Biomarkers of cobalamin (vitamin B-12) status in the epidemiologic setting: a critical overview of context, applications, and performance characteristics of cobalamin, methylmalonic acid, and holotranscobalamin II

Ralph Carmel

ABSTRACT
Cobalamin deficiency is relatively common, but the great majority of cases in epidemiologic surveys have subclinical cobalamin deficiency (SCCD), not classical clinical deficiency. Because SCCD has no known clinical expression, its diagnosis depends solely on biochemical biomarkers, whose optimal application becomes crucial yet remains unsettled. This review critically examines the current diagnostic concepts, tools, and interpretations. Their exploration begins with understanding that SCCD differs from clinical deficiency not just in degree of deficiency but in fundamental pathophysiology, causes, likelihood and rate of progression, and known health risks (whose causation by SCCD awaits proof by randomized clinical trials). Conclusions from SCCD data, therefore, often may not apply to clinical deficiency and vice versa. Although many investigators view cobalamin testing as unreliable, cobalamin, like all diagnostic biomarkers, performs satisfactorily in clinical deficiency but less well in SCCD. The lack of a diagnostic gold standard limits the ability to weigh the performance characteristics of metabolic biomarkers such as methylmalonic acid (MMA) and holotranscobalamin II, whose specificities remain incompletely defined outside their relations to each other. Variable cutoff selections affect diagnostic conclusions heavily and need to be much better rationalized. The maximization of reliability and specificity of diagnosis is far more important today than the identification of ever-earlier stages of SCCD. The limitations of all current biomarkers make the combination of ≥2 test result abnormalities, such as cobalamin and MMA, the most reliable approach to diagnosing deficiency in the research setting; reliance on one test alone courts frequent misdiagnosis. Much work remains to be done.

INTRODUCTION
The time appears ripe, 2 decades into the era of population studies of metabolic cobalamin status, such as NHANES, to critically review the concepts and the performances of the diagnostic biomarkers. Such a review must consider the diverse perspectives of epidemiologic information users, whether investigators who mine and reanalyze the data or clinicians who try to adapt relevant findings to medical care.

The current state of uncertainty about cobalamin deficiency and about its reliable diagnosis especially requires critical thinking. This uncertainty is most obvious in the clinical arena, but its ramifications extend to and must be understood in the public health arena. Even though the public health focus is almost entirely subsumed by subclinical cobalamin deficiency (SCCD) and not clinical deficiency, understanding the complicated subject in either arena begins with the clinical perspective.

Clinicians have become uncertain about what to do with the many asymptomatic patients with mild biochemical abnormalities who increasingly displace ill patients as the face of cobalamin deficiency in medical practice, as important distinctions between clinical disease and biochemical changes blur (1). This incompletely understood SCCD raises quandaries: should such patients be treated, monitored, or perhaps ignored? And, if treated, for how long, with what doses, and to what endpoints? The use of oral cobalamin without a precise diagnosis dominates the landscape and confuses patients and physicians. Moreover, diagnostic abilities have atrophied, most obviously in the disastrous loss of ability to test cobalamin absorption status, which had been crucial to informed prognosis and management. Deficiency that arises from intrinsic factor (IF)–related malabsorption usually progresses inexorably (and was once lethal), whereas nonmalabsorptive or minimally malabsorptive deficiency progresses very slowly or not at all (2).

In the public health arena, the emphasis on frequencies and finding the earliest possible cases made insufficient room for the important distinctions between the cobalamin deficiencies, their causes, and their outcomes. The proper balance between the pursuit of ultimate sensitivity of biochemical detection and overdagnosis

1 From the Departments of Medicine, New York Methodist Hospital, Brooklyn, NY, and Weill Medical College of Cornell University, New York, NY.
2 Presented at the conference “NHANES Monitoring of Biomarkers of Folate and Vitamin B-12 Status: a Roundtable Review,” held in Rockville, MD, 15–16 July 2010.
3 The author’s work that this article reviews was supported by a grant, DK32640, from the NIH. The National Center for Health Statistics of the Centers for Disease Control and Prevention and the Office of Dietary Supplements of the NIH cosponsored and supported the roundtable.
4 Address correspondence to R Carmel, Department of Medicine, New York Methodist Hospital, 506 Sixth Street, Brooklyn, NY 11215. E-mail: rac9001@nyp.org.

doi: 10.3945/ajcn.111.013441.
awaits rigorous determination of the effect that the diagnosis of SCCD actually has on health and the benefits of intervention.

THE EVOLUTION OF DIAGNOSTIC TESTING OF COBALAMIN DEFICIENCY

The early history of testing

From 1949 to the late 1980s cobalamin testing consisted almost exclusively of serum cobalamin measurement. The chief goal was to diagnose or rule out the uncommon but very serious deficiency state caused in 94% of cases by pernicious anemia (the irreversible malabsorptive disease defined by loss of gastric IF secretion) and intestinal diseases that prevent IF uptake (3). The clinical hallmark of cobalamin deficiency was megaloblastic anemia (>70% of cases) and/or, somewhat less often, neurologic dysfunction (4).

Readers accustomed to current views of cobalamin as an insensitive test may be surprised to learn that sensitivities were in fact 95–97% in medical encounters (probably because clinical deficiency appears only after its underlying severe malabsorption had sufficiently depleted cobalamin stores). Specificity was always viewed as the greater problem because unexplained low cobalamin concentrations seemed frequent. Not only did folate deficiency regularly depress serum cobalamin, but pregnancy does so even more often (4). Many other low cobalamin concentrations without clinical signs or symptoms had no obvious explanations.

The first useful functional metabolic tests appeared in the 1960s. Urine methylmalonic acid (MMA) became measurable (5) but was relatively insensitive. A highly sensitive tool identified cobalamin and folate deficiencies by their impairment of deoxyuridine suppression (6). Demand for these cumbersome tests was limited.

Metabolic markers and subclinical deficiency

Two developments in the mid-1980s changed the scope of cobalamin deficiency and its determination. One was the first documentation of minimal, often “subclinical,” deficiency in a series of studies from my laboratory in patients who lacked anemia, macrocytosis, or deficiency symptoms (7–12). We used the deoxyuridine suppression test, which has proven to be as sensitive as MMA and homocysteine testing (13). Equally important, the deficiencies only rarely arose from disruption of IF-mediated absorption, such as pernicious anemia (8, 10, 14). Our discovery of SCCD indicated that many unexplained low cobalamin concentrations described in the earlier clinical literature were neither spurious nor artificial.

The second, nearly concurrent, development was the introduction of assays that reliably detect minute elevations of MMA and homocysteine, the known substrates of cobalamin-mediated reactions (15–18). These assays facilitated large surveys that proved that SCCD was much more common than classical clinical deficiency. The criteria and the performance characteristics of the diagnostic markers that define SCCD and clinical deficiency are contrasted in Table 1.

SCCD as a public health issue

Because of SCCD’s high frequency, cobalamin deficiency emerged as a potential public health issue. Research focused on determining the prevalence of biochemical deficiency in populations, especially the elderly. Cross-sectional studies also explored whether health risks accompanied the subclinical state. Serious suspicions have been raised in relation to neurologic and cognitive risks, but the mixed record is not detailed here. The most critical gap remains the absence of clinical trials to determine whether any of the statistical links reflect causation.

In all this activity, investigators have lost sight of important distinctions between clinical deficiency and SCCD (1, 19, 20), which are not merely sequential stations along a continuum of depletion. The sharp disparities begin with pathogenesis, which is clear in clinical deficiency but uncertain in most cases of SCCD. The causes usually dictate whether the deficiency is permanent and progressive (eg, pernicious anemia) or transient and ephemeral. The belief that SCCD must inevitably progress to clinical deficiency misappropriated the behavior of pernicious anemia and continues to color misperceptions about SCCD. The several courses proposed for SCCD (Figure 1) originate from scattered published observations but the natural histories of SCCD remain inadequately studied. Long-term longitudinal studies (21) have been rare.

SCCD usually lacks the severe IF-related malabsorption that leads inexorably to clinical deficiency. Absorption is normal in most cases, except for a mild and occasionally antibiotic-responsive malabsorption limited to food-bound cobalamin in 30–50% of cases (8, 10, 11, 14, 19). Claims of higher malabsorption rates (22) used dubious diagnostic criteria (fictitiously named “Carmel’s criteria”) instead of testing absorption (2). The limited role of malabsorption (14) may help explain the long latency and frequent spontaneous metabolic remissions in SCCD.

These differences illustrate why the facts and therapeutic imperatives of clinical deficiency do not apply automatically to SCCD. Conversely, population surveys generate data relevant mostly to SCCD that may translate poorly to clinical patients; the frequency of cobalamin-related biochemical abnormalities is ∼5–15% in the elderly (23–25), whereas the rate of pernicious anemia is only 1.9% (26). The increasingly routine reliance on metabolic testing in clinical medicine is an example of often wasteful translation from epidemiologic research (1). Macrocytosis, the hallmark and forerunner of the anemia of clinical cobalamin deficiency, is by definition irrelevant to SCCD and frequently nonspecific (27, 28); but population surveys often misattribute normocytic anemia to cobalamin deficiency (29–31).

Methodologic and interpretive concerns in SCCD

Achievement of the highest sensitivity and earliest detection of SCCD has captured substantial diagnostic attention. The advantages are uncertain, and the attempt increases risks of overdiagnosis because the much-needed diagnostic “gold” standard has remained elusive. Another concern for epidemiologic research is that relevant cobalamin deficiency exists only in very small subsets within large populations and applying broad statistical brushstrokes can obscure the critical minorities. Analyses that identify (and distinguish between) neglected subsets, such as persons who have clinical deficiency or malabsorption, may mitigate distortion and add valuable depth to the flat landscape of broad population analyses.
Macrocytosis or macrocytic anemia with elevated serum MMA or plasma total homocysteine produces a normal mean corpuscular volume. Microcytosis caused by iron deficiency or thalassemia that coexist with macrocytosis in 5–10% of cases of pernicious anemia; the combination of the 2 because they are key for its diagnosis. Maladies requires reclassification as clinical deficiency is not clear. Macrocytosis unrelated to cobalamin, such as alcohol abuse (the most common cause of macrocytosis), myelodysplastic syndrome, or copper deficiency.

### DIAGNOSTIC MARKERS

The preeminence of serum cobalamin as a diagnostic tool has come under challenge because of its reduced sensitivity for SCCD. This decrease is not surprising because most biomarkers lose sensitivity when applied to mild or marginal conditions. The lack of a gold standard complicates diagnostic evaluations and decisions, especially when diagnosis depends completely on biomarkers, as it does in SCCD. (Reliance on biomarkers is lower in the clinical arena, where clinical manifestations provide a diagnostic anchor and serum cobalamin often serves only a confirmatory function.) Because all biomarkers lack sufficient sensitivity and specificity to be reliable in SCCD, reliance on any single diagnostic test is risky in epidemiologic research and public health surveys. Which tests to use and how many results must be abnormal to provide reasonable diagnostic certainty of deficiency depend on many circumstances, and opinions vary.

All the biomarkers and their assays have idiosyncratic requirements and confounders, but some influences seem universal. Renal dysfunction elevates all biomarker concentrations; MMA and homocysteine correlate with even minor changes in creatinine and other indexes of glomerular filtration (32, 33), whereas cobalamin and holotranscobalamin II (holo-TC II) seem to rise significantly only with advanced renal dysfunction and by uncertain mechanisms (32, 34). Genetic influences on serum concentrations, transport, and use are becoming increasingly apparent. However, cutoff selection remains the most important external influence on biomarker data and requires careful, agenda-free science.

### SERUM COBALAMIN MEASUREMENT

#### Methodologic notes

Cobalamin assays require extraction of cobalamin from its binding proteins in plasma and conversion of the native cobalamins to cyanocobalamin, the nonphysiologic form that serves as the assay standard. Quantitation, initially based on the growth-promoting function of cobalamin on microbes, was replaced by radioisotopic competitive binding techniques using IF as the assay standard. Transcobalamin (TC) I, also called haptocorrin, whose affinity for nonfunctional corrin analogs produces spuriously high “cobalamin” values in sera that contain substantial amounts of such analogs, must not contaminate the IF (35).

Although a few reference laboratories prefer microbiologic assays, immunoenzymatic luminescence methods that rely on competitive binding by IF have now replaced isotopic assays. The newer methods need attention. They involve proprietary secrets, and their manufacturers do not always observe transparency (36) or track assay performance closely. A laboratory survey reported diagnostic misclassifications (37). Unfortunately, monitoring and proficiency testing are currently suboptimal because they do not use optimal low-cobalamin control specimens.
The most worrisome immunoenzymatic assay failures have been occasional systematic overestimates that seem limited to samples from overtly cobalamin-deficient patients, often with pernicious anemia, and without commensurate errors in normal sera (38–40). These errors have involved many manufacturers’ assays at various times; in one report, 3 of 5 commercial assays failed to give the expected low cobalamin results in one or both of 2 patients (40). Sometimes the error is demonstrably due to the kit’s failure to denature the sample’s anti-IF antibody, which competes with the kit’s anti-IF antibody. An error that affected 73% of cobalamin-deficient specimens was undetected for months (38). A clinical study attributed an unprecedented 86% rate of falsely normal cobalamin results to insensitivity of cobalamin assays at various times; in one report, 3 of 5 commercial assays failed to give the expected low cobalamin results in one or both of 2 patients (40). Sometimes the error is demonstrably due to the kit’s failure to denature the sample’s anti-IF antibody, which competes with the kit’s anti-IF antibody. An error that affected 73% of cobalamin-deficient specimens was undetected for months (38). A clinical study attributed an unprecedented 86% rate of falsely normal cobalamin results to insensitivity of cobalamin assays at various times; in one report, 3 of 5 commercial assays failed to give the expected low cobalamin results in one or both of 2 patients (40). Sometimes the error is demonstrably due to the kit’s failure to denature the sample’s anti-IF antibody, which competes with the kit’s anti-IF antibody. An error that affected 73% of cobalamin-deficient specimens was undetected for months (38). A clinical study attributed an unprecedented 86% rate of falsely normal cobalamin results to insensitivity of cobalamin assays at various times; in one report, 3 of 5 commercial assays failed to give the expected low cobalamin results in one or both of 2 patients (40). Sometimes the error is demonstrably due to the kit’s failure to denature the sample’s anti-IF antibody, which competes with the kit’s anti-IF antibody. An error that affected 73% of cobalamin-deficient specimens was undetected for months (38). A clinical study attributed an unprecedented 86% rate of falsely normal cobalamin results to insensitivity of cobalamin assays at various times; in one report, 3 of 5 commercial assays failed to give the expected low cobalamin results in one or both of 2 patients (40). Sometimes the error is demonstrably due to the kit’s failure to denature the sample’s anti-IF antibody, which competes with the kit’s anti-IF antibody. An error that affected 73% of cobalamin-deficient specimens was undetected for months (38). A clinical study attributed an unprecedented 86% rate of falsely normal cobalamin results to insensitivity of cobalamin assays at various times; in one report, 3 of 5 commercial assays failed to give the expected low cobalamin results in one or both of 2 patients (40). Sometimes the error is demonstrably due to the kit’s failure to denature the sample’s anti-IF antibody, which competes with the kit’s anti-IF antibody. An error that affected 73% of cobalamin-deficient specimens was undetected for months (38). A clinical study attributed an unprecedented 86% rate of falsely normal cobalamin results to insensitivity of cobalamin assays at various times; in one report, 3 of 5 commercial assays failed to give the expected low cobalamin results in one or both of 2 patients (40). Sometimes the error is demonstrably due to the kit’s failure to denature the sample’s anti-IF antibody, which competes with the kit’s anti-IF antibody. An error that affected 73% of cobalamin-deficient specimens was undetected for months (38). A clinical study attributed an unprecedented 86% rate of falsely normal cobalamin results to insensitivity of cobalamin assays at various times; in one report, 3 of 5 commercial assays failed to give the expected low cobalamin results in one or both of 2 patients (40). Sometimes the error is demonstrably due to the kit’s failure to denature the sample’s anti-IF antibody, which competes with the kit’s anti-IF antibody. An error that affected 73% of cobalamin-deficient specimens was undetected for months (38).
FIGURE 2. The influence of higher cutoffs for serum cobalamin concentrations on the frequency of true deficiency in the cases labeled as "deficient" by each cutoff. True deficiency was defined by each study's abnormal methylymalonic acid (MMA) and homocysteine results. Results of 4 surveys with metabolic data that provided the frequencies (or permitted their calculations) are shown here. This figure summarizes the cobalamin concentrations and cutoffs in ng/L ranges in the top line and in pmol/L in smaller font in the second line (1 ng = 0.738 pmol), which closely approximate the different points in all 4 studies. The arrowheads that bracket the frequency rates of metabolically defined deficiency in the center box (201–350 ng/L) delineate the relevant ranges of cobalamin values. The further data breakdowns available in reference 23 provide more discrete subset rates (italicized in smaller font).1Reference 47, 2Reference 23, 3Reference 46, 4Reference 24.

5–12% at traditional cutoffs. This impact of cutoff manipulation calls attention to the difficulties that surround the choices.

Other interpretive issues

Spuriously low cobalamin concentrations in pregnancy and fetal deficiency limit cobalamin specificity (4). Variations in plasma cobalamin-binding proteins can produce low or elevated cobalamin concentrations because TC I normally carries >80% of plasma cobalamin and often determines cobalamin concentrations (52). TC I deficiency may account for as much as 15% of low cobalamin concentrations (53); if so, it may cause more low cobalamin concentrations than does pernicious anemia. TC I deficiency often has a genetic origin (54); mutations in both alleles cause cobalamin concentrations of <100 pmol/L, whereas the more common heterozygous state causes mildly low or borderline cobalamin concentrations (53, 54).

Spuriously high cobalamin concentrations often accompany elevated TC I levels. The best known causes of dramatic elevations, such as chronic myelogenous leukemia and some cancers, are rare (4). Equally high cobalamin concentrations result from antibodies to TC II (often after hydroxocobalamin injections) or TC I (55–57), which may account for 8% of high cobalamin concentrations (58). More widespread are the more moderate cobalamin and TC I elevations found in renal failure (34), in blacks compared with other ethnic groups (59), and in association with homozygosity for α,1,2-fucosyltransferase gene polymorphisms (60).

SERUM MMA MEASUREMENT

The metabolic dependence of isomerization of L-methylmalonyl coenzyme A by methylmalonyl coenzyme A mutase with 5′-deoxyadenosylcobalamin as a cofactor, and its independence from folate metabolism, have made MMA an attractive biomarker of cobalamin insufficiency.

Methodologic notes

Gas chromatographic assay of MMA in urine has been available since the late 1950s (5) and a more sensitive gas chromatography/mass spectrometry method appeared in 1979 (61). Sensitive capillary gas chromatography/mass spectrometry assays with various modifications soon made it possible to measure small concentrations in serum accurately and precisely (15–18). As automation progressed, serum MMA determination became more widely available.

Performance characteristics

The sensitivity of serum MMA elevation for cobalamin deficiency in patients with overt pernicious anemia was usually >95% (15, 17, 18). Most MMA elevations were large, often in the thousands of nmol/L (15). In these and similar (62) clinical settings, MMA, like homocysteine and holo-TC II (62), performed only marginally better than serum cobalamin. All biomarkers

Table 2

Influence of diagnostic cutoff selection on the frequency of “abnormal” serum cobalamin concentrations in epidemiologic surveys of the elderly1

<table>
<thead>
<tr>
<th>Source location, year (reference)</th>
<th>Study population</th>
<th>n</th>
<th>Compared cobalamin cutoffs</th>
<th>Frequencies of “abnormal” cobalamin that resulted</th>
</tr>
</thead>
<tbody>
<tr>
<td>Denver, 1992 (47)</td>
<td>Elderly outpatients</td>
<td>152</td>
<td>&lt;148 vs &lt;221</td>
<td>8.5 vs 25.0</td>
</tr>
<tr>
<td>Framingham, 1994 (23)</td>
<td>Elderly town dwellers</td>
<td>548</td>
<td>&lt;148 vs &lt;228</td>
<td>5.3 vs 40.5</td>
</tr>
<tr>
<td>Netherlands, 1998 (46)</td>
<td>Elderly town dwellers</td>
<td>105</td>
<td>&lt;150 vs &lt;260</td>
<td>24.8 vs 60.1</td>
</tr>
<tr>
<td>Oklahoma City, 1998 (48)</td>
<td>Elderly outpatients</td>
<td>303</td>
<td>&lt;148 vs &lt;221</td>
<td>6.3 vs 16.2</td>
</tr>
<tr>
<td>United States, 1999 (49)</td>
<td>Disabled elderly town-dwelling women</td>
<td>762</td>
<td>&lt;148 vs &lt;228</td>
<td>6.2 vs 33.5</td>
</tr>
<tr>
<td>Los Angeles, 1999 (24)</td>
<td>Elderly town dwellers and outpatients</td>
<td>591</td>
<td>&lt;140 vs &lt;258</td>
<td>11.8 vs 50.4</td>
</tr>
<tr>
<td>United Kingdom, 2007 (25)</td>
<td>Elderly town dwellers in medical registries</td>
<td>2403</td>
<td>&lt;150 vs &lt;200 vs &lt;300</td>
<td>8.6 vs 29.3 vs 71.7</td>
</tr>
<tr>
<td>Norway, 2000 (50)</td>
<td>Adults aged 47–49 y</td>
<td>3684</td>
<td>&lt;150 vs &lt;200 vs &lt;400</td>
<td>0.4 vs 3.1 vs 64.5</td>
</tr>
<tr>
<td>Norway, 2000 (50)</td>
<td>Adults aged 71–74 y</td>
<td>3262</td>
<td>&lt;150 vs &lt;200 vs &lt;400</td>
<td>3.1 vs 6.7 vs 67.6</td>
</tr>
<tr>
<td>Georgia, 2010 (51)</td>
<td>Octogenarians</td>
<td>79</td>
<td>&lt;148 vs &lt;185 vs &lt;258</td>
<td>7.6 vs 11.4 vs 38.0</td>
</tr>
<tr>
<td>Georgia, 2010 (51)</td>
<td>Centenarians</td>
<td>215</td>
<td>&lt;148 vs &lt;185 vs &lt;258</td>
<td>11.6 vs 22.8 vs 39.1</td>
</tr>
</tbody>
</table>

1 All the surveys focused on elderly subjects, but the Norwegian survey (50) also included middle-aged adults.
appear to be less sensitive in SCCD, but no biochemical gold standard exists against which to compare them reliably. MMA values are usually only borderline or mildly abnormal in SCCD and, unlike in clinical deficiency, they are rarely >800 nmol/L.

MMA elevation is consistently more frequent than low cobalamin concentrations in population surveys (23, 24, 30, 50). Cutoff selections determine the frequency, but this disparity with cobalamin suggests MMA’s superior sensitivity for cobalamin deficiency (indeed, a normal MMA value supports the likelihood of normal cobalamin status even when cobalamin concentration is low). Nevertheless, at least some of the isolated MMA elevations reflect limitations in MMA specificity, which has been difficult to determine reliably. A large Norwegian survey identified creatinine, cobalamin, age, and sex as the major identifiable influences on serum MMA, but these explained only 16% of MMA variation (50). A smaller study showed that cobalamin predicted only 2% of the serum MMA value, whereas cystatin C, an index of glomerular filtration, predicted 8% (33).

Additional influences on MMA surely exist and require systematic study. Some authors have suggested, but not documented, a role for plasma volume contraction in some unexplained MMA elevations (15). Antibiotics sometimes lower MMA concentrations, which suggests influences from active propionate metabolism by intestinal bacteria (18, 63). Neither heterozygosity for genetic methylmalonic aciduria (64) nor MMA-related polymorphisms seem, thus far, to elevate serum MMA significantly. The observation of elevated MMA concentrations during the first year of life (65) remains a puzzle. The values return to normal spontaneously, but statistical association with lower cobalamin and higher homocysteine values (both still normal) and reduction of MMA concentrations by cobalamin therapy raise the possibility of mild cobalamin deficiency (66).

Given its undetermined specificity, MMA cannot be the gold standard for cobalamin deficiency (eg, as used in the holo-TC II literature). Moreover, a follow-up of 432 untreated, asymptomatic patients with MMA >280 nmol/L showed that MMA fluctuated over 1–3.9 y (CV: 34%) (21). MMA elevation progressed in only 16% of cases, compared with spontaneous decline in 44% and no change in the remainder. As these findings suggest, minor changes, whether spontaneous or after treatment (41), should not be overinterpreted (42).

The influence of cutoff selection

Published MMA cutoffs have ranged from 210 to 480 nmol/L, sometimes within the same laboratory. Frequencies of MMA abnormality, and therefore by extension often of SCDD, vary inversely with the selected cutoff. The Hordaland data (50) aptly show the effects of cutoff migration (Table 3). As another example, the MMA cutoffs in the 4 surveys in Figure 2 (23, 24, 46, 47) ranged from 240 to 370 nmol/L, and their 5–44% frequencies of MMA abnormality diagnose bivariate second-order to the cutoffs.

The most commonly applied MMA cutoff is ≈270 nmol/L. Many laboratories defined cutoffs by 3 or 2 SD from the mean (≈>370 or >270 nmol/L, respectively). Others made physiologic choices (eg, demonstration or projection of MMA concentrations in cobalamin-repleted subjects) as low as 210 nmol/L. (67). The maximally suppressed MMA concentration might represent true normality but it is not necessarily practical or desirable in clinical or epidemiologic settings. Such MMA response can help, as it helps when assessing therapeutic effectiveness (17), but abnormal values can also regress to the mean. It is also possible that some improvements of mildly elevated or even normal MMA merely signify fuller saturation of the methylmalonyl coenzyme A mutase by its cobalamin cofactor. On balance, when studies use MMA alone to define cobalamin deficiency, unusually low MMA cutoffs such as 210 nmol/L (30, 31) compound the risk of overdiagnosis (R Carmel, unpublished observations, 2010). A potential data-based approach to defining biomarker cutoffs may be to construct 2-phase regression models to determine where their slopes change in relation to other biomarkers (68).

### Table 3

Influence of diagnostic cutoff selection on frequency of “abnormal” serum methylmalonic acid concentrations, as illustrated in middle-aged and elderly subjects in the Hordaland survey

<table>
<thead>
<tr>
<th>Selected MMA cutoff</th>
<th>3684 adults aged 47–74 y</th>
<th>3262 adults aged 71–74 y</th>
</tr>
</thead>
<tbody>
<tr>
<td>&gt;210 nmol/L</td>
<td>14.7</td>
<td>36.7</td>
</tr>
<tr>
<td>&gt;260 nmol/L</td>
<td>5.3</td>
<td>17.7</td>
</tr>
<tr>
<td>&gt;370 nmol/L</td>
<td>1.2</td>
<td>5.2</td>
</tr>
<tr>
<td>&gt;750 nmol/L</td>
<td>0.1</td>
<td>0.9</td>
</tr>
</tbody>
</table>

1 This survey showed lower rates of cobalamin deficiency than most population surveys to date. Data are from reference 50. MMA, methylmalonic acid.

2 The frequencies of “abnormal” results with the use of these cutoffs may be higher in other populations.

3 This highly stringent cutoff is also likely to select a much higher proportion of subjects with clinical deficiency that arises from malabsorption (eg, pernicious anemia) and a lower proportion of subjects with subclinical deficiency than do lower cutoffs.

### HOLOTRANS Cobalamin II

TC I and TC II (those who call TC I haptocorrin refer to TC II as TC) carry cobalamin in the bloodstream. Holo-TC II, the TC II fraction with attached cobalamin, forms at the ileal cell and delivers its cobalamin to tissues so rapidly that <20–30% of plasma cobalamin normally exists as holo-TC II at any time. The remaining cobalamin circulates in holo-TC I, which turns over slowly with no discernible delivery to cells and is considered metabolically unavailable.

Two subtly different and, in fact, competing diagnostic claims about holo-TC II (69) are that low concentrations either indicate impaired holo-TC II transfer into the bloodstream from the ileal cell (ie, a marker of cobalamin malabsorption) or denote inadequate availability to tissues (ie, metabolic insufficiency). Both claims may be partially true (70), an ambiguity that ultimately renders the test nonspecific for either claim. Malabsorption and deficiency can and often do exist without each other (69). Transient (eg, drug-induced) malabsorption might exist for only a few weeks or months and produce low holo-TC II concentrations without ever creating cobalamin deficiency, which requires
years of storage depletion. To complicate holo-TC II status further, the gut may not be the sole contributor to circulating holo-TC II (69).

Some authors view holo-TC II abnormality as the earliest possible marker of cobalamin deficiency because it occurs in patients with normal concentrations of all other biomarkers. Current research cannot prove or disprove such conclusions, which court tautology. The full range of alternative influences on holo-TC II concentrations has not been adequately explored.

Methodologic notes
Lindemans et al (71) suggested in 1983 that measurement of holo-TC II, whose cobalamin is available to cells, may provide a truer snapshot of cobalamin status than total cobalamin. Documentation was limited for many years because holo-TC II concentrations are very small and early methods were crude and often indirect (69). Reliable study became possible with the reports of an accurate enzyme-linked immunosorbent assay and radioimmunoassay in 2002 (72, 73). An automated 2-step sandwich microparticle enzyme immunoassay centered on an antibody to holo-TC II (74) replaced the radioimmunoassay in 2008. Reference intervals are similar but not identical in the 3 methods; plasma samples might also yield slightly higher values than do serum samples.

Performance characteristics
Careful clinical assessment has been limited, in part because cobalamin absorption testing became unavailable after the late 1990s. Reliable studies must resolve what holo-TC II reflects and explore potential confounders of results.

Performance assessment has largely relied on statistical comparisons in populations that did not include well-characterized patients. Studies compared sensitivity and specificity with those of cobalamin and often used MMA as the benchmark despite its undefined specificity. Most receiver operator characteristics–based comparisons and similar analyses of sensitivity and specificity have shown that holo-TC II modestly outperforms total cobalamin, with areas under the curves of 0.75–0.90 and 0.72–0.85, respectively (25, 62, 75, 76). One study (25) concluded that neither test is suitable for the screening of asymptomatic populations (ie, SCCD) because false-positive results outnumber true-positive ones. A clinical study in 49 patients with early but clinically expressed deficiency concluded that neither metabolic biomarkers nor holo-TC II outperformed cobalamin as a predictor of response to cobalamin therapy (62).

The clinical specificity of holo-TC II, as opposed to its comparability against MMA, is unclear. Holo-TC II rises in renal failure (34), but modest renal insufficiency does not affect holo-TC II or total cobalamin as much as it affects MMA and total homocysteine concentrations (32). Holo-TC II becomes normal in patients with pernicious anemia after correction of their cobalamin deficiency but continues to be mildly but significantly lower than in control subjects (70), which supports absorption status as an independent influence on holo-TC II values. Whether prolonged fasting can similarly affect holo-TC II concentrations is unknown. A frequent genetic variant, the substitution of proline for arginine in TCN2 codon 259, may impair holo-TC II function; its effect on serum holo-TC II concentrations is unsettled (77–79). Suggested but unexplored clinical confounders of holo-TC II include oral contraceptive use, myelodysplasia and other hematologic conditions, histiocytic disorders, folate disorders, and alcoholism (76, 80, 81). A proposed diagnostic advantage for holo-TC II over cobalamin in pregnancy (82) is inconclusive: holo-TC II was not measured before pregnancy but its concentrations rose postpartum, which suggests that an unrecognized holo-TC II decline had, in fact, occurred during pregnancy.

The influence of cutoff choices
Laboratories have used holo-TC II cutoffs as high as 50 pmol/L. A multicenter evaluation of laboratory performance noted cutoffs between 11 and 41 pmol/L in 8 studies, and the authors favored the higher values as more representative of normality (76). However, the automated assay that dominates the commercial market identified a 95% CI cutoff of 19 pmol/L (74), which was even lower than the 24 pmol/L cutoff in its predecessor assay (72). The influence of cutoffs on outcomes awaits formal study. Only one group has suggested a sex difference in values, with a lower cutoff for women aged <45 y (83).

Other interpretive issues
Miller et al (75) reported that holo-TC II and total cobalamin results agreed in 93% of cases (Table 4, groups 1 and 4). The crux to understanding how the 2 biomarkers’ performances differ lies within the remaining 7%, of whom only 1 in the 2 markers indicated cobalamin deficiency. Clinical, metabolic, and absorptive investigation of this subset would be invaluable.

PLASMA TOTAL HOMOCYSTEINE
Plasma homocysteine has largely comparable sensitivity to MMA for cobalamin deficiency and is also excellent for monitoring response to cobalamin therapy (17). However, homocysteine is too limited by its confounders to serve as a cobalamin status biomarker and I will not discuss it in detail. Its equal susceptibility to folate changes prevents its use to distinguish between cobalamin and folate deficiencies in epidemiologic surveys such as the National Health and Nutrition Examination Survey. Although folic acid fortification may have left cobalamin deficiency as the chief influence on homocysteine concentrations among the elderly in the United States (84), folate status explains

<table>
<thead>
<tr>
<th>Group</th>
<th>Cobalamin results</th>
<th>Holo-TC II results</th>
<th>n</th>
<th>Frequency in study population %</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Low</td>
<td>Low</td>
<td>32</td>
<td>5.3</td>
</tr>
<tr>
<td>2</td>
<td>Normal</td>
<td>Low</td>
<td>28</td>
<td>4.6</td>
</tr>
<tr>
<td>3</td>
<td>Low</td>
<td>Normal</td>
<td>14</td>
<td>2.3</td>
</tr>
<tr>
<td>4</td>
<td>Normal</td>
<td>Normal</td>
<td>533</td>
<td>87.8</td>
</tr>
</tbody>
</table>

1 The cutoff used to differentiate low from normal cobalamin results was 148 pmol/L. The cutoff used to differentiate low from normal holo-TC II results was 35 pmol/L. Data are from reference 73.
15% of homocysteine variation, compared with 4% for cobalamin status in European countries without fortification (33). Homocysteine is also the most susceptible of the 4 biomarkers to preanalytic influences of sample collection and processing, and it is the only one with sex-related differences in reference intervals that require stratified analyses. A review of expert opinion provides further discussion of the homocysteine assay and its interpretations (85).

CONCLUSIONS AND RECOMMENDATIONS

Issues of concepts and goals

Both public health experts and clinicians must consider all the implications of SCCD’s heavy predominance in population surveys in contrast to medical encounters. SCCD differs fundamentally, not just in degree, from clinical deficiency. That SCCD rarely arises from IF-related malabsorption may explain its unlikely or delayed progression to clinical deficiency. The differences make unfiltered equations of findings and expectations between the public health and clinical sectors potentially confusing for both sectors. Epidemiologists have reported potentially important associations between SCCD and neurocognitive changes, but only prospective clinical trials can transform statistical associations into causative health risks and prove that intervention brings benefits. Randomized trials may be unethical in clinical cobalamin deficiency, but they are ethical, and indeed imperative, in SCCD.

More comprehensive efforts must be marshalled to satisfy the US Food and Drug Administration’s charge that researchers monitor and clarify folic acid fortification’s effect on cobalamin deficiency. Despite some intrinsic limitations, only large surveys can identify enough potential subjects. However, proper focus on the population subsets at greatest risk may require means to differentiate between normallaboratory SCCD and incipient stages of pernicious anemia (28); more diverse and sensitive neurologic and cognitive tests are also needed.

Epidemiologic focus should extend to cobalamin status in young children (MMA is often high in the first year of life), persons with *Helicobacter pylori* infection, and chronic users of drugs that suppress gastric acid secretion. Attention to cobalamin excess and its effects may also become necessary. Daily 1000–2000 μg cyanocobalamin self-supplementation has grown, and this nonphysiologic cobalamin can accumulate substantially in cells under such circumstances (86). Red cells could become a valuable research resource.

Methodologic and interpretive issues

The greatest intrinsic source of diagnostic uncertainty in cobalamin status assessment is the absence of a gold standard for deficiency. This barrier and the limitations of all markers dictate that researchers no longer rely on any single test. A model for combining tests in all future research, preferably with the use of markers of different cobalamin properties and nonoverlapping confounder influences, is presented in Table 5. Studies can combine a marker of cobalamin content, such as serum cobalamin or holo-TC II, with a metabolic function indicator, such as serum MMA. Despite minor statistical advantages over cobalamin, holo-TC II awaits systematic study of subjects defined by relevant characteristics that include absorption status, not just by their MMA status, and of confounding influences on holo-TC II results. When considering everything, the best diagnostic combination currently appears to be cobalamin and MMA. Reliable diagnosis of deficiency requires abnormal results in both chosen tests. Basing diagnoses on an abnormality of either test alone compounds their nonspecificities and guarantees overdiagnosis of deficiency. Research must also better assess the specificity of MMA elevation in SCCD.

The greatest extrinsic source of diagnostic uncertainty (and noncomparability of studies) for all biomarkers is the highly variable use of cutoffs. Experts should develop guidelines for appropriate criteria and methods that laboratories can use when defining cutoffs.

<table>
<thead>
<tr>
<th>Diagnostic interpretation</th>
</tr>
</thead>
<tbody>
<tr>
<td>Abnormal</td>
</tr>
<tr>
<td>Normal</td>
</tr>
<tr>
<td>Abnormal</td>
</tr>
<tr>
<td>Normal</td>
</tr>
</tbody>
</table>

**TABLE 5**

Suggested principles for optimal diagnostic testing for cobalamin deficiency in epidemiologic research in the absence of a diagnostic gold standard

- Principles and assumptions are as follows: 1) Reliance on one biomarker alone, whether the result is normal or abnormal, is inadequate for reliable biochemical diagnosis when clinical assessment is unavailable or clinical deficiency is absent. 2) The use of all 4 biomarkers, although informative sometimes, is excessive and multiplies the likelihood of diagnostic categorization disagreements. 3) The choice should include at least one indicator of vitamin amount [total cobalamin or holo-transcobalamin II (holo-TC II)] and at least one indicator of metabolic function [methylmalonic acid (MMA) or homocysteine]. 4) Plasma homocysteine, although reasonably sensitive, has disabling specificity problems that prevent its selection as one of the key biomarkers. 5) Although holo-TC II has slightly better sensitivity and specificity than total cobalamin, the sparse information on alternative influences on holo-TC II (ie, diagnostic confounding and specificity) makes cobalamin a superior choice. 6) This schema is less applicable to clinical cobalamin deficiency than to subclinical cobalamin deficiency because clinical findings can modify some of the interpretations. The diagnostic interpretations shown are consistent with the principle that biochemical diagnosis of deficiency can be certain only when more than one biochemical abnormality is evident. In some cases, this rule might miss a diagnosis.

- Serum cobalamin is preferable to holo-TC II because research has better characterized it and its shortcomings. If holo-TC II testing is substituted for cobalamin testing in this column, the diagnostic interpretations in the right-hand column will be similar.

- The normal cobalamin result in the face of an abnormal MMA result suggests either metabolic cobalamin deficiency with a falsely normal serum cobalamin concentration or nondeficiency with a falsely elevated MMA concentration. The odds may favor the former formulation but the uncertainty mandates against a diagnosis of deficiency in a research study. (In the clinical setting, other information on the patient can bear on the diagnostic interpretation; in a research study, the possibility of doing so depends on the study design.)

- Falsely low cobalamin: suspect transcobalamin I deficiency or folate deficiency if the subject is not pregnant. (If holo-TC II was tested instead of cobalamin, the diagnostic possibilities include a falsely low holo-TC II, whose causes remain poorly defined, and cobalamin malabsorption without cobalamin deficiency.)
The current cobalamin assay methods also require attention. The limited transparency of commercial sources and possibly frequent malfunctions of many immunoenzymatic methods that disproportionately affect samples from subjects with cobalamin deficiency and can easily escape detection are disquieting. The inclusion of control samples that contain anti-IF antibody may improve detection of the error. The microbiological cobalamin assay may be a viable alternative for large reference laboratories.

Planning for future capabilities

The ability to identify IF-related malabsorption, primarily pernicious anemia, would greatly improve the categorization of cobalamin deficiency in surveys. Such information is likely to be as relevant to deficiency as the biochemical expression of deficiency, and may also hold the key to understanding the putative neurologic complications when folic acid intake is high (28). Inclusion of the anti-IF antibody assay, despite its low sensitivity, seems the best available option after the disappearance of reliable absorption tests. The antibody test has near-100% specificity for pernicious anemia and requires only a serum sample, and laboratories may measure it post hoc in samples with low cobalamin concentrations. An improved capability for measurement of non-functional analogs of cobalamin, which can accumulate in some subjects and may have adverse implications, may also be useful to future surveys.

The author had responsibility for all parts of the manuscript. He has submitted a patent application for work related to transcobalamin I deficiency and its genetic aspects. The author did not declare a conflict of interest.

REFERENCES

28. Carmel R. Does high folic acid intake affect unrecognized cobalamin deficiency, and how will we know it if we see it? Am J Clin Nutr 2009;90:1449–50.
35. Kolhouse JF, Kondo H, Allen NC, Podell E, Allen RH. Cobalamin analogues are present in human plasma and can mask cobalamin deficiency.


