HHS Public Access

Author manuscript

Am J Clin Nutr. Author manuscript; available in PMC 2023 March 29.

Published in final edited form as:

Am J Clin Nutr. 2020 March 01; 111(3): 601–612. doi:10.1093/ajcn/nqz307.

Population RBC folate concentrations can be accurately estimated from measured whole blood folate, measured hemoglobin, and predicted serum folate—cross-sectional data from the NHANES 1988–2010

Ming Zhang[1], **Maya R Sternberg**[1], **Lorraine F Yeung**[2], **Christine M Pfeiffer**[1]

[1]National Center for Environmental Health, CDC, Atlanta, GA, USA

[2]National Center on Birth Defects and Developmental Disabilities, CDC, Atlanta, GA, USA

Abstract

Background: RBC folate (RBF) is an indicator of folate status and risk of neural-tube defects. It is calculated from whole blood folate (WBF), serum folate (SFOL), and hematocrit (Hct). SFOL and/or Hct are sometimes unavailable; hemoglobin (Hb) is generally available in surveys.

Objectives: We assessed the ability of different RBF approximations to generate population data in women aged 12–49 y.

Methods: Using SFOL, RBF, Hct, Hb, and mean corpuscular Hb content (MCHC) from prefortification (1988–1994) and postfortification (1999–2006, 2007–2010) NHANES we applied 6 approaches: *#1*) assume SFOL = 0; *#2*) impute SFOL (population median); *#3*) impute Hct (population median); *#4*) estimate Hct (Hb/MCHC); *#5*) assume SFOL = 0 and estimate Hct; and *#6*) predict SFOL (from WBF) and estimate Hct. For each approach, we calculated the paired percentage difference to the "true" RBF and estimated various statistics.

Results: For 2007–2010 (unweighted data), the median relative difference from "true" RBF was lowest for approaches #2 (−0.74%), #4 (−0.96%), and #6 (−1.15%), intermediate for #3 (−3.36%), and highest for #5 (4.96%) and #1 (5.78%). The 95% agreement limits were smallest for approach #1 (2.33%, 13.0%) and largest for #3 (−20.8%, 11.3%). Approach #2 showed concentration-dependence (negative compared with positive differences at low compared with high RBF). Using weighted data, we found similar patterns across approaches for mean relative differences by demographic subgroup for all 3 time periods.

Address correspondence to CMP (cfp8@cdc.gov).

The authors' responsibilities were as follows—CMP and MZ: developed the study concept and were responsible for drafting the manuscript; MRS: performed the statistical data analysis; and all authors: were responsible for critical revision of the manuscript and read and approved the final manuscript.

The authors report no conflicts of interest.

The findings and conclusions in this report are those of the authors and do not necessarily represent the official views or positions of the CDC/Agency for Toxic Substances and Disease Registry.

Data described in the article, code book, and analytic code are publicly and freely available without restriction at https://wwwn.cdc.gov/nchs/nhanes/Default.aspx.

Supplemental Tables 1–6 and Supplemental Figures 1–2 are available from the "Supplementary data" link in the online posting of the article and from the same link in the online table of contents at https://academic.oup.com/ajcn/.

Conclusions: We obtained the best agreement between estimated and "true" RBF when we predicted SFOL using a regression equation obtained from a subset of samples (approach #6). Alternatively, the consistent overestimation of RBF when assuming SFOL = 0 (~6%) could be addressed by adjusting the data (approach #5). Similar observations for pre- and postfortification periods suggest applicability to low and high folate status situations, but should be confirmed elsewhere. To estimate RBF, at least WBF and Hb are needed. *Am J Clin Nutr* 2020;111:601–612.

Keywords

NHANES; microbiologic assay; radioassay; hematocrit; low-resource countries

Introduction

National nutrition surveys often assess folate status because folate deficiency and insufficiency cause public health problems such as increased risks of megaloblastic anemia and neural-tube defects (NTDs), respectively (1). An insufficiency cutoff for RBC folate (RBF) concentrations in women of reproductive age has been available since 2015 to assess the risk of NTDs (2). Countries have become more interested in assessing population folate status and monitoring the effects of folate interventions. There are limitations in using plasma folate as an indicator of NTD risk, because the relation between RBF and plasma folate is modified by BMI, genotype, and substantially by low plasma vitamin B-12 (3).

RBF is a good indicator of tissue stores and long-term status (4). Typically, whole blood is hemolyzed with ascorbic acid solution followed by prompt freezing of the hemolysate. To accurately calculate RBF (conventional formula), the measured whole blood folate (WBF) is corrected for the serum folate (SFOL) contribution and normalized to the proportion of packed red cells as measured by hematocrit (Hct), i.e., [WBF − SFOL*(1 − Hct)]/Hct. This requires extra resources in terms of specimen volume (to ensure sufficient blood for serum and whole blood hemolysate), time, and cost (to conduct the analysis of Hct and folate for 2 specimens for thousands of samples).

Thus, the question has arisen as to whether acceptably accurate RBF results can be obtained using different approaches. This is relevant to the nutrition community and particularly to low- and middle-income countries planning nutrition surveys. Because SFOL concentrations are much lower than those of RBF, using a simplified RBF calculation formula that ignores SFOL (i.e., assumes SFOL = 0) is an option (5, 6). However, failure to correct RBF for the SFOL contribution can give a false picture of the clinical folate status, especially in patients with high SFOL concentrations (7). Hemoglobin (Hb) concentration and packed cell volume (i.e., Hct) show a strong linear correlation only marginally affected by the mean corpuscular volume (8). Thus, Hb survey data can be used instead of measuring Hct. We are not aware of a systematic evaluation of different RBF calculation approaches with 1 or 2 components missing in the formula.

Our main objective was to use NHANES data to assess the comparability of conventionally calculated RBF, referred to as "true" RBF, with RBF approximations from 6 approaches. Because of the interest in folate insufficiency, our study focused on women of reproductive age in NHANES 2007–2010 (microbiologic assay). We also included data from pre– (1988–

1994) and post–folic acid fortification (1999–2006) periods (radioassay) to cover a broader folate concentration range and a different assay type. In a supplemental analysis for persons aged ≥ 1 y (2007–2010) we expanded to other demographic groups. We first assessed the comparability of various computational strategies to the "true" RBF using unweighted data and then used weighted statistics to interpret the findings in different population subgroups.

Methods

Participants and survey design

The NHANES has been collecting cross-sectional data on the health and nutritional status of the civilian noninstitutionalized US population, first as periodic surveys (from the early 1970s to the middle 1990s) and since 1999 as a continuous survey conducted in 2-y survey periods. The survey is conducted by the National Center for Health Statistics at the CDC and has a stratified, multistage, probability sample design. The NHANES combines home interviews with health tests performed in a Mobile Examination Center, where biologic samples are collected for biochemical analyses. Interview and examination response rates for each survey period are publicly available (9). All respondents gave their informed consent, and the NHANES protocol was approved by the National Center for Health Statistics Research Ethics Review Board.

Biomarker measurement

Serum and whole blood hemolysate samples were analyzed for folate by a CDC laboratory using different methods over time: the Bio-Rad Quantaphase I radioassay (1988–1991), the Quantaphase II radioassay (1991–2006), and the microbiologic assay (2007–2010) (10). Assay adjustments were applied to the 1988–1991 folate data before their public release to account for method differences between the Quantaphase I and II (11). The Quantaphase II radioassay produced results that were 29% lower than the microbiologic assay for serum samples and 45% lower for whole blood samples (12). We did not use the published regression equations to adjust the Quantaphase II data to the microbiologic assay (12) because our goal was to assess how each of the 6 approximations behaved within each time period and assay. We were not interested to compare folate concentrations over time standardized to a common assay. Long-term CVs were 4.0–7.0% for serum folate and 4.0–6.0% for RBF for the radioassay, and 4.7–10% for serum folate and 7.5–14% for RBF for the microbiologic assay (10). Hb, Hct, and mean corpuscular Hb content (MCHC) were measured as part of the complete blood count using a 5-part differential Beckman Coulter method.

Study variables

We used publicly available data for the following biomarkers: SFOL, RBF, Hct, Hb, and MCHC. Owing to the special interest in folate status in women of reproductive age, our main analysis used data from women aged 12–49 y from the pre–folic acid fortification period (1988–1994), the early postfortification period (1999–2006), and the late postfortification period (2007–2010). For brevity, we refer to this group as "women." We categorized the data into 3 age groups (12–19, 20–39, and 40–49 y) and 3 race/Hispanic origin groups (Mexican American, non-Hispanic black, and non-Hispanic white).

In a supplemental analysis, we used data from all participants aged ≥1 y limited to the late postfortification period (2007–2010). We categorized the data into 6 age groups (1–5, 6–11, 12–19, 20–39, 40–59, and ≥ 60 y), 2 sex groups (male and female), and 3 race/Hispanic origin groups (Mexican American, non-Hispanic black, and non-Hispanic white). For 2007–2010, we also examined blood folate concentrations in women by fasting status (self-reported time since last food or drink was consumed; <3, 3 to <8, and ≥8 h) and use of folic acid–containing dietary supplements (self-reported consumption of supplements in the past 30 d; yes and no), 2 variables previously shown to be associated with folate biomarker concentrations (13, 14).

Statistical analysis

Statistical analyses were carried out using SAS for Windows software version 9.4 (SAS Institute) and SAS callable SUDAAN software version 11 (RTI) to account for the complex survey design. We used multiyear Mobile Examination Center survey weights to account for the unequal probabilities of selection, adjustment for nonresponse, and poststratification. SUDAAN uses Taylor series linearization to calculate variance estimates. We did not exclude any participants but restricted our data set to participants with complete data for all the variables (Supplemental Figure 1). For the main analysis, our analytical sample consisted of 6275 (1988–1994), 9236 (1999–2006), and 3951 (2007–2010) women with complete biomarker data. For the supplemental analysis, our analytical sample consisted of 16,815 (2007–2010) persons ≥1 y of age with complete biomarker data.

To prepare the data for statistical analysis, we conducted 2 steps: we estimated the population median values for the variables of interest (SFOL, Hct, and MCHC) to be used as imputed values in RBF estimation; and we estimated RBF for each NHANES participant using the 6 approaches. For the first step, we calculated the measured WBF concentration for each participant using the publicly available data for RBF, SFOL, and Hct: WBF = (RBF*Hct) + SFOL*(1.0 − Hct); Hct (%) was converted to a decimal. We then calculated median (95% CI) concentrations for folate (in RBCs, whole blood, serum, and predicted serum) and hematologic biomarkers (Hct, Hb, MCHC, and estimated Hct) overall and by demographic subgroup for each time period (weighted). To assess the quality of the predicted SFOL and estimated Hct, we plotted the relative difference between the predicted and measured SFOL and between the estimated and measured Hct for all 3 time periods (Bland–Altman).

In the second step, we calculated the approximated RBF using the 6 approaches, manipulating either 1 or 2 variables at a time (Figure 1). In approach #1, we assumed SFOL = 0, which is equivalent to the simple formula for RBF (WBF/Hct). In approach #2, we imputed SFOL at an approximate estimate of the population median [10 nmol/L (1988–1994), 30 nmol/L (1999–2006), and 40 nmol/L (2007–2010)]. The actual weighted median SFOL concentrations across time and population were as follows: women (1988–1994), 11.2 nmol/L; women (1999–2006), 26.6 nmol/L; women (2007–2010), 37.1 nmol/L; persons ≥ 1 y (2007–2010), 39.8 nmol/L. In approach #3, we imputed Hct at 40% (0.4), the approximate estimate of the population median. There were only small differences in the actual weighted median Hct across time and population: women (1988–1994), 39.3%; women (1999–2006),

39.8%; women (2007–2010), 38.9%; persons ≥ 1 y (2007–2010), 40.8%. In approach #4, we estimated Hct as the measured Hb divided by an imputed MCHC of 340 g/L, a value near the estimated population median MCHC. There were only small differences in the actual median MCHC across time and population: women (1988–1994), 337 g/L; women (1999–2006), 339 g/L; women (2007–2010), 343 g/L; persons ≥ 1 y (2007–2010), 344 g/L. In approach #5, we assumed SFOL = 0 and estimated Hct (Hb/MCHC).

Finally, in approach #6, we predicted SFOL and estimated Hct (Hb/MCHC). SFOL was predicted separately for NHANES 1988–1994, 1999–2006, and 2007–2010 using unweighted simple linear regression derived from measured WBF as the *x*-variable and measured SFOL as the *y*-variable. For the main analysis in women, we obtained a simple random sample of 150 women between the ages of 12 and 49 y from each NHANES time period to predict SFOL (Supplemental Table 1). For the supplemental analysis in persons aged ≥1 y for 2007–2010, we used WBF data from a simple random sample of 150 persons aged ≥1 y to predict SFOL (Supplemental Table 1).

The first goal of our statistical analysis was to describe the impact of the different estimation approaches and how they compared to the "true" RBF by using unweighted analysis. For each of the 6 approaches, we plotted the paired relative difference between the estimated and "true" RBF values against the mean of the 2 RBF values (Bland–Altman) and reported the nonparametric 95% agreement limits and median relative difference (2.5th, 5th, 50th, 95th, and 97.5th). We also calculated for each approach the percentage of NHANES participants with estimated RBF concentrations >±10% relative difference, arbitrarily defined as larger relative difference. Our second goal was to describe estimated RBF concentrations by demographic variables representative of the US population. For this, we calculated the weighted RBF geometric mean overall and by demographic group for each approach. We also calculated the weighted mean percentage difference to the "true" RBF value overall and by demographic group. For approaches #5 and #6 that address both variables of interest (SFOL and Hct), we performed 2 additional calculations using data from NHANES 2007–2010 (women): the mean percentage difference to the "true" RBF value by fasting status or supplement use and the central 95% reference interval by demographic subgroup. Lastly, we calculated for each of the 6 approaches the weighted prevalence of low folate status using different RBF cutoffs [<305 nmol/L for folate deficiency representing risk of megaloblastic anemia and <748 nmol/L for folate insufficiency representing risk of NTDs (15)], using data from NHANES 2007–2010. We evaluated the diagnostic characteristics of each approach for folate insufficiency by calculating the sensitivity and specificity relative to using the "true" RBF value.

Results

Characteristics of the components used for RBF calculation

Weighted median concentrations of RBF, WBF, and SFOL in women overall and by age and race/Hispanic origin group showed the lowest values during the 1988–1994 prefortification period (measured by radioassay), intermediate values during the 1999–2006 early postfortification period (also measured by radioassay), and the highest values during the 2007–2010 late postfortification period (measured by microbiologic assay which

generates higher results than the radioassay) (Table 1). The weighted median predicted SFOL concentrations (using WBF data from a random subset of women to predict SFOL from each time period) appeared to be generally higher than the measured SFOL concentrations, except for women 12–19 y of age. The Bland–Altman plot showed a close to zero unweighted median relative difference between the predicted and measured SFOL values for the 3 time periods, but wide 95% agreement limits (1988–1994: 13%; −48%, 177%; 1999–2006: 1%; −48%, 125%; and 2007–2010: 5%; −46%, 153%) (Supplemental Figure 2A–C). Furthermore, each plot showed a concentration-dependent pattern.

Weighted median values for the 3 measured hematologic variables (Hct, Hb, and MCHC) were similar across the 3 time periods and generally similar by age or race/Hispanic origin group, except for non-Hispanic black women (Table 2). The weighted median estimated Hct (calculated from the measured Hb and an approximate MCHC of 340 g/L) appeared to be similar to the measured Hct. The Bland–Altman plot showed an unweighted median (95% agreement limit) relative difference between the estimated and measured Hct of −1.5% (−7.8%, 3.2%), −0.2% (−5.9%, 4.1%), and 1.0% (−5.4%, 6.0%) for the 3 successive time periods, indicating a fairly narrow range among the differences (Supplemental Figure 2D–F). Each plot showed a slight concentration-dependent pattern.

Estimated RBF in women for NHANES 2007–2010 (unweighted analysis)

The Bland–Altman plots (Figure 2) showed fairly consistent differences across the RBF concentration distribution, except for approach #2, which showed noticeable concentration-dependence, with negative differences at low and positive differences at high RBF concentrations. The median relative difference was lowest for approaches #2 (−0.74%), #4 (−0.96%), and #6 (−1.15%), intermediate for approach #3 (−3.36%), and highest for approaches #5 (4.96%) and #1 (5.78%) (Table 3). The widths of the agreement limits covering 95% of the samples were narrowest for approaches #1 (2.33%, 13.0%) and #4 (−5.44%, 5.46%), intermediate for approaches #2 (−8.62%, 54.49%), #5 (−1.19%, 14.7%), and #6 (−6.93%, 7.14%), and widest for approach #3 (−20.8%, 11.3%). The percentage of women having an estimated RBF concentration exceeding ±10% relative difference was <2% for approaches #4 and #6 and 2.4% for approach #2. On the other hand, both approaches #1 and #5 overestimated the true RBF by 9% and 12%, respectively, with minimal underestimation (<1%). Approach #3 underestimated the true RBF by 18% and overestimated by 4%.

Estimated RBF in women by demographic group for NHANES 1988–2010 (weighted analysis)

The RBF geometric means for the 6 approaches showed similar patterns for each of the 3 time periods, both overall and by demographic group (Table 4). The weighted mean relative differences between the estimated and “true” RBF values using the 6 approaches showed similar patterns within each of the 3 time periods (Table 5). Approaches #1 (assume SFOL = 0) and #5 (assume SFOL = 0 and estimate Hct) produced similar outcomes with overall differences of 5.5% and 6.6% (1988–1994), 7.6% and 7.8% (1999–2006), and 6.2% and 5.1% (2007–2010), respectively; age and race/Hispanic origin groups also showed consistently higher RBF values of the same magnitude. Approach #2 (impute

SFOL) produced overall differences of <±1% for each time period and for most age and race/Hispanic origin groups. Approach #3 (impute Hct) generated overall ≤3% lower estimated RBF values for each time period and by age group, but differences were larger for some race/Hispanic origin groups (e.g., ≤7% lower estimated RBF values for non-Hispanic black women). Approach #4 (estimate Hct) produced overall differences of <±1% for each time period and for most age and race/Hispanic origin groups (non-Hispanic black women showed larger differences ≤3%). Finally, approach #6 (predict SFOL and estimate Hct) generated overall differences of ≤±1% for each time period and for most age and race/Hispanic origin groups (non-Hispanic black women showed larger differences ≤3%).

Prevalence of low folate status in women for NHANES 2007–2010 (weighted analysis)

Based on the "true" RBF data, the prevalence of folate deficiency (risk of megaloblastic anemia, RBF <305 nmol/L) was 0.27% (95% CI: 0.12%, 0.57%; data not shown). All RBF approaches produced folate deficiency prevalence estimates of <1% (data not shown). The prevalence of folate insufficiency (risk of NTDs, RBF <748 nmol/L) was 22.5% overall and 20.5%, 38.1%, and 18.2% for Mexican-American, non-Hispanic black, and non-Hispanic white women, respectively (Table 6). When we compared the prevalence of insufficiency across each approach, we found the same patterns as observed with concentrations but in the opposite direction. Approaches #1 and #5 produced prevalences ~5 percentage points lower than "true" RBF and resulted in 100% specificity but lower sensitivity (77.4% and 80.7%, respectively). Approaches #2, #4, and #6 produced prevalences within a few percentage points of "true" RBF and both sensitivity and specificity were ≥ 97%. Approach #3 produced a close prevalence estimate overall, but not by race/Hispanic origin (~5 and 7 percentage points higher for Mexican-American and non-Hispanic black women, respectively), and the sensitivity and specificity were close to 95% each.

Central 95% reference intervals for "true" and estimated RBF in women for NHANES 2007–2010 (weighted analysis)

Compared to the "true" RBF, the central 95% reference interval for RBF by approach #5 was ~5% higher at both ends of the distribution regardless of age or race/Hispanic origin (Supplemental Table 2). For approach #6, the overall reference interval was very similar to that of "true" RBF. For some demographic groups we observed slightly lower or higher reference intervals, but differences were small.

Association of estimated RBF with fasting status and supplement use in women for NHANES 2007–2010 (weighted analysis)

Fasting status was significantly associated with Hct ($P = 0.001$), SFOL ($P = 0.0316$), and RBF ($P = 0.023$). Folic acid–containing supplement use was significantly associated with SFOL ($P < 0.0001$) and RBF ($P < 0.0001$), but not with Hct ($P = 0.46$). For approaches #5 and #6, the differences between the estimated and "true" RBF values were quite similar regardless of the length of the fast or the use of supplements (approach #5: ~5%; approach #6: about −1%) (Supplemental Table 3).

Supplemental analysis in persons aged ≥1 y for NHANES 2007–2010 (weighted analysis)

We expanded the scope of our analysis to assess the same RBF approaches in the overall population aged ≥1 y. The descriptive data for RBF, WBF, SFOL (Supplemental Table 4), and for hematological variables (Supplemental Table 5) were similar to the 2007–2010 data for women (Tables 1, 2). As observed for women, we also saw differences between the predicted and measured SFOL values in persons aged ≥1 y, whereas the estimated and measured Hct values were similar. We observed the same patterns for the 6 RBF approaches as seen for women (Supplemental Table 6). However, there were differences by age group for each approach, with children aged 1–5 and 6–11 y showing larger differences between estimated and "true" RBF values than other age groups. Furthermore, there were also moderate sex and race/Hispanic origin differences between the estimated and "true" RBF values for each approach.

Discussion

This report is to our knowledge the first systematic evaluation that compares several approaches to estimate RBF for population folate status assessment. At a minimum, WBF and Hct or WBF and Hb data are needed to estimate RBF. When we assumed SFOL = 0, RBF results were biased on average ~6% high. When we predicted SFOL from WBF, RBF results produced close agreement with the "true" RBF values. Using either an imputed SFOL or an imputed Hct had disadvantages such as concentration-dependent differences or wide agreement limits. Using an estimated Hct produced close agreement, but still requires a SFOL measurement.

Eliminating the need to measure Hct would simplify field operations. There are 2 strategies to achieve that. One could use an imputed Hct value close to the population median (approach #3), or one could estimate the Hct from Hb because of the strong correlation between these 2 hematologic parameters (approach #4) (8). The use of an imputed Hct does not take into account the wide range of Hct values between individuals [e.g., 21.8–59.1% in NHANES 2015–2016 (16)]. Furthermore, Hct is used as a factor in the RBF calculation formula. These are some of the reasons that cause the wider agreement limits (−20.8%, 11.3%) between the estimated and "true" RBF values, making approach #3 less advisable. The accuracy of the second strategy depends largely on the consistency of the MCHC across the population and on the chosen imputed MCHC value. By using an approximate population median MCHC of 340 g/L, we were within ±1% of the median MCHC value of most population groups. However, given that the MCHC decreases with microcytic (iron deficiency) and increases with macrocytic (folate and/or vitamin B-12 deficiency) anemia (17), each survey should have knowledge of their population-specific MCHC value. Because approach #4 used the measured Hb for each person rather than a constant Hct value (as in approach #3), this approach produced tighter agreement limits (−5.44%, 5.46%). Furthermore, the folate insufficiency prevalence was close to that of "true" RBF, and both the sensitivity and specificity were ≥98%, suggesting that most people who were either folate sufficient or insufficient were correctly identified. Thus, if Hb is but Hct is not available in a nutrition survey, this approach could be useful. However, SFOL measurements are still needed, resulting in only limited resource savings. Furthermore, our findings are

only transferable if the assay used to measure Hb produces data of similar accuracy and variability as in NHANES.

SFOL contributes only a small fraction to WBF in the hemolysate. For example, during 2007–2010, the overall ratio of SFOL to RBF in women was 0.039 and it varied only slightly by age group (range: 0.037–0.044), race/Hispanic origin group (range: 0.037–0.039), or supplement use (range: 0.037–0.042) (data not shown). The overall ratio was similar in persons aged ≥1 y (0.040), but it was higher in children aged 1–5 y (0.054) and 6–11 y (0.055) than in other age groups (range: 0.035–0.044) (data not shown). Given these relatively small differences, it is reasonable to evaluate the possibility of eliminating the measurement of SFOL to save time and resources. This could be accomplished through multiple strategies: assuming SFOL = 0, imputing SFOL, or predicting SFOL. Replacing each individual SFOL with a constant imputed value (approach #2) is not accurate because RBF concentrations are underestimated at the lower end and overestimated at the upper end of the distribution. Assuming SFOL = 0 (approaches #1 and #5) consistently overestimated RBF concentrations by ~6% regardless of the assay or time period, underestimated the prevalence of folate insufficiency by ~5 percentage points, and thus produced some false negative results (sensitivity: ~80%). The simplicity of this approach may hold promise for epidemiologic applications if the cutoff value was raised to obtain the correct prevalence or the RBF data were multiplied by a factor to correct for the overestimation. This approach would save significant time and materials and allow countries in resource-limited environments to generate RBF population data and assess NTD risk. However, if the folate assay used responds differently to serum compared with whole blood, the adjustment factor may be different. Thus, laboratories not using the microbiologic assay may have to determine the correct adjustment factor by analyzing SFOL in a subset of the study population.

Although the predicted SFOL agreed poorly with the measured SFOL and cannot be used to interpret population folate status, using a predicted SFOL and estimated Hct to estimate RBF emerged as an accurate approach (#6) where the RBF data can be used directly for population folate status assessment. This approach produced close estimates to the “true” RBF values across the concentration distribution with tight agreement limits (−6.93%, 7.14%). Data for all 3 time periods and both assays generally agreed within ±1% of the “true” RBF value, the folate insufficiency prevalence was close to that of “true” RBF, and both the sensitivity and specificity were ≥97%. Neither fasting status nor use of folic acid–containing supplements appeared to noticeably affect the estimated RBF. Although SFOL has to be measured in a subset of the study population (~150 samples compared with thousands of survey samples) in order to generate a valid prediction equation (the same assay should be used as for the entire study), this approach provides ~50% savings compared with measuring both SFOL and WBF folate for the entire study. It requires 3 simple computational steps: generate a regression equation from subset analysis, predict SFOL, and estimate Hct for all survey participants.

Major strengths of this study were that we investigated multiple approaches to estimate RBF and we looked at each approach from several angles, including different population groups (women and persons aged ≥1 y), time periods (pre- and postfortification), and

assays (radioassay and microbiologic assay). In addition to evaluating the approaches using unweighted data, we generated nationally representative estimates by demographic subgroup to demonstrate the impact of these approaches for epidemiologic applications. Weaknesses of our study are that other than fasting status and supplement use, we did not explore factors that may influence components of the RBF calculation formula (e.g., pregnancy, obesity, or smoking status). However, we showed that <2% of estimated RBF results deviated by >±10% from the "true" RBF values with approach #6. For pragmatic reasons we used approximate population medians as imputed values instead of the exact values at each time period. Last, adjusting data using statistically derived equations can have undesirable consequences such as an underestimation of the SEs affecting statistical inferences and potentially biased reference intervals or prevalence estimates (18). These potential effects have to be weighed against the savings in time and resources when simplifying the RBF calculation. When resources permit, the conventional approach should be used to calculate the "true" RBF values.

In summary, this comprehensive study using nationally representative data from multiple time periods in NHANES demonstrated the potential to simplify the estimation of RBF in epidemiologic studies. The 2 most promising approaches requiring the least amount of resources were to assume SFOL = 0 and estimate Hct (#5) or to predict SFOL and estimate Hct (#6). The former approach could generate estimated values similar to the "true" RBF values if a correction factor is used to compensate for the overestimation of RBF concentrations; however, the correction factor may be assay-dependent. The latter approach produces estimated values similar to the "true" RBF values, but requires the measurement of SFOL in a subset of the study population to generate the SFOL prediction equation. These approaches require some advance knowledge of the behavior of MCHC and/or SFOL to ascertain whether these results can be replicated in other populations. Although these findings are encouraging, they should be replicated in other populations that may have even lower folate status than the US population during prefortification and/or higher rates of anemia. We expect that the use of the folate microbiologic assay may increase in future population surveys. This is due to an ongoing project of harmonizing this assay (19) and developing laboratory capacity at the regional level for these measurements (20). That is expected to lead to more comparable folate biomarker results.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

We thank Donna J LaVoie, Neelima Paladugula, Bridgette Haynes, and Daniel Rabinowitz (CDC's National Center for Environmental Health) for performing the serum and RBC folate analysis.

This work was performed under employment of the US Federal government and the authors did not receive any outside funding.

Abbreviations used:

Hb	hemoglobin

Hct	hematocrit
MCHC	mean corpuscular hemoglobin content
NTD	neural-tube defect
RBF	RBC folate
SFOL	serum folate
WBF	whole blood folate

References

1. Rogers LR, Cordero AM, Pfeiffer CM, Hausman DB, Tsang BL, De-Regil LM, Rosenthal J, Razzaghi H, Wong EC, Weakland AP,, et al. Global folate status in women of reproductive age: a systematic review with emphasis on methodological issues. Ann N Y Acad Sci 2018;1431:35–57. [PubMed: 30239016]
2. WHO. Guideline: optimal serum and red blood cell folate concentrations in women of reproductive age for prevention of neural tube defects [Internet]. Geneva: World Health Organization; 2015 [cited 4 February, 2019]. Available from: http://apps.who.int/iris/bitstream/10665/161988/1/9789241549042_eng.pdf?ua=1.
3. Chen M-Y, Rose CE, Qi YP, Williams JL, Yeung LF, Berry RJ, Hao L, Cannon MJ, Crider KS. Defining the plasma folate concentration associated with the red blood cell folate concentration threshold for optimal neural tube defects prevention: a population-based, randomized trial of folic acid supplementation. Am J Clin Nutr 2019;109:1452–61. [PubMed: 31005964]
4. Bailey LB, Stover PJ, McNulty H, Fenech MF, Gregory JF III, Mills JL, Pfeiffer CM, Fazili Z, Zhang M, Ueland PM,, et al. Biomarkers of nutrition for development—folate review. J Nutr 2015;145:1636S–80S. [PubMed: 26451605]
5. Molloy AM, Scott JM. Microbiological assay for serum, plasma, and red cell folate using cryopreserved, microtiter plate method. Methods Enzymol 1997;281:43–53. [PubMed: 9250965]
6. Serdula MK, Nichols EK, Aburto NJ, Masa'd H, Obaid B, Wirth J, Tarawneh M, Barham R, Hijawi B, Sullivan KM,, et al. Micronutrient status in Jordan: 2002 and 2010. Eur J Clin Nutr 2014;68:1124–8. [PubMed: 24986824]
7. Kelleher BP, O'Broin SD. High serum folate and the calculation of red cell folate. Clin Lab Haematol 1995;17(2):204–5. [PubMed: 8536429]
8. O'Connor G, Molloy AM, Daly L, Scott JM. Deriving a useful packed cell volume estimate from haemoglobin analysis. J Clin Pathol 1994;47:78–9. [PubMed: 8132815]
9. Centers for Disease Control and Prevention, National Center for Health Statistics. Response rates & CPS population totals, National Health and Nutrition Examination Survey [Internet]. Hyattsville, MD: CDC; 2018 [cited 10 February, 2019]. Available from: https://wwwn.cdc.gov/nchs/nhanes/ResponseRates.aspx.
10. Pfeiffer CM, Hughes JP, Lacher DA, Bailey RL, Berry RJ, Zhang M, Yetley EA, Rader JI, Sempos CT, Johnson CL. Estimation of trends in serum and RBC folate in the U.S. population from pre- to postfortification using assay-adjusted data from the NHANES 1988–2010. J Nutr 2012;142:886–93. [PubMed: 22437563]
11. Raiten DJ, Fisher KD. Assessment of folate methodology used in the third National Health and Nutrition Examination Survey (NHANES III, 1988–1994). J Nutr 1995;125(Suppl):1371S–98S. [PubMed: 7738698]
12. Pfeiffer CM, Hughes JP, Durazo-Arvizu RA, Lacher DA, Sempos CT, Zhang M, Yetley EA, Johnson CL. Changes in measurement procedure from a radioassay to a microbiologic assay necessitates adjustment of serum and RBC folate concentrations in the U.S. population from the NHANES 1988–2010. J Nutr 2012;142:894–900. [PubMed: 22437557]
13. Pfeiffer CM, Sternberg MR, Schleicher RL, Rybak ME. Dietary supplement use and smoking were strongly related to biomarkers of water-soluble vitamin status after adjusting for socioeconomic

and lifestyle factors in a representative sample of US adults. J Nutr 2013;143:957S–65S. [PubMed: 23576641]

14. Haynes BMH, Pfeiffer CM, Sternberg MR, Schleicher RL. Selected pre-analytical and physiological factors are weakly to moderately associated with twenty-nine biomarkers of diet and nutrition, NHANES 2003–2006. J Nutr 2013;143:1001S–10S. [PubMed: 23596168]
15. Pfeiffer CM, Sternberg MR, Hamner HC, Crider KS, Lacher DA, Rogers LM, Bailey RL, Yetley EA. Applying inappropriate cutpoints leads to misinterpretation of folate status in the US population. Am J Clin Nutr 2016;104:1607–15. [PubMed: 27680995]
16. Centers for Disease Control and Prevention, National Center for Health Statistics. National Health and Nutrition Examination Survey. 2015–2016 data documentation, codebook and frequencies. Complete blood count with 5-part differential – whole blood (CDC_I) [Internet]. Hyattsville, MD: CDC; 2017 [cited 22 May, 2019]. Available from: https://wwwn.cdc.gov/Nchs/Nhanes/2015-2016/CBC_I.htm.
17. O'Broin SD, Kelleher BP, Davoren A, Gunter EW. Field-study screening of blood folate concentrations: specimen stability and finger-stick sampling. Am J Clin Nutr 1997;66:1398–405. [PubMed: 9394692]
18. Sternberg M Multiple imputation to evaluate the impact of an assay change in national surveys. Stat Med 2017;36:2697–719. [PubMed: 28419523]
19. Zhang M, Sternberg MR, Pfeiffer CM. Harmonizing the calibrator and microorganism used in the folate microbiologic assay leads to improvements in serum and whole blood folate comparability as shown in a CDC round robin study. J Nutr 2018;148: 807–17. [PubMed: 30053280]
20. Pfeiffer CM, Zhang M, Jabbar S. Framework for laboratory harmonization of folate measurements in low- and middle-income countries and regions. Ann N Y Acad Sci 2018;1414: 96–108. [PubMed: 29377148]

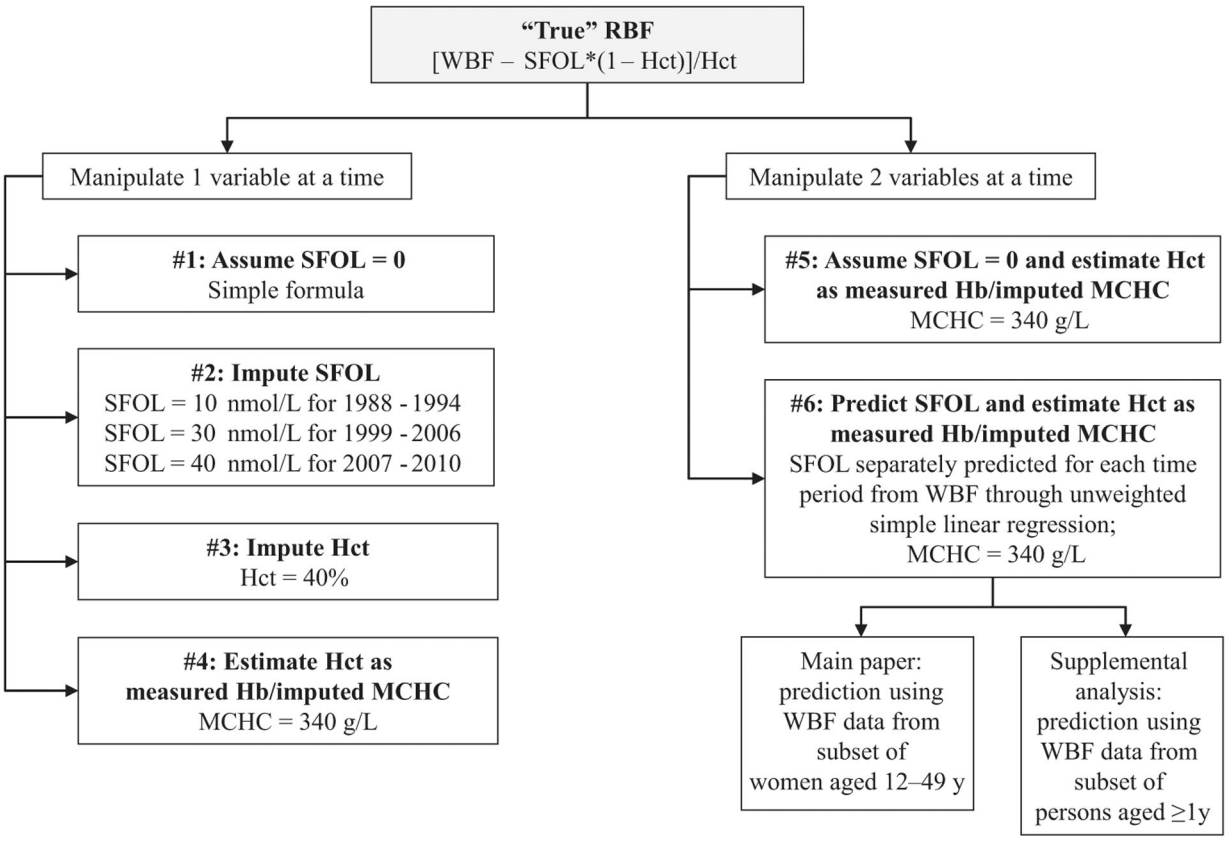

FIGURE 1.

Study design showing the 6 approaches used to estimate RBC folate concentrations. Hb, hemoglobin; Hct, hematocrit; MCHC, mean corpuscular hemoglobin content; RBF, RBC folate; SFOL, serum folate; WBF, whole blood folate.

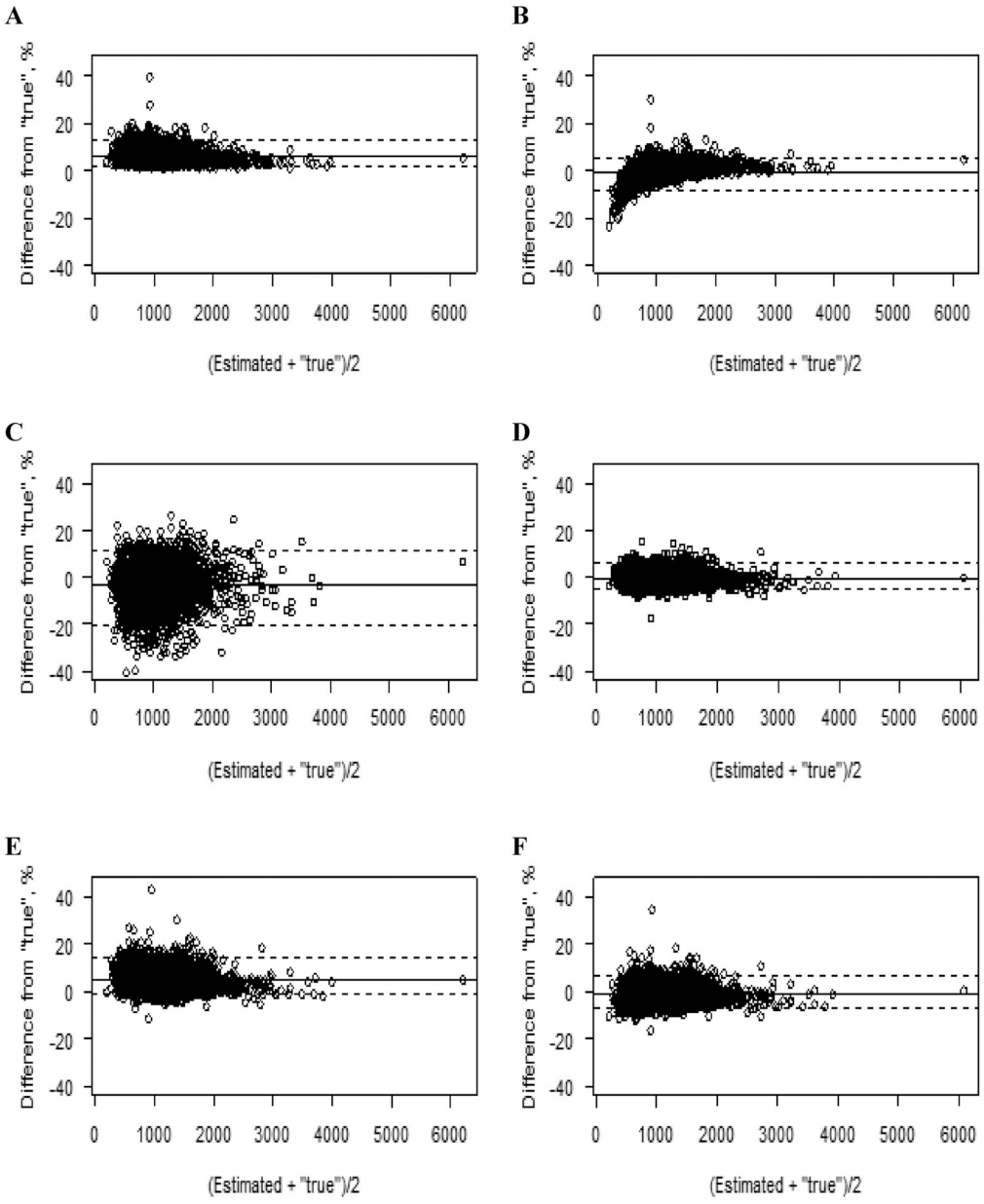

FIGURE 2.

Bland–Altman plot showing the unweighted relative difference between estimated and "true" RBC folate values in women aged 12–49 y using 6 different approaches, NHANES 2007–2010. The solid line represents the median difference, whereas the dashed lines represent the nonparametric "agreement limits" capturing 95% of the data, $n = 3951$. (A) Assume SFOL = 0; (B) impute SFOL; (C) impute Hct; (D) estimate Hct; (E) assume SFOL = 0 and estimate Hct; (F) predict SFOL and estimate Hct. Hct, hematocrit; SFOL, serum folate.

TABLE 1

Folate concentrations by demographic group, women aged 12–49 y, NHANES 1988–2010[1]

Survey group	Sample size, *n*	RBF,[2] nmol/L	WBF, nmol/L	SFOL, nmol/L	SFOL_predicted,[3] nmol/L
1988–1994					
Overall	6275	363 (349, 381)	150 (143, 157)	11.2 (10.5, 11.8)	13.2 (12.7, 13.8)
Age, y					
12–19	1548	327 (307, 348)	136 (128, 146)	11.4 (10.7, 12.4)	12.2 (11.6, 13.0)
20–39	3445	362 (346, 377)	149 (143, 156)	11.0 (10.2, 11.8)	13.2 (12.7, 13.7)
40–49	1282	403 (382, 425)	163 (155, 175)	11.3 (10.2, 12.5)	14.3 (13.5, 15.2)
Race/Hispanic origin					
MA	1961	354 (333, 378)	143 (139, 153)	10.3 (9.64, 11.0)	12.7 (12.4, 13.5)
NHB	2121	282 (276, 290)	115 (110, 118)	9.19 (8.89,9.40)	10.5 (10.1, 10.7)
NHW	1894	386 (361,403)	160 (150, 170)	11.8 (10.9, 12.7)	14.0 (13.2, 14.8)
1999–2006					
Overall	9236	580 (569, 594)	249 (244, 255)	26.6 (26.0, 27.2)	28.1 (27.6, 28.5)
Age, y					
12–19	4130	534 (523, 544)	229 (224, 235)	27.4 (26.5,28.1)	26.3 (25.8, 26.7)
20–39	3589	576 (561, 594)	247 (241, 254)	25.9 (25.2, 26.7)	27.8 (27.3, 28.5)
40–49	1517	631 (615, 646)	269 (261, 279)	27.4 (26.4, 28.4)	29.9 (29.0, 30.8)
Race/Hispanic origin					
MA	2750	559 (546, 568)	239 (232, 244)	24.7 (23.9, 25.4)	27.1 (26.5, 27.6)
NHB	2365	470 (458, 480)	194 (189, 195)	22.1 (21.3,23.1)	23.1 (22.6, 23.2)
NHW	3304	615 (598, 630)	266 (259, 274)	28.4 (27.2, 29.4)	29.4 (28.9, 30.3)
2007–2010					
Overall	3951	1010 (986, 1030)	413 (403, 423)	37.1 (36.0, 38.4)	38.4 (37.7, 39.1)
Age, y					
12–19	1042	933 (896, 976)	387 (369, 401)	40.5 (37.8, 42.7)	36.8 (35.4, 37.7)
20–39	1894	1000 (978, 1040)	413 (402, 423)	36.2 (34.9, 37.6)	38.3 (37.5, 39.0)
40–49	1015	1080 (1020, 1150)	447 (419, 473)	37.5 (35.1,39.4)	40.6 (38.7, 42.4)
Race/Hispanic origin					

Am J Clin Nutr. Author manuscript; available in PMC 2023 March 29.

Survey group	Sample size, *n*	RBF,[2] nmol/L	WBF, nmol/L	SFOL, nmol/L	SFOL_predicted,[3] nmol/L
MA	867	980 (926, 1020)	398 (380, 414)	35.4 (34.0, 36.5)	37.4 (36.2, 38.4)
NHB	785	831 (799, 868)	329 (319, 343)	29.5 (28.2, 30.7)	32.7 (32.0, 33.5)
NHW	1575	1070 (1030, 1110)	446 (428, 465)	39.7 (37.8, 42.4)	40.6 (39.3, 42.0)

[1] Values are sample sizes and weighted median (95% CI) folate concentrations measured by the BioRad Quantaphase II radioassay (1988–1994 and 1999–2006) or microbiologic assay (2007–2010). Hct, hematocrit; MA, Mexican American; NHB, non-Hispanic black; NHW, non-Hispanic white; RBF, RBC folate; SFOL, serum folate; WBF, whole blood folate.

[2] Calculated using measured WBF, SFOL, and Hct: RBF = [WBF – SFOL*(1 Hct)]/Hct.

[3] SFOL predicted using unweighted simple linear regression derived from measured WBF in a simple random sample of women aged 12–49 y ($n = 150$) as the x-variable and measured SFOL as the y-variable from NHANES 1988–1994 (predicted SFOL = 1.35023 + 0.07943*WBF), 1999–2006 (predicted SFOL = 5.52173 + 0.09046*WBF), and 2007–2010 (predicted SFOL = 10.59233 + 0.06712*WBF), respectively.

TABLE 2

Hematologic variables by demographic group, women aged 12–49 y, NHANES 1988–2010[1]

Group	Hct, %	Hb, g/L	MCHC, g/L	Hct_estimated,[2] %
1988–1994				
Overall	39.3 (39.0, 39.6)	132 (131, 133)	337 (336, 338)	38.9 (38.6,39.1)
Age, y				
12–19	39.3 (39.0, 39.7)	132 (131, 133)	336 (335, 338)	38.8 (38.4,39.1)
20–39	39.2 (39.0, 39.5)	132 (131, 133)	337 (336, 338)	38.9 (38.6, 39.2)
40–49	39.3 (39.0, 39.8)	132 (131, 134)	337 (336, 338)	38.9 (38.6, 39.3)
Race/Hispanic origin				
MA	39.0 (38.7, 39.4)	132 (130, 133)	336 (335, 338)	38.7 (38.3, 39.0)
NHB	37.6 (37.4, 37.8)	125 (124, 125)	331 (330,332)	36.6 (36.4, 36.8)
NHW	39.6 (39.4, 39.8)	134 (133, 135)	338 (337, 339)	39.3 (39.0, 39.6)
1999–2006				
Overall	39.8 (39.6, 40.1)	136 (135, 136)	339 (338, 340)	39.9 (39.7,40.1)
Age, y				
12–19	39.8 (39.6, 40.1)	136 (135, 136)	340 (339, 341)	39.9 (39.6, 40.1)
20–39	39.7 (39.4, 40.0)	135 (134, 136)	339 (339, 340)	39.7 (39.5, 40.0)
40–49	40.1 (39.7, 40.6)	137 (135, 138)	339 (338, 340)	40.2 (39.8, 40.5)
Race/Hispanic origin				
MA	39.5 (39.2, 39.9)	134 (134, 135)	339 (338, 340)	39.5 (39.3, 39.8)
NHB	38.0 (37.8, 38.2)	127 (126, 128)	333 (332, 335)	37.3 (37.0, 37.5)
NHW	40.2 (40.0, 40.6)	137 (137, 138)	341 (340,341)	40.4 (40.2, 40.6)
2007–2010				
Overall	38.9 (38.6, 39.1)	134 (133, 135)	343 (342, 345)	39.4 (39.1,39.6)
Age, y				
12–19	38.6 (38.3, 38.9)	133 (132, 134)	343 (341, 345)	39.2 (38.9, 39.5)
20–39	38.9 (38.5, 39.2)	134 (133, 135)	344 (342, 345)	39.4 (39.1,39.7)
40–49	39.0 (38.7, 39.3)	134 (133, 135)	343 (341, 345)	39.4 (39.1,39.7)
Race/Hispanic origin				
MA	38.1 (37.8, 38.6)	131 (130, 133)	343 (341,345)	38.6 (38.3, 39.0)

Group	Hct, %	Hb, g/L	MCHC, g/L	Hct_estimated,[2] %
NHB	37.2 (36.9, 37.5)	126 (125, 126)	338 (336, 340)	36.9 (36.7, 37.2)
NHW	39.3 (39.0, 39.6)	136 (135, 136)	345 (343, 347)	39.9 (39.6, 40.1)

[1]Values are weighted medians (95% CIs); for sample sizes, see Table 1. Hb, hemoglobin; Hct, hematocrit; MA, Mexican American; MCHC, mean corpuscular hemoglobin content; NHB, non-Hispanic black; NHW, non-Hispanic white.

[2]Calculated as measured Hb/imputed MCHC (340 g/L for NHANES 1988–2010).

TABLE 3

Nonparametric unweighted agreement limits and median relative difference between estimated and "true" RBF using 6 different approaches, women aged 12–49 y, NHANES 2007–2010[1]

	RBF estimation approach					
Percentile	**#1: Assume SFOL = 0[2]**	**#2: Impute SFOL[3]**	**#3: Impute Hct[4]**	**#4: Estimate Hct[5]**	**#5: Assume SFOL = 0 and estimate Hct[6]**	**#6: Predict SFOL and estimate Hct[7]**
2.5	2.33	− 8.62	− 20.8	− 5.44	− 1.19	− 6.93
5	2.76	− 7.10	− 17.0	− 4.84	− 0.34	− 6.04
50	5.78	− 0.74	− 3.36	− 0.96	4.96	− 1.15
95	11.1	4.24	8.97	4.04	12.8	5.47
97.5	13.0	5.49	11.3	5.46	14.7	7.14

[1] Values are percentage differences between each RBF estimation approach and the "true" RBF value as measured by microbiologic assay (2007–2010), n = 3951. Hb, hemoglobin; Hct, hematocrit; MCHC, mean corpuscular hemoglobin content; RBF, RBC folate; SFOL, serum folate; WBF, whole blood folate.

[2] RBF estimated using measured WBF and Hct without correcting for measured SFOL: RBF = WBF/Hct.

[3] RBF estimated using measured WBF and Hct and an imputed SFOL of 40 nmol/L (2007–2010): RBF = [WBF − imputed SFOL*(1 − Hct)]/Hct.

[4] RBF estimated using measured WBF and SFOL and an imputed Hct of 40%: RBF = [WBF – SFOL*(1 − 0.4)]/0.4.

[5] RBF estimated using measured WBF and SFOL and an estimated Hct: RBF = [WBF – SFOL*(1 − estimated Hct)]/estimated Hct; estimated Hct = measured Hb/imputed MCHC MCHC 340 g/L.

[6] RBF estimated using measured WBF and an estimated Hct without correcting for measured SFOL: RBF = WBF/estimated Hct; estimated Hct = measured Hb/imputed MCHC using MCHC 340 g/L.

[7] RBF estimated using measured WBF, predicted SFOL, and estimated Hct: RBF = [WBF − predicted SFOL*(1 − estimated Hct)]/estimated Hct; SFOL predicted using unweighted simple linear regression derived from measured WBF in a simple random sample of women aged 12–49 y (n = 150) as the x-variable and measured SFOL as the y-variable from NHANES 2007–2010 (predicted SFOL = 10.59233+0.06712*WBF); estimated Hct measured Hb/imputed MCHC using MCHC = 340 g/L.

Am J Clin Nutr. Author manuscript; available in PMC 2023 March 29.

TABLE 4

"True" and estimated RBF concentrations by demographic group using different approaches, women aged 12–49 y, NHANES 1988–2010[1]

		RBF estimation approach					
Group	**"True" RBF**	**#1: Assume SFOL = 0[2]**	**#2: Impute SFOL[3]**	**#3: Impute Hct[4]**	**#4: Estimate Hct[5]**	**#5: Assume SFOL = 0 and estimate Hct[6]**	**#6: Predict SFOL and estimate Hct[7]**
1988–1994							
Overall	373 (360, 386)	393 (379, 407)	375 (362, 389)	364 (351, 378)	376 (364, 389)	397 (384,411)	375 (363, 388)
Age, y							
12–19	334 (317, 353)	354 (336, 374)	337 (319, 357)	328 (311, 346)	338 (321, 356)	359 (340, 378)	339 (321, 357)
20–39	372 (360, 384)	392 (379, 405)	374 (362, 388)	363 (351, 376)	376 (364, 388)	396 (383, 409)	374 (362, 387)
40–49	407 (384, 431)	428 (403, 454)	410 (386, 436)	399 (376, 423)	411 (388, 436)	433 (408, 459)	409 (386, 434)
Race/Hispanic origin							
MA	365 (346, 384)	383 (364, 404)	366 (347, 386)	353 (335, 372)	369 (351, 388)	388 (370, 408)	367 (349, 386)
NHB	287 (279, 296)	305 (297, 314)	286 (278, 295)	269 (262, 277)	295 (287, 304)	314 (305, 323)	295 (287, 304)
NHW	391 (375, 407)	411 (395, 429)	394 (378,412)	386 (370, 402)	393 (378, 410)	414 (398, 432)	392 (376, 409)
1999–2006							
Overall	588 (577, 599)	632 (620, 644)	583 (571,595)	583 (572, 594)	589 (578, 600)	633 (621, 645)	590 (579, 601)
Age, y							
12–19	538 (527, 549)	582 (570, 594)	533 (522, 545)	534 (522, 546)	538 (527, 549)	582 (571, 594)	541 (530,553)
20–39	586 (573, 600)	630 (615, 644)	580 (566, 595)	580 (566, 593)	587 (574, 600)	630 (616, 645)	588 (574, 601)
40–49	626 (610, 643)	671 (654, 689)	622 (605, 640)	624 (608, 641)	628 (612, 645)	673 (656, 691)	629 (612, 646)
Race/Hispanic origin							
MA	575 (564, 586)	616 (605, 627)	567 (556, 578)	566 (555, 578)	576 (565, 587)	617 (606, 628)	575 (564, 586)
NHB	475 (465, 486)	516 (504, 527)	463 (451, 474)	449 (439, 459)	484 (474, 495)	526 (514, 538)	484 (473, 496)
NHW	620 (604, 636)	665 (649, 682)	618 (601,635)	622 (607, 638)	618 (603, 634)	664 (647, 681)	620 (604, 637)
2007–2010							
Overall	1010 (986, 1030)	1070 (1050, 1100)	1000 (977, 1030)	975 (951, 1000)	1000 (977, 1020)	1060 (1040, 1080)	999 (976, 1020)
Age, y							
12–19	937 (907, 968)	1000 (971, 1030)	934 (903, 966)	905 (875, 936)	927 (898, 956)	990 (960, 1020)	932 (903, 961)
20–39	1010 (982, 1040)	1070 (1040, 1100)	1000 (972,1030)	974 (949, 1000)	998 (972, 1020)	1060 (1030, 1080)	996 (969, 1020)
40–49	1060 (1010, 1110)	1130 (1070, 1180)	1060 (1000, 1110)	1030 (980, 1080)	1050 (1010, 1100)	1120 (1070, 1170)	1050 (1010, 1100)

Group	"True" RBF	RBF estimation approach					
		#1: Assume SFOL = 0[2]	#2: Impute SFOL[3]	#3: Impute Hct[4]	#4: Estimate Hct[5]	#5: Assume SFOL = 0 and estimate Hct[6]	#6: Predict SFOL and estimate Hct[7]
Race/Hispanic origin							
MA	979 (935, 1030)	1040 (994,1090)	970 (925, 1020)	933 (893, 974)	971 (925, 1020)	1030 (983, 1080)	969 (924, 1020)
NHB	853 (823, 884)	907 (875, 940)	833 (801, 866)	789 (761, 817)	858 (829, 888)	912 (882, 944)	854 (825, 884)
NHW	1070 (1030, 1100)	1130 (1090, 1170)	1060 (1020, 1100)	1040 (1010, 1080)	1050 (1010, 1090)	1110 (1080, 1150)	1050 (1010, 1090)

[1] Values are weighted geometric means (95% CIs) (nmol/L) for RBF measured by the BioRad Quantaphase II radioassay (1988–1994 and 1999–2006) or microbiologic assay (2007–2010); for sample sizes, see Table 1. Hb, hemoglobin; Hct, hematocrit; MA, Mexican American; MCHC, mean corpuscular hemoglobin content; NHB, non-Hispanic black; NHW, non-Hispanic white; RBF, RBC folate; SFOL, serum folate; WBF, whole blood folate.

[2] RBF estimated using measured WBF and Hct without correcting for measured SFOL: RBF = WBF/Hct.

[3] RBF estimated using measured WBF and Hct and an imputed SFOL of 10 (1988–1994), 30 (1999–2006), or 40 nmol/L (2007–2010): RBF = [WBF − imputed SFOL*(1 − Hct)]/Hct.

[4] RBF estimated using measured WBF and SFOL and an imputed Hct of 40%: RBF = [WBF − SFOL*(1 − 0.4)]/0.4.

[5] RBF estimated using measured WBF and SFOL and an estimated Hct: RBF = [WBF − SFOL*(1 − estimated Hct)]/estimated Hct; estimated Hct = measured Hb/imputed MCHC using MCHC = 340 g/L.

[6] RBF estimated using measured WBF and an estimated Hct without correcting for measured SFOL: RBF = WBF/estimated Hct; estimated Hct = measured Hb/imputed MCHC using MCHC = 340 g/L.

[7] RBF estimated using measured WBF, predicted SFOL, and estimated Hct: RBF = [WBF − predicted SFOL*(1 − estimated Hct)]/estimated Hct; SFOL predicted using unweighted simple linear regression derived from measured WBF in a simple random sample of women aged 12–49 y (n = 150) as the x-variable and measured SFOL as the y-variable from NHANES 1988–1994 (predicted SFOL = 1.35023 + 0.07943*WBF), 1999–2006 (predicted SFOL = 5.52173 + 0.09046*WBF), and 2007–2010 (predicted SFOL = 10.59233 + 0.06712*WBF), respectively; estimated Hct = measured Hb/imputed MCHC using MCHC = 340 g/L.

TABLE 5

Relative difference between estimated and "true" RBF by demographic group using different approaches, women aged 12–49 y, NHANES 1988–2010[1]

	RBF estimation approach					
Group	**#1: Assume SFOL = 0[2]**	**#2: Impute SFOL[3]**	**#3: Impute Hct[4]**	**#4: Estimate Hct[5]**	**#5: Assume SFOL = 0 and estimate Hct[6]**	**#6: Predict SFOL and estimate Hct[7]**
1988–1994						
Overall	5.5 (5.3, 5.6)	0.80 (0.56, 1.0)	−2.0 (−2.4, −1.5)	1.1 (0.81, 1.3)	6.6 (6.4, 6.9)	0.81 (0.57, 1.0)
Age, y						
12–19	6.0 (5.7, 6.3)	0.98 (0.58, 1.4)	− 1.8 (−2.5, −1.0)	1.2 (0.83, 1.5)	7.3 (6.9, 7.7)	1.4 (1.0, 1.7)
20–39	5.4 (5.2, 5.6)	0.72 (0.47, 0.96)	−2.1 (−2.6, −1.6)	1.0 (0.74, 1.3)	6.5 (6.2, 6.8)	0.67 (0.37, 0.97)
40–49	5.2 (4.9, 5.5)	0.88 (0.48, 1.3)	− 1.7 (−2.4, −0.94)	1.2 (0.82, 1.5)	6.4 (6.0, 6.8)	0.70 (0.35, 1.0)
Race/Hispanic origin						
MA	5.2 (5.0, 5.3)	0.45 (0.18,0.72)	−2.8 (−3.6, −2.1)	1.3 (0.80, 1.8)	6.6 (6.0, 7.1)	0.70 (0.18, 1.2)
NHB	6.2 (5.9, 6.6)	−0.18 (−0.48,0.11)	− 5.9 (−6.3, −5.5)	2.8 (2.4, 3.1)	9.3 (8.8, 9.8)	2.9 (2.4, 3.3)
NHW	5.4 (5.1, 5.6)	1.0 (0.69, 1.4)	− 1.0 (−1.6, −0.46)	0.70 (0.41,0.99)	6.1 (5.9, 6.4)	0.41 (0.16, 0.66)
1999–2006						
Overall	7.6 (7.4, 7.7)	−0.76 (−0.94, −0.58)	−0.51 (−1.0, −0.02)	0.19 (−0.04, 0.42)	7.8 (7.5, 8.0)	0.36 (0.12, 0.60)
Age, y						
12–19	8.2 (8.1, 8.4)	− 0.70 (−0.92, −0.47)	−0.52 (−1.1, 0.03)	0.07 (−0.13, 0.28)	8.3 (8.1, 8.6)	0.77 (0.53, 1.0)
20–39	7.5 (7.3, 7.7)	−0.94 (−1.2, −0.70)	−0.80 (−1.3, −0.31)	0.15 (−0.10, 0.39)	7.6 (7.3, 7.9)	0.21 (−0.07,0.48)
40–49	7.3 (7.0, 7.5)	− 0.50 (−0.79, −0.20)	0.01 (−0.74, 0.76)	0.35 (0.08, 0.61)	7.7 (7.2, 8.1)	0.36 (−0.02, 0.73)
Race/Hispanic origin						
MA	7.2 (7.0, 7.4)	− 1.3 (−1.5, −1.1)	− 1.2 (−2.0, −0.42)	0.22 (−0.08, 0.53)	7.4 (7.0, 7.8)	−0.02 (−0.39, 0.35)
NHB	8.5 (8.2, 8.8)	− 2.6 (−2.9, −2.2)	−5.1 (−5.6, −4.7)	1.9 (1.6, 2.2)	11 (10, 11)	2.3 (1.9, 2.7)
NHW	7.4 (7.3, 7.6)	− 0.26 (−0.50, −0.02)	0.62 (0.04, 1.2)	−0.20 (−0.43, 0.02)	7.2 (6.9, 7.4)	− 0.00 (−0.24, 0.24)
2007–2010						
Overall	6.2 (6.0, 6.4)	− 0.66 (−0.86, −0.46)	−3.1 (−3.7, −2.5)	−0.97 (−1.4, 0.58)	5.1 (4.6, 5.6)	−0.98 (−1.4, −0.56)
Age, y						
12–19	7.0 (6.6, 7.3)	− 0.29 (−0.62, 0.04)	−3.2 (−4.1, −2.2)	− 1.1 (−1.5, −0.64)	5.7 (5.2, 6.2)	−0.53 (−0.97, −0.09)
20–39	6.0 (5.8, 6.2)	−0.83 (−1.1, −0.59)	−3.1 (−3.8, −2.5)	− 1.1 (−1.5, −0.71)	4.8 (4.3, 5.3)	− 1.2 (−1.7, −0.77)
40–49	6.0 (5.7, 6.2)	−0.60 (−1.0, −0.17)	−2.9 (−3.7, −2.1)	−0.71 (−1.2, −0.22)	5.2 (4.6, 5.7)	−0.81 (−1.3, −0.34)

Group	RBF estimation approach					
	#1: Assume SFOL = 0[2]	#2: Impute SFOL[3]	#3: Impute Hct[4]	#4: Estimate Hct[5]	#5: Assume SFOL = 0 and estimate Hct[6]	#6: Predict SFOL and estimate Hct[7]
Race/Hispanic origin						
MA	6.2 (5.9, 6.4)	−0.91 (−1.2, −0.65)	− 4.4 (−5.2, −3.6)	−0.87 (−1.2, −0.49)	5.2 (4.8, 5.6)	−1.0 (−1.4, −0.62)
NHB	6.3 (6.1, 6.5)	−2.3 (−2.6, −1.9)	−7.1 (−7.7, −6.6)	0.62 (−0.03, 1.28)	7.0 (6.2, 7.8)	0.18 (−0.55, 0.91)
NHW	6.1 (5.8, 6.3)	− 0.24 (−0.56, −0.08)	− 1.8 (−2.6, −1.0)	− 1.4 (−1.8, −0.99)	4.5 (4.0, 5.0)	− 1.3 (−1.8, −0.86)

[1]Values represent weighted mean percentage differences (95% CIs) between each RBF estimation approach and the "true" RBF value measured by the BioRad Quantaphase II radioassay (1988–1994 and 1999–2006) or microbiologic assay (2007–2010); for sample sizes, see Table 1. Hb, hemoglobin; Hct, hematocrit; MA, Mexican American; MCHC, mean corpuscular hemoglobin content; NHB, non-Hispanic black; NHW, non-Hispanic white; RBF, RBC folate; SFOL, serum folate; WBF, whole blood folate.

[2]RBF estimated using measured WBF and Hct without correcting for measured SFOL: RBF = WBF/Hct.

[3]RBF estimated using measured WBF and Hct and an imputed SFOL of 10 (1988–1994), 30 (1999–2006), or 40 nmol/L (2007–2010): RBF = [WBF − imputed SFOL*(1 − Hct)]/Hct.

[4]RBF estimated using measured WBF and SFOL and an imputed Hct of 40%: RBF = [WBF – SFOL*(1 − 0.4)]/0.4.

[5]RBF estimated using measured WBF and SFOL and an estimated Hct: RBF = [WBF – SFOL*(1 − estimated Hct)]/estimated Hct; estimated Hct = measured Hb/imputed MCHC using MCHC = 340 g/L.

[6]RBF estimated using measured WBF and an estimated Hct without correcting for measured SFOL: RBF = WBF/estimated Hct; estimated Hct = measured Hb/imputed MCHC using MCHC = 340 g/L.

[7]RBF estimated using measured WBF, predicted SFOL, and estimated Hct: RBF = [WBF − predicted SFOL*(1 − estimated Hct)]/estimated Hct; SFOL predicted using unweighted simple linear regression derived from measured WBF in a simple random sample of women aged 12–49 y (n = 150) as the x-variable and measured SFOL as the y-variable from NHANES 1988–1994 (predicted SFOL = 1.35023+0.07943*WBF), 1999–2006 (predicted SFOL = 5.52173+0.09046*WBF), and 2007–2010 (predicted SFOL = 10.59233+0.06712*WBF), respectively; estimated Hct = measured Hb/imputed MCHC using MCHC = 340 g/L.

TABLE 6

Prevalence of insufficient folate status for RBF (<748 nmol/L) using "true" and estimated RBF data by race/Hispanic origin and diagnostic sensitivity and specificity, women aged 12–49 y, NHANES 2007–2010[1]

		RBF approximation approach					
Race/Hispanic origin	**"True" RBF**	**#1: Assume SFOL = 0[2]**	**#2: Impute SFOL[3]**	**#3: Impute Hct[4]**	**#4: Estimate Hct[5]**	**#5: Assume SFOL = 0 and estimate Hct[6]**	**#6: Predict SFOL and estimate Hct[7]**
Prevalence, %							
Overall	22.5 (20.5, 24.6)	17.4 (15.6, 19.4)	24.0 (21.8,26.2)	25.1 (23.1,27.2)	23.0 (21.1,25.1)	18.2 (16.5, 20.0)	23.4 (21.4,25.6)
MA	20.5 (16.8, 24.7)	17.0 (13.6,21.0)	23.1 (19.2, 27.6)	25.1 (21.5, 29.2)	21.7 (18.3, 25.6)	17.5 (14.1,21.5)	22.7 (18.8, 27.2)
NHB	38.1 (35.1,41.2)	30.9 (27.4, 34.6)	40.6 (37.5, 43.8)	45.4 (41.6, 49.2)	36.9 (34.0, 40.0)	30.8 (27.8, 34.1)	37.4 (34.4, 40.4)
NHW	18.2 (15.8,21.0)	13.7 (11.3, 16.4)	19.5 (16.9, 22.5)	19.5 (16.8, 22.4)	19.2 (16.7, 22.0)	14.6 (12.4, 17.1)	19.4 (16.8, 22.3)
Sensitivity/specificity							
Overall	n/a	77.4/100	99.1/97.8	93.7/94.8	98.1/98.7	80.7/100	96.9/97.9

[1] Values represent weighted percentage prevalences (95% CIs); for sample sizes, see Table 1. Hct, hematocrit; MA, Mexican American; MCHC, mean corpuscular hemoglobin content; NHB, non-Hispanic black; NHW, non-Hispanic white; RBF, RBC folate; SFOL, serum folate; WBF, whole blood folate.

[2] RBF estimated using measured WBF and Hct without correcting for measured SFOL: RBF = WBF/Hct.

[3] RBF estimated using measured WBF and Hct and an imputed SFOL of 40 nmol/L: RBF = [WBF – 40*(1 − Hct)]/Hct.

[4] RBF estimated using measured WBF and SFOL and an imputed Hct of 40%: RBF = [WBF – SFOL*(1 − 0.4)]/0.4.

[5] RBF estimated using measured WBF and SFOL and an estimated Hct: RBF = [WBF – SFOL*(1 − estimated Hct)]/estimated Hct; estimated Hct = measured Hb/imputed MCHC using MCHC 340 g/L.

[6] RBF estimated using measured WBF and an estimated Hct without correcting for measured SFOL: RBF = WBF/estimated Hct; estimated Hct = measured Hb/imputed MCHC using MCHC = 340 g/L.

[7] RBF estimated using measured WBF, predicted SFOL, and estimated Hct: RBF = [WBF − predicted SFOL*(1 − estimated Hct)]/estimated Hct; SFOL predicted using unweighted simple linear regression derived from measured WBF in a simple random sample of women aged 12–49 y (n = 150) as the x-variable and measured SFOL as the y-variable from NHANES 2007–2010 (predicted SFOL = 10.59233+0.06712*WBF); estimated Hct = measured Hb/imputed MCHC using MCHC = 340 g/L.