greatly reduce the transition into the category of obesity (16). A similar effect was found for young adults in the Coronary Artery Risk Development in Young Adults (CARDIA) study (19), in which high dairy intakes reduced the transition from overweight at entry to obesity by 30% and the transition from normal blood pressure to hypertension by 62%.

In any event, my purpose in the present article is not to explore in depth the role of calcium in the various disorders that may fall into this heading of calcium paradox disease, but to point out that any morbidity of the various sorts catalogued above, which would not occur if calcium intake were higher, reflects nutritional deficiency. Such morbidities should thus be classified and managed as deficiency diseases.

VITAMIN D

The index disease for vitamin D is rickets (or osteomalacia). The disease is generally considered to come about because of malabsorption of calcium and phosphorus, which is correct as far as it goes. As the efficiency of calcium absorption decreases, PTH secretion increases. The latter stimulates synthesis of calcitriol, which tends to improve calcium absorption efficiency but evidently not enough to make up the difference, thus leading to still further increases in PTH concentration. At the same time, increased PTH concentrations lower the renal threshold for phosphorus, which, together with reduced phosphorus absorption, produces hypophosphatemia. It is this hypophosphatemia that impairs osteoblast and chondroblast cell function and leads to disordered metaphyseal growth plates and the characteristic histologic features of rickets and osteomalacia.

One should have thought that the impaired calcium absorption produced by vitamin D deficiency would produce osteoporosis, as has been shown directly for low calcium intakes. Until recently though, impaired calcium absorption has not been formally factored into the canonical picture of vitamin D deficiency disease. However, in his heuristically important reconceptualization of vitamin D deficiency, Parfitt (20) defined 3 degrees of what he termed "hypovitaminosis D osteopathy." The first is characterized simply by calcium malabsorption without histologic abnormalities, which predictably results in osteoporosis. The second is also characterized by calcium malabsorption (and osteoporosis), with early histologic features of osteomalacia but without associated symptoms or clinical laboratory findings indicative of osteomalacia. Only with the third, or most severe, degree of vitamin D depletion are there both clinical and histologic rickets or osteomalacia. The principal clinical manifestation of stages 1 and 2 is, therefore, osteoporosis, not rickets or osteomalacia; and if the vitamin D input never becomes low enough to produce clinical rickets or osteomalacia, then its deficiency is expressed solely as osteoporosis. Thus, stages 1 and 2 of hypovitaminosis D osteopathy constitute a long-latency vitamin D deficiency disease that operates through the same basic mechanism as that of the index disease.

The modest vitamin D fortification of milk and a few other foods over the past several decades, coupled with the use of vitamin D supplements in children, has eliminated most cases of stage-3 vitamin D deficiency in North America. However, these same stratagems have not been sufficient to prevent the lesser degrees of deficiency (*see* below). Indeed, the stratagems were not designed to do so. The vitamin D requirements are pegged to the prevention of stage-3 deficiency, and there still remains a presumption that if one does not have rickets or osteomalacia, then one has sufficient vitamin D. This position is no longer tenable, not just for the theoretical reasons just outlined, but because a rapidly growing body of evidence indicates malfunctioning and morbidity, which are correctable with vitamin D, in persons who do not have the index disease. For example, increasing serum 25-hydroxyvitamin D [25(OH)D] concentrations from $\approx$50 nmol/L to $\approx$80 nmol/L (both values within the usual reference range) improves calcium absorption efficiency by nearly two-thirds (21) and reduces osteoporotic fracture risk by one-third (22). Thus, it is now incontrovertible that vitamin D deficiency that is less extreme than that required to produce rickets or osteomalacia nevertheless produces disease, although of a long-latency character.

Both the short-latency deficiency disease (rickets or osteomalacia) and the long-latency deficiency disease (osteoporosis) operate through the same fundamental physiologic mechanism, although at different levels of repletion. However, as with calcium, it is now becoming evident that vitamin D operates through other mechanisms as well.

It has been known for $>$60 y that there is an inverse association between sun exposure and cancer mortality (23). Early on, it was also recognized that there was both a direct association between squamous cell carcinoma and sun exposure and an inverse association between squamous cell and internal cancers. In fact, noting these 2 associations, early theorists suggested that development of squamous cell carcinoma conferred immunity against various visceral cancers. The negative correlation between sun exposure and several common cancers is very robust and has been documented repeatedly (eg, *see* references 24–26). It is now understood to be a function not of immunity per se, but of serum 25(OH)D concentrations, and, of course, does not require development of skin cancer at all.

More recently, it was shown that many tissues have calcitriol receptors, an initially puzzling observation because the tissues involved bore no obvious relation to the calcium economy. (Such puzzlement reflected the presumption of one disease and one mechanism for each nutrient.) Many of these same tissues have also been shown to express 1-α-hydroxylase and thus are able to make calcitriol for themselves. Both in disease states such as some of the myelodysplasias and psoriasis and in various cancer cell lines in culture, calcitriol has been shown to induce cell differentiation and to control cell proliferation (27).

Thus, vitamin D can be said to have at least 2 distinct groups of functions. One is classically endocrine: calcitriol synthesis increases in response to PTH; the resulting product circulates in the blood to reach target tissues (gut and bone), where it exerts its proper effect and, through negative feedback (augmenting calcium inputs into extracellular fluid), regulates PTH secretion and hence its own production. The other is autocrine (or perhaps paracrine). The cells or tissues concerned make and

degrade the hormone locally to regulate their own proliferation and differentiation.

Because the endocrine function is driven by PTH, it is relatively independent of serum 25(OH)D concentrations, which explains why calcitriol concentrations may well be entirely normal in patients with florid osteomalacia. It is likely, however, that the autocrine function is dependent on the substrate concentration of 25(OH)D. For this reason, persons with a low 25(OH)D concentration (whether because they have little solar exposure or because they have extensive pigmentation that limits solar vitamin D synthesis in the skin) are less able to make calcitriol within the tissues concerned in an amount sufficient to exert the controls over cell proliferation that are needed to reduce oncogenesis. Thus, although circulating blood calcitriol concentrations may be normal in vitamin D deficiency states, tissue concentrations, which are dependent on substrate 25(OH)D concentrations, may be low. Several investigators have shown that prostate cancer incidence and mortality are inversely associated both with solar UV radiation and with serum 25(OH)D concentrations (24, 25). Also, as is widely recognized, African American males have both the lowest serum 25(OH)D concentrations and the highest prostate cancer incidence of any population group in the United States. Such an association could well be regarded as merely coincidental if considered in isolation. However, it is both predicted and explained by the current understanding of the vitamin D autocrine system.

Collateral evidence of this potentially important autocrine role of vitamin D is seen in the recent use of calcitriol and calcitriol analogs as components of combined chemotherapy programs for various advanced cancers (28, 29). Such use should not be dismissed as entirely "pharmacologic," because the apparent protection against cancer provided by 25(OH)D occurs in the range of serum values found naturally in various human populations. Grant (24), for example, has estimated that perhaps as much as 20% of the breast cancer burden of Europe is a manifestation of vitamin D deficiency. To the extent that this conclusion is correct, such cancers represent a long-latency deficiency disorder involving an entirely different mechanism from that of the classical vitamin D deficiency index disease.

Only relatively recently has reliable measurement of serum 25(OH)D concentrations been available, and most of the physiology of vitamin D had been worked out before that time and thus was unconnected to specific levels of vitamin D repletion. For example, although the Food and Nutrition Board had no difficulty in identifying 25(OH)D as the functional indicator for vitamin D status, they were not able, with the data that existed at that time, to assign numerical values to the lower limit of the normal range or to assign cutoff values for various vitamin D activities (6). As a result, current measurements are usually related to laboratory "reference" ranges (which is inevitably circular inasmuch as such ranges record what is observed in people who are considered "normal" only because they do not have recognizable rickets or osteomalacia). The studies to which I just alluded (21, 22), together with a large body of data relating PTH concentrations to circulating 25(OH)D concentrations (eg, see references 30 and 31), indicate that the lower end of an acceptable normal range must be $\approx$80 nmol/L. (By contrast, the lower end of most reference ranges is closer to 40 nmol/L.) In recent years, nutritional scientists have referred to values $<$20 nmol/L as "deficient" (because they were reproducibly associated with osteomalacia or rickets), values above $\approx$80 as "normal," and values in between as "insufficient," without a clear consensus as to where the boundary might lie between "insufficient" and "normal." From the standpoint of the calcium economy, it now seems clear from the studies just cited that, at least for people living in the United Kingdom and North America, values $<$80 nmol/L are deficient. The awkward term *insufficiency* ought to be dropped. Its use simply reflected the point made earlier that the index disease for vitamin D was osteomalacia or rickets. If one had it, one was "deficient"; and if one did not, one could not be "deficient."

Where the boundary between optimal and suboptimal serum 25(OH)D concentrations may fall with respect to the emerging autocrine function of vitamin D is still quite unclear. Its elucidation in the years ahead may be one of the most important tasks confronting nutritional science, if for no other reason than because the payoff could be so large.

FOLIC ACID

For both calcium and vitamin D, we have come to recognize that low intakes produce not only the index disease but also long-latency diseases previously unrecognized, some of which arise through the same mechanisms responsible for the index disease and some of which arise through mechanisms totally unrelated and, until recently, unimagined. Are these 2 nutrients the only instances of such polyvalent functional activity? That seems quite unlikely, and there are suggestive parallels with folic acid that seem worth exploring briefly. In this analysis, I use folic acid as a surrogate for perhaps many other nutrients.

The index disease for folic acid deficiency is megaloblastic anemia, which is due ultimately to failure of folic acid–mediated one-carbon transfers in purine and pyrimidine synthesis, including specifically the conversion of uracil to thymine. The result is inappropriate insertion of uridine into DNA in place of thymidine, leading somehow to chromosomal fragmentation. Given the rapid cellular turnover involved in hemopoiesis and intestinal mucosal replacement, the associated megaloblastic changes represent short-latency deficiency diseases, which, like rickets, are manifestations of severe deficiency. Also of short latency and probably occurring through the same mechanism (but at folic acid intakes high enough to prevent anemia) are neural tube defects (NTDs) that develop in the early stages of fetal development.

The biochemical details of this index mechanism have been worked out for some time, as has the additional role of folic acid (together with vitamins B-6 and B-12) in the regeneration of methionine from homocysteine. Until recently, this latter reaction, which is of obvious biochemical importance, was less clearly associated with functional consequences, let alone disease. One of the insights from recent advances in molecular biology is the realization that high nuclear concentrations of methyl donors are important for the methylation of DNA and that methionine is an important component of that system. DNA methylation, in turn, is a critical regulator of gene expression and a controller of transposon replication, a process that contributes to oncogenesis. It has been recognized for some time that there is an inverse association between folic acid intake and the development of certain cancers (eg, *see* reference 32), but it was not until DNA methylation was understood that a plausible mechanism was presented. Oncogenesis is a long, slow process, and if it were augmented by

TABLE 1
Suggested categorization of deficiency diseases associated with 3 illustrative nutrients[1]

	Index			Nonindex			
Nutrient	Disease	Mechanism	Latency	Disease	Mechanism	Latency	Intake[2]
Calcium	Osteoporosis	Depletion of nutrient reserve	Long	Oxalate urolithiasis	Reduced intestinal binding of oxalic acid	Short	Equal to index
				Colon cancer promotion	Reduced intestinal binding of fatty and bile acids	Long	Equal to index
				"Calcium paradox" diseases[3]	Elevation of intracellular [Ca^{2+}]	Long	Uncertain
Vitamin D	Rickets and osteomalacia	Intestinal malabsorption of calcium and phosphorous	Short	Osteoporosis	Same as for index disease	Long	Above index
				Oncogenesis	Reduced control of cell differentiation by 1-α-hydroxylation of 25(OH)D	Long	Above index
Folic acid	Megaloblastic anemia	Impaired DNA synthesis	Short	Neural tube defects	Same as for index disease	Short	Above index
				Oncogenesis	Reduced DNA methylation by methionine	Long	Uncertain
				Connective tissue degenerative disorders[4]	Irreversible degradation of elastic proteins by homocysteine	Long	Uncertain

[1] 25(OH)D, 25-hydroxyvitamin D.
[2] Relative to the intake required to prevent the index disease.
[3] Including some portion of the disease burden of hypertension, preeclampsia, obesity, insulin resistance, polycystic ovary syndrome, and premenstrual syndrome.
[4] Including some portion of the disease burden of occlusive vascular disease, osteoporosis, dementia, and presbyopia.

folic acid deficiency, the result would be a long-latency deficiency disorder involving a mechanism different from that of the index disease.

But there is yet another part to this story. In the absence of adequate folic acid, homocysteine accumulates and its concentration in circulating blood increases. Homocysteine is one of the most reactive biological molecules, and its concentration in blood is dependent not only on folic acid but also on vitamins B-6 and B-12. High folic acid intakes lower plasma homocysteine, on average, to <10 μmol/L, and it is likely that some such concentration should be the target of folic acid supplementation. Although the B Vitamin Panel of the Food and Nutrition Board took erythrocyte folic acid as the primary functional indicator for folic acid status, they nevertheless explicitly incorporated plasma homocysteine concentration into their estimation of the average requirement (33).

Plasma homocysteine has repeatedly been identified as a strong independent risk factor for cardiovascular disease, as well as several other disorders. It has been suggested that homocysteine damages slowly-turning-over connective tissues by irreversible binding to sulfhydryl groups in the epidermal growth factor–like repeats of fibrillin and similar structural proteins. Krumdieck and Prince (34), for example, have called attention to the close parallels between the hallmark manifestations of homocystinuria (with serum homocysteine concentrations typically >100 μm/L), ie, occlusive vascular disease, osteoporosis, mental deterioration, and ectopia lentis, and the counterpart manifestations of "normal" aging (with serum homocysteine concentrations between 10 and 100 μm/L), ie, occlusive vascular disease, osteoporosis, dementia, and presbyopia. To the extent that this formulation is correct, we would have manifestations of yet other long-latency deficiency diseases, which would also operate through mechanisms distinct from the mechanism of the index disease.

COMMENT

Nutritional scientists have become increasingly aware of a probable role for nutrients in reducing the disease burden of several chronic diseases. However, these scientists have generally been careful to distinguish such activity from the more proper "nutrient" effects, as manifested, for example, in the prevention of the index diseases of the nutrients concerned. I believe that this distinction is unhelpful and have attempted in this article to indicate that inadequate intakes of specific nutrients may produce more than one disease, may produce them by more than one mechanism, and may require several years for the consequent morbidity to be sufficiently evident to be clinically recognizable as "disease." Once it has been clearly established that such disorders are produced by an inadequate intake of a given nutrient and are preventable by a higher intake, they should be precisely termed "deficiency diseases," just as are the index diseases more classically associated with the nutrients concerned.

For 3 illustrative nutrients, how the various related deficiency diseases fit into this scheme of short and long latencies and index and nonindex mechanisms, as well as the relative inputs required to prevent these diseases, is shown in **Table 1**. Note particularly that the intakes required for prevention of the nonindex diseases, although not precisely known for many of the disorders, are otherwise as high or higher than required to prevent the index disease. For example, preventing the osteo-

porosis of vitamin D deficiency requires ≈4 times the vitamin D input needed to prevent rickets, and the NTDs of folic acid deficiency require 2–4 times the intake needed to prevent anemia. Indeed, that relation follows almost automatically from this scheme, because nonindex diseases that are preventable at intakes lower than the intake needed to prevent the index disease could not exist in persons who were free of the index disease.

Table 1 is offered as an expression of a conceptual framework in which nutrient-related health status outcomes can be organized. It is intended to be illustrative, not exhaustive. For example, the problem of falls in the elderly, which several studies have shown can be appreciably reduced by elevating serum 25(OH)D concentrations into the range that is optimal for bone status, is not listed but is arguably as important as the other nonindex diseases related to vitamin D. Yet another nonindex mechanism is evidently at work here. Finally, it is necessary to stress once again that, for most of the diseases cited in Table 1 (index as well as nonindex), causation is multifactorial, and the scheme in Table 1 is applicable solely to that portion of the disease burden that is produced primarily by inadequacy of the nutrient concerned.

The challenges that nutritional science faces with respect to long-latency and nonindex deficiencies are 2-fold: a scientific challenge and a policy challenge. Although both are difficult, the latter is probably the more formidable.

The scientific challenge is to find how to establish the role played by low nutrient intake in various morbid outcomes when the clinical manifestation may take years to develop and when other, nonnutrient factors may either contribute in their own right to the total morbidity burden or be required for disease expression. The usual randomized controlled trial design, which is the only study type that permits strong causal inference, would probably not be available for many of the nutrients and disease endpoints concerned (for both ethical and practical reasons). At the same time, epidemiologic and ecologic studies, which are useful in identifying associations, can be notoriously misleading, as recent experience with estrogen and coronary artery disease has made painfully apparent (35). I know of no obvious solution to this challenge. (Indeed, it is a challenge precisely because it is difficult.) Nor do I believe that there is likely to be a single solution applicable to every nutrient.

However, I suggest that one may find some help in the experience obtained in pinning down the carcinogenicity of tobacco products. This too was a situation involving a long-latency disease outcome, in which randomized controlled trials would have been both impractical and unethical. (With nutrition, one looks for benefit rather than harm, but the investigational challenges in inferring causality are the same either way.) The usual principles for supporting causal inference from observational data should hold in the nutritional context. These include, in no particular order, biological plausibility, correct temporal sequence, dose-response relations, experiments of nature found in inborn errors of metabolism, and demonstration of the causal connection in animal models. These principles are all well understood in a general way, and what I suspect may have been lacking up till now was the conviction within the field of nutrition that long-latency deficiency diseases exist, that they are nutritional problems, and that the use of such inferential and investigative stratagems may be both appropriate and necessary.

The policy issues may be more vexing. Both folic acid and calcium present interesting examples. Recognizing the role of folic acid in NTDs, the Food and Nutrition Board first recommended the fortification of food through the Food and Drug Administration in 1974 (36), which was 24 y before mandatory fortification was ultimately implemented. Then, when fortification was mandated in 1996 (for implementation by 1998), the level of fortification was less than that recommended by many scientific and professional groups. For example, the American Academy of Pediatrics, the professional group most directly involved with NTDs, had urged a level of fortification 2.5 times the level that was finally adopted (37). Estimates of NTD reduction in infants in the post-1998 period have been correspondingly modest [19% fewer NTDs, as contrasted with an anticipated 40–50% reduction with a proposed, higher level of fortification (38).]

A similar story can be told for calcium. In 1988 the Surgeon General suggested a program of calcium fortification of food as one means of combating the widespread problem of calcium deficiency in the US population (39). The Food and Drug Administration, which is a part of the same branch of government, did not act on that suggestion. Then in 1992, a citizen petition was filed with the agency to produce the same effect; the petition called for quite modest levels of calcium fortification of grain products (40). The Food and Drug Administration failed to follow its own rules and tabled the petition without response for several years until, on further intervention by the petitioner, it was finally rejected on the grounds that the agency did not have the staff to do the necessary background work.

It is difficult either to understand or to justify the resistance of regulatory authorities to changes in current practices that is typified by these examples (but not confined to folic acid and calcium by any means). This is particularly true in the case of folic acid, for which the result has been a tragic crop of infants with unnecessary NTDs, which are preventable with current knowledge but are still occurring even today because of inadequate protection.

Regulatory bodies obviously cannot respond to every shift in the winds of public opinion, and some caution is to be commended. Nevertheless, some middle ground ought to be found. The manifest difficulty in doing so is one of the reasons why the policy side of this challenge may be more difficult than the scientific side.

Yet another aspect of the latter problem is the position of nutritional policymakers that "we won't change without proof." The irony of that seemingly unassailable position was captured by Walter Willett in an interview he gave to Gary Taubes (41): "They say, 'You really need a high level of proof to change the recommendations,' which is ironic, because they never had a high level of proof to set them."

The most difficult part of this challenge, I suspect, is finding the will to settle on nutrient intake recommendations that are biologically defensible while we wait for evidence that lower intakes may be safe or higher intakes more beneficial. In many instances, because the current recommendations are based on the prevention of the index disease only, they can no longer be said to be biologically defensible. The preagricultural human diet, insofar as it can be reconstructed, may well be a better starting point for policy. Such a diet cannot be known in detail,

but as several investigators (42–44) have shown, the diet probably would have had at least the following features: high protein intake, low glycemic index, high calcium intake, high folic acid intake, an alkaline ash residue, and (for reasons of latitude and skin exposure) high vitamin D input. It is in this nutritional context that human physiology evolved, and it is to this context that human physiology is adapted. The burden of proof should fall on those who say that these more natural conditions are not needed and that lower intakes are safe.

REFERENCES

1. Forbes RM, Weingartner KE, Parker HM, Bell RR, Erdman JW Jr. Bioavailability to rats of zinc, magnesium and calcium in casein-, egg-, and soy protein–containing diets. J Nutr 1979;109:1652–60.
2. National Research Council. Nutrient requirements of nonhuman primates. Washington, DC: National Academy of Sciences, 1978:56–8.
3. Kanis JA, Passmore R. Calcium supplementation of the diet-I. BMJ 1989;298:137–40.
4. Kanis JA, Passmore R. Calcium supplementation of the diet-II. BMJ 1989;298:205–8.
5. National Research Council. Recommended dietary allowances, 10th ed. Washington, DC: National Academy of Sciences, National Academy Press, 1989.
6. Food and Nutrition Board, Institute of Medicine. Dietary reference intakes for calcium, magnesium, phosphorus, vitamin D, and fluoride. Washington, DC: National Academy Press, 1997.
7. Heaney RP. Ethnicity, bone status, and the calcium requirement. Nutr Res 2002;22:153–78.
8. Chapuy MC, Arlot ME, Duboeuf F, et al. Vitamin D_3 and calcium to prevent hip fractures in elderly women. N Engl J Med 1992;327:1637–42.
9. Dawson-Hughes B, Harris SS, Krall EA, Dallal GE. Effect of calcium and vitamin D supplementation on bone density in men and women 65 years of age or older. N Engl J Med 1997;337:670–6.
10. Heaney RP. Is the paradigm shifting? Bone (in press).
11. Palmieri GMA, Nutting DF, Bhattacharya SK, Bertorini TE, Williams JC. Parathyroid ablation in dystrophic hamsters. J Clin Invest 1981; 68:646–54.
12. Zemel MB, Shi H, Greer B, DiRienzo D, Zemel PC. Regulation of adiposity by dietary calcium. FASEB J 2000;14:1132–8.
13. Brickman AS, Nyby MD, von Hungen K, Eggena P, Tuck ML. Calcitropic hormones, platelet calcium, and blood pressure in essential hypertension. Hypertension 1990;16:515–22.
14. Fujita T, Palmieri GMA. Calcium paradox disease: calcium deficiency prompting secondary hyperparathyroidism and cellular calcium overload. J Bone Miner Metab 2000;18:109–25.
15. Davies KM, Heaney RP, Recker RR, et al. Calcium intake and body weight. J Clin Endocrinol Metab 2000;85:4635–8.
16. Heaney RP. Normalizing calcium intake: projected population effects on body weight. J Nutr 2003;133:268S–70S.
17. Parikh SJ, Yanovski JA. Calcium intake and adiposity. Am J Clin Nutr 2003;77:281–7.
18. Rose G, Day S. The population mean predicts the number of deviant individuals. BMJ 1990;301:1031–4.
19. Pereira MA, Jacobs DR Jr, Van Horn L, Slattery ML, Kartashov AI, Ludwig DS. Dairy consumption, obesity, and the insulin resistance syndrome in young adults. JAMA 2002;287:2081–9.
20. Parfitt AM. Osteomalacia and related disorders. In: Avioli LV, Krane SM, eds. Metabolic bone disease and clinically related disorders. 2nd ed. Philadelphia: WB Saunders, 1990:329–96.
21. Heaney RP, Dowell MS, Hale CA, Bendich A. Calcium absorption varies within the reference range for serum 25-hydroxyvitamin D. J Am Coll Nutr 2003;22:142–6.
22. Trivedi DP, Doll R, Khaw KT. Effect of four monthly oral vitamin D_3 (cholecalciferol) supplementation on fractures and mortality in men and women living in the community: randomised double blind controlled trial. BMJ 2003;326:469–74.
23. Apperly FL. The relation of solar radiation to cancer mortality in North America. Cancer Res 1941;1:191–5.
24. Grant WB. An ecologic study of dietary and solar ultraviolet-B links to breast carcinoma mortality rates. Cancer 2002;94:272–81.
25. Ahonen MH, Tenkanen L, Teppo L, Hakama M, Tuohimaa P. Prostate cancer risk and prediagnostic serum 25-hydroxyvitamin D levels (Finland). Cancer Causes Control 2000;11:847–52.
26. Grant WB. An estimate of premature cancer mortality in the U.S. due to inadequate doses of solar ultraviolet-B radiation. Cancer 2002;94: 1867–75.
27. Holick MF. Vitamin D: importance for prevention of cancer, type I diabetes, heart disease, and osteoporosis. Am J Clin Nutr (in press).
28. Beer TM, Eilers KM, Garzotto M, Egorin MJ, Lowe BA, Henner WD. Weekly high-dose calcitriol and docetaxel in metastatic androgen-independent prostate cancer. J Clin Oncol 2003;21:123–8.
29. Solit DB, Smaletz O, DeLaCruz A, et al. Phase I study of high dose calcitriol ($1,25(OH)_2$ vitamin D_3) and zoledronic acid in patients with prostate cancer. 2002. Internet: http://www.asco.org/ac/1,1003,_12-002489-00_18-002002-00_19-00777-00_28-00RESULTPAGE,00.asp (accessed 20 Jan 2003).
30. Chapuy MC, Preziosi P, Maamer M, et al. Prevalence of vitamin D insufficiency in an adult normal population. Osteoporos Int 1997;7: 439–43.
31. Thomas MK, Lloyd-Jones DM, Thadhani RI, et al. Hypovitaminosis D in medical inpatients. N Engl J Med 1998;338:777–83.
32. Giovannucci E, Stampfer MJ, Colditz GA, et al. Multivitamin use, folate, and colon cancer in women in the Nurses' Health Study. Ann Intern Med 1998;129:517–24.
33. Food and Nutrition Board, Institute of Medicine. Dietary reference intakes for thiamin, riboflavin, niacin, vitamin B_6, folate, vitamin B_{12}, pantothenic acid, biotin, and choline. Washington, DC: National Academy Press, 1998.
34. Krumdieck CL, Prince CW. Mechanisms of homocysteine toxicity on connective tissues: implications for the morbidity of aging. J Nutr 2000;130:365S–8S.
35. Rossouw JE, Anderson GL, Prentice RL, et al. Risks and benefits of estrogen plus progestin in healthy postmenopausal women: principal results from the Women's Health Initiative randomized controlled trial. JAMA 2002;288:321–33.
36. Food and Nutrition Board. Proposed fortification for cereal-grain products. Washington, DC: National Academy of Sciences, 1974.
37. AAP. Position statement on increasing folic acid levels to prevent neural tube birth defects. 1998. Internet: http://www.aap.org/advocacy/wshing/folic498/html (accessed 20 Jan 2003).
38. Honein MA, Paulozzi LJ, Mathews TJ, Erickson JD, Wong LYC. Impact of folic acid fortification of the US food supply on the occurrence of neural tube defects. JAMA 2001;285:2981–6.
39. US Department of Health and Human Services, Public Health Service. The Surgeon General's report on nutrition and health. Washington, DC: US Government Printing Office, 1988. [US DHHS publication (PHS) 88-50211.]
40. Department of Health and Human Services, Public Health Service, Food and Drug Administration. Letter to Paul A Lachance and Harold L Newmark, dated June 19, 2000. Docket no. 92P-0064. Washington, DC: US Department of Health and Human Services, 2000.
41. Taubes G. The soft science of dietary fat. Science 2001;291:2536–45.
42. Sebastian A, Frassetto LA, Sellmeyer DE, Merriam RL, Morris RC Jr. Estimation of the net acid load of the diet of ancestral preagricultural *Homo sapiens* and their hominid ancestors. Am J Clin Nutr 2002;76: 1308–16.
43. Cordain L, Brand Miller J, Eaton SB, Mann N, Holt SHA, Speth JD. Plant-animal subsistence ratios and macronutrient energy estimations in worldwide hunter-gatherer diets. Am J Clin Nutr 2000;71:682–92.
44. Eaton B, Nelson DA. Calcium in evolutionary perspective. Am J Clin Nutr 1991;54(suppl):281S–7S.