ORIGINAL RESEARCH

CURRENT DEVELOPMENTS IN NUTRITION

Research Methodology and Study Design

Regression to the Mean: A Statistical Phenomenon of Worthy Consideration in Anemia Research

Kelsey M Cochrane,[1,2] Brock A Williams,[1,2] Jordie AJ Fischer,[1,2] Kaitlyn LI Samson,[1,2] Lulu X Pei,[1,3] and Crystal D Karakochuk[1,2]

[1]Food, Nutrition, and Health, Faculty of Land and Food Systems, The University of British Columbia, Vancouver, British Columbia, Canada; [2]British Columbia Children's Hospital Research Institute, Healthy Starts, Vancouver, British Columbia, Canada; and [3]Department of Statistics, Faculty of Science, The University of British Columbia, Vancouver, British Columbia, Canada

ABSTRACT

Background: Regression to the mean (RTM) is a statistical phenomenon where second measurements are more likely to be closer to the mean. This is particularly observed in those with baseline values further from the mean. Anemic individuals (hemoglobin <120 g/L) are often recruited when evaluating iron supplementation programs, as they are more likely to elicit a greater hemoglobin response; however, they are also at greater risk for RTM as their baseline values are lower than the overall population mean.

Objective: The aim was to calculate and apply RTM to a previously conducted iron supplementation trial of women in Cambodia at increasingly severe baseline anemia cutoffs (hemoglobin <120 g/L, <115 g/L, and <110 g/L).

Methods: Women received either 60 mg/d iron ($n = 191$) or placebo ($n = 185$) for 12 wk. Hemoglobin was measured at baseline and at 12 wk (endline), and change in hemoglobin was calculated in each group for each cutoff. RTM was calculated in the placebo group at each cutoff and applied to the change observed at each cutoff in the iron group to obtain the RTM-free effect.

Results: In the placebo group, mean change in hemoglobin increased as cutoffs became more extreme (0.9 g/L to 1.9 g/L in those with baseline hemoglobin <120 g/L and <110 g/L, respectively). RTM estimates similarly increased: 1.0 g/L (<120 g/L), 1.3 g/L (<115 g/L), and 1.8 g/L (<110g/L). When applying RTM to the iron group, we found that ~10% of the "treatment effect" could be attributable to RTM at each cutoff. However, iron supplementation was still effective in increasing hemoglobin, with an increased effect in those with lower baseline values, as proven by the RTM-free effect at each cutoff: 8.7 g/L (<120 g/L), 10.9 g/L (<115 g/L), and 13.6g/L (<110 g/L).

Conclusions: RTM may have accounted for ~10% of the observed change in hemoglobin following iron supplementation; however, appropriate use of a placebo group in the statistical analyses of the trial controls for this potential RTM effect. *Curr Dev Nutr* 2020;4:nzaa152.

Keywords: regression to the mean, anemia, hemoglobin, iron supplementation, nonpregnant women, Cambodia

Manuscript received July 19, 2020. Initial review completed September 9, 2020. Revision accepted September 22, 2020. Published online September 24, 2020.
The authors reported no funding received for this work.
Author disclosures: The authors report no conflicts of interest.
Supplemental Figure 1 is 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/cdn/.
Address correspondence to CDK (e-mail: crystal.karakochuk@ubc.ca).

Introduction

Regression to the mean (RTM) is a statistical phenomenon that was first described in 1877 (1). The effect of RTM is highly probable when repeated or pre- and posttest measures are taken in individuals above or below a relevant clinical cutoff or set deviation from the population mean (2). When there is a lack of perfect correlation between 2 measures due to biological or measurement variability, the second measure is more likely to be closer to the mean (3, 4). This RTM effect is greater when baseline levels are at the more extreme (high or low) ends of the distribution or the further that individuals deviate from the mean (5). In the absence of an appropriate placebo group (which accounts for the effect of RTM in between-group statistical analyses), this can lead researchers to misinterpret treatment or intervention effects when a significant portion of, or the entire effect, may be due to RTM (5–7). Additionally, to accurately estimate within-group treatment effects (e.g., RTM-free effects), RTM is calculated in the placebo group and then applied to the observed change in the intervention group. Misinterpretations due to RTM can have implications for the application of research findings to clinical practice or public policy development.

While this phenomenon has been explored and analyzed in various disease states, such as osteoporosis (8), hypertension (9, 10), hypercholesteremia (11, 12), and obesity (3), it has not been widely studied in other areas of nutrition research, such as nutritional hematology or anemia. There are numerous examples of anemia and iron supplementation clinical trials that have been conducted without the use of a

placebo group and/or which recruit only anemic individuals, increasing the possibility for misinterpretation of results due to an RTM effect (13–17). Overall, there is limited awareness among public health officials, nutrition researchers, and clinicians of the importance of this concept and the possible effect on anemia intervention results. Misinterpretation of intervention effects due to the presence of RTM is prevalent in nutrition research, as clinical trials are often done in those who are most "at risk" (with baseline values at the extreme ends of a distribution) (2).

Hemoglobin concentrations are routinely measured as diagnostic criteria for anemia (defined as hemoglobin <120 g/L in nonpregnant women) (18). Anemia rates are commonly measured and reported as a proxy of the nutritional status of a population, such as in national Demographic and Health Surveys and by the WHO. Further, recommendations for country- or population-specific iron supplementation programs are largely based on the prevalence of anemic individuals (19), whose hemoglobin concentrations may be most responsive to iron supplementation, but who are also at higher risk of RTM. Additionally, it is commonly believed that the effect of iron may be modified by baseline iron and hemoglobin status (20); the extent to which RTM contributes to this phenomenon is unknown. In studies of this nature, the calculation of an RTM-free effect can help illustrate the true intervention effect.

The aims of the current study were as follows: *1*) to apply the concept of RTM to a previously conducted iron supplementation trial among nonpregnant women in Cambodia (21) and quantify the effect of iron on repeat hemoglobin measures in those with increasing severity of anemia at baseline (hemoglobin <120 g/L, <115 g/L, and <110 g/L) and *2*) to apply RTM estimates to the treatment effect of iron on hemoglobin for estimation of the RTM-free effect. We hypothesized that as baseline anemia values became more severe, we would observe an increased change in hemoglobin from baseline to 12 wk (endline), which may be partially explained by RTM.

Methods

The original trial

The original trial took place in Kampong Chhnang Province, Cambodia (2015), with ethics granted from the Clinical Research Ethics Board at the University of British Columbia in Vancouver, Canada (H15–00933), and the National Ethics Committee for Health Research in Phnom Penh, Cambodia (110-NECHR). Full details for the original study can be found elsewhere (21) and at clinicaltrials.gov (NCT02481375), but in brief, $n = 809$ women were randomly assigned to daily oral supplementation for 12 wk with either 60 mg elemental iron ($n = 201$; iron group), 14 other micronutrients ($n = 202$), iron and 14 other micronutrients ($n = 206$), or placebo ($n = 200$). Both participants and researchers were blinded to the supplementation group allocations. The study aimed to recruit anemic women [screened as anemic based on a capillary blood sample tested with the HemoCue Hb 301 (HemoCue AB)] who were otherwise healthy, nonpregnant, and aged between 18 and 45 y. Exclusion criteria included the use of medications or food supplements in the previous 3 mo. Once enrolled, all women were treated for possible helminth infection via a deworming tablet (500 mg mebendazole). Venous blood samples were obtained at baseline and at 12 wk (endline) for measurement of hemoglobin concentration (grams/liter) using an automated hematology analyzer (Sysmex XN-1000; Sysmex Corporation). Although capillary samples were used to enroll women who were suspected to have anemia, venous blood samples measured by the automated hematology analyzer (the "gold standard") showed that only 58% of women were anemic at baseline (21). To accomplish the aims of this research, anonymized trial data from the placebo and iron groups were used.

Assessing the data for calculation of RTM

Prior to the calculation of RTM and RTM-free effects, several steps were taken to assess the data in these 2 groups. First, only those with no missing hemoglobin values (measured in venous blood using the hematology analyzer) at baseline and 12 wk (endline) were included. One participant was excluded from the placebo group who had a change in hemoglobin of >20 g/L that could not be biologically explained. Thus, a total of $n = 185$ from the placebo group and $n = 191$ from the iron group were included in the final analyses. Second, the distributions of baseline and endline hemoglobin concentrations in the placebo group were assessed (as the placebo group is used for calculation of RTM); distributions were very slightly left-skewed (see **Supplemental Figure 1**). Calculation of RTM using skewed data may underestimate the RTM effect (22). However, if transforming the data leads to a loss of interpretability, use of the original data can be justified, and CIs can be used to help capture the variability around the true RTM estimate (22). Given the loss of interpretability associated with transformation in this case and the approximate normality of the data through visualization, use of the original untransformed data in calculations, with the presentation of 95% CIs, is justified.

Figures to visually examine RTM

Scatterplots and distribution plots were used to visually assess the data for the occurrence of RTM in the placebo group. First, a scatterplot was created to depict the change in hemoglobin from baseline to 12 wk (endline) against baseline hemoglobin, with a fitted regression line. Second, a distribution plot of hemoglobin at baseline and 12 wk (endline), which identifies participants who had a baseline hemoglobin further than 1 SD above or below the group mean, was created to offer a visual demonstration of RTM upon repeated measurements in those with more extreme baseline values.

Determining cutoffs for the calculation of RTM

The next step in the analysis was to "a priori" select a cutoff used to calculate RTM. The cutoff allows calculation of RTM for all individuals with a baseline measurement above or below that value, and typically represents a high-risk or clinically relevant group (22). Three cutoffs of increasingly extreme baseline hemoglobin values were chosen: <120 g/L (cutoff for anemia for nonpregnant women, thus the most clinically relevant), <115 g/L, and <110 g/L. By evaluating increasingly extreme cutoffs, we were able to investigate whether RTM increases in those with baseline values further from the group mean (which would be expected); as baseline values become more extreme from a clinical perspective (e.g., increased severity of anemia), and are therefore further from the population mean, susceptibility to RTM increases. Additionally, these cutoffs allow each cutoff group to maintain a reasonable sample size for the statistical analyses (<120 g/L: $n = 95$;

TABLE 1 Baseline characteristics of the study population[1]

	Group	
	Placebo	Iron
Total participants, *n*	185	191
Age, y	30 ± 8	31 ± 8
Hemoglobin, g/L	117.1 ± 12.1	115.5 ± 13.8
Median (IQR)	119 (78, 143)	118 (62, 143)
Anemia (Hb <120 g/L), *n*/total *n* (%)	95/185 (51)	115/191 (60)
Ferritin,[2] μg/L	49.3 (15.1, 61.0)	54.3 (17.8, 81.4)
Iron deficiency (ferritin[2] <15 μg/L), *n*/total *n* (%)	46/185 (25)	39/190 (21)
Iron deficiency anemia (ferritin[2] <15 μg/L and Hb <120 g/L), *n*/total *n* (%)	38/185 (21)	34/190 (18)
Genetic hemoglobin disorders, *n*/total *n* (%)	129/185 (70)	151/191 (79)
Inflammatory markers, *n*/total *n* (%)		
Acute inflammation (CRP >5 mg/L)	4/185 (2)	9/190 (5)
Chronic inflammation (AGP >1 g/L)	13/185 (7)	15/190 (8)

[1]Values are means ± SDs or medians (IQRs) unless otherwise noted. Estimates are reported in those from the placebo and iron groups with no missing baseline or endline Hb values, and exclusion of $n = 1$ participant in the placebo group with an unexplained change in Hb >20 g/L. AGP, α1-acid glycoprotein; CRP, C-reactive protein; Hb, hemoglobin.
[2]Ferritin adjusted for inflammation using correction factors (25).

<110 g/L: $n = 72$; <110 g/L: $n = 49$). Although there is no required sample size for calculation of RTM, as it is a nonparametric approach that does not rely on an assumed distribution, the calculation does require use of parameters such as within- and between-group SDs, which are altered by sample size. Additionally, the sample size of each cutoff group in this study is similar to those in previously published reports (2, 23). For those at each cutoff in the placebo group, the mean ± SD for change in hemoglobin from baseline to 12 wk (endline) was calculated and a paired *t* test was conducted.

Calculation of RTM and the RTM-free effect

The RTM effect was calculated in the placebo group, and then applied to the iron group for estimation of the RTM-free effect. Given the random allocation of participants to treatment groups, it is expected that RTM would have occurred consistently across all study participants (23). The RTM effect (estimate and 95% CI) for each cutoff in the placebo group was calculated using the following formula, which has been previously described in numerous reports (2, 22–24):

$$\text{RTM effect} = \frac{\sigma_w^2}{\sqrt{\sigma_w^2 + \sigma_b^2}} \times \text{G}(z) \qquad (1)$$

Here σ_w^2 is the within-subject variance, σ_b^2 is the between-subject variance, and $\text{G}(z) = \frac{\phi(z)}{1-\Phi(z)}$, where $\phi(z)$ is the standard normal probability density function and $\Phi(z)$ is the standard normal cumulative distribution function. As hemoglobin concentrations below the selected cutoffs indicate high-risk, the *z*-score was defined as $z = \frac{\mu - c}{\sigma}$, where c is the cutoff (120 g/L, 115 g/L, or 110 g/L), μ is the mean baseline hemoglobin of the placebo group, and σ is the sum of within-subject and between-subject variances ($\sigma_w^2 + \sigma_b^2$). Bootstrap CIs are presented with point estimates to reflect variability in sampling. The RTM estimate can be compared with the observed change in the placebo group to assess how much of this unexplained variation (which is not due to a treatment effect) may be accounted for by RTM.

To calculate the RTM-free effect in the iron group, the RTM estimate (from the placebo group) for each cutoff was subtracted from the respective mean change in hemoglobin from baseline to 12 wk (endline) at each cutoff in the iron group. Calculations were completed in Stata 16.0 (StataCorp) and R version 3.6.3 (R Core Team 2020).

Results

Baseline characteristics of the study population are described in **Table 1**. Overall, the mean ± SD age of women was 30 ± 8 y and >50% of participants were anemic at baseline (hemoglobin <120 g/L based on the hematology analyzer). Mean ± SD hemoglobin (grams/liter) at baseline in the placebo and iron groups was 117.1 ± 12.1 and 115.5 ± 13.8, respectively. Greater than 20% of women had low iron stores at baseline [inflammation-adjusted (25) ferritin <15 μg/L], ~5% had indication of acute or chronic inflammation [C-reactive protein (CRP) >5 mg/L and α1-acid glycoprotein (AGP) >1 g/L, respectively], and ≥70% had a genetic hemoglobin disorder (most commonly, the hemoglobin E homozygous variant or α-thalassemia).

The presence of RTM is visually apparent in the placebo group as per both the scatterplot (**Figure 1**) and distribution plot (**Figure 2**). In Figure 1, women with more extreme baseline values (both high and low), experienced a greater change in hemoglobin from baseline to 12 wk (endline) than those with baseline values closer to the group mean. Further, the direction of change suggests that individuals at both extreme ends of the distribution regressed towards the mean upon the second measurement. In other words, women with lower baseline values experienced a greater positive change and women with higher baseline values experienced a greater negative change. This is also evident in Figure 2, as most women with baseline hemoglobin concentrations >1 SD away from the mean travelled towards the group mean upon the second measurement.

The mean ± SD of baseline, endline, and total change in hemoglobin for women in the placebo group and *P* values for paired *t* tests of baseline and endline values are presented in **Table 2**. Although, theoretically, minor consistent variation from baseline to endline should occur in the placebo group, change in hemoglobin increased and became more statistically significant as the cutoff became more extreme. For

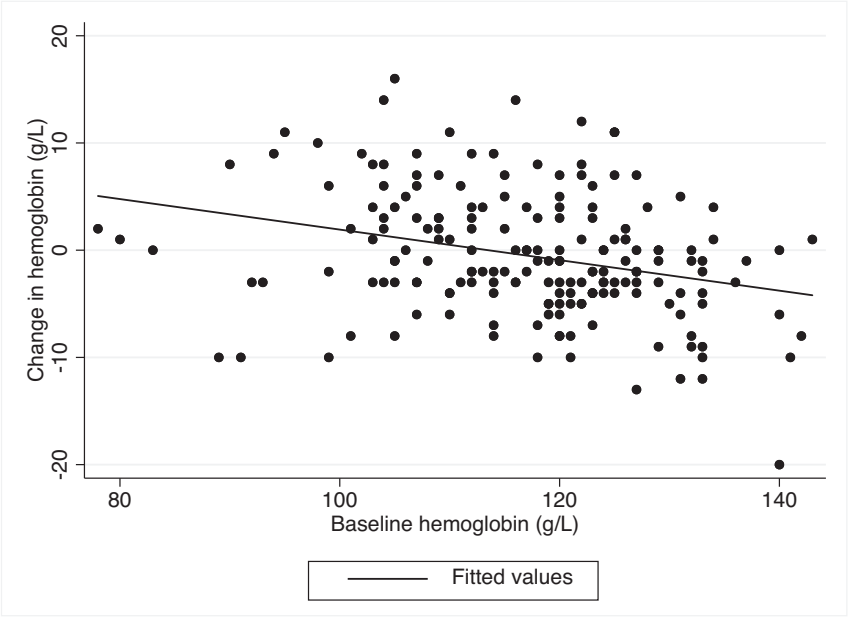

FIGURE 1 Scatterplot of baseline hemoglobin against change in hemoglobin over 12 wk among women in the placebo group ($n = 185$). Each point represents an individual participant from the placebo group and the solid line represents the fitted regression line for this group.

instance, when no cutoff is applied, the mean change in hemoglobin (grams/liter) in the placebo group is small (-0.51 ± 5.8) and not statistically significant ($P = 0.881$, paired t test); however, once baseline hemoglobin values decrease below 115 g/L, the change becomes statistically significant ($P < 0.05$, paired t test). As this group received no treatment (iron), these results may be at least partially explained by RTM.

In **Table 3**, we present RTM estimates and 95% CIs in the placebo group, as well as the RTM-free effect in the iron group and the observed change in hemoglobin (mean ± SD) in both groups at each cutoff (<120 g/L, <115 g/L, and <110 g/L). RTM estimates in the placebo group increase as the hemoglobin cutoffs become more extreme. When comparing the RTM estimate with the mean change in hemoglobin from baseline to 12 wk (endline) in the placebo group, it appears that most (if not all) of the change at each cutoff is accounted for by RTM. For example, those with baseline hemoglobin <120 g/L experienced a mean change in hemoglobin of 0.9 g/L, and the RTM estimate for this group was 1.0 g/L; thus, we speculate that the change experienced in this group is likely due to RTM. In the iron group, mean change in hemoglobin from baseline to endline also increases as the cutoffs become more extreme (Table 3). However, when applying RTM estimates to the observed change in the iron group at each cutoff, RTM did not account for the full treatment effect of iron; the percentage of the treatment effect in the iron group that may be due to RTM remained relatively consistent across all cutoffs (~10%). In other words, the RTM-free effect (e.g., genuine treatment effect) of iron supplementation is ~10% lower than the observed change from baseline to 12 wk (endline) in the iron group. Nonetheless, despite the occurrence of RTM in this study, iron supplementation appears to have a genuine treatment effect on improving hemoglobin concentrations, which increases in those with lower baseline values.

Discussion

To our knowledge, this is the first study to report the effect of RTM on repeated hemoglobin measurements (baseline and endline values) from an iron supplementation trial. These findings indicate that RTM was present and may have accounted for ~10% of the treatment effect of iron supplementation on increasing hemoglobin concentrations in the studied population. However, despite occurrence of RTM, iron supplementation was still more effective in those with lower baseline hemoglobin concentrations, based on calculation of the RTM-free effect at increasingly severe anemic cutoffs (hemoglobin <120 g/L, <115 g/L, and <110 g/L).

When calculating RTM, it is assumed that any variation from baseline to endline in the study outcomes among those receiving no intervention is due to RTM (22). However, we acknowledge that there are other factors that may have contributed to the increase in hemoglobin concentration observed in the placebo group. First, all women in the trial were provided with a deworming tablet at baseline. Women in the placebo group who were positive for parasitic infection may have benefited from this intervention, thus showing an increased hemoglobin concentration after 12 wk. However, parasitic infection was not measured among women enrolled in the trial; therefore, we cannot ascertain if this was indeed a potential cause of increased hemoglobin among women in the placebo group. Any dietary changes throughout the intervention, including seasonal food availability or motivation to improve dietary habits (as a participant of a nutrition trial), may have also

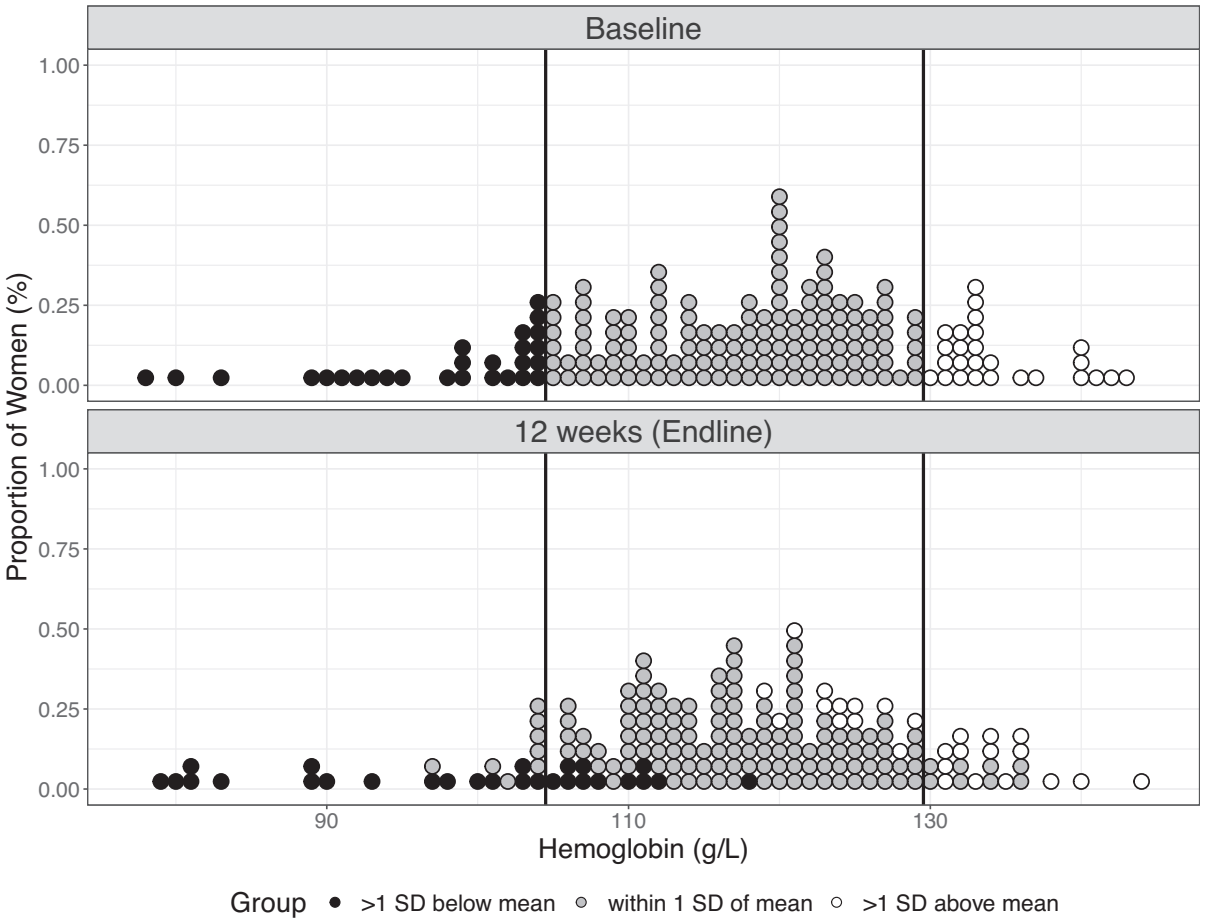

FIGURE 2 Distribution plot of hemoglobin concentration at baseline and 12 wk (endline) among women in the placebo group. Black solid lines are set 1 SD away from the group mean. Black points represent individuals with baseline hemoglobin 1 SD below the mean, white points represent individuals with baseline hemoglobin 1 SD above the mean, and gray points represent individuals with baseline hemoglobin within 1 SD of the mean. Hemoglobin concentrations at baseline and 12 wk (endline) are displayed in the top and bottom graphs, respectively. Individuals >1 SD away from the mean (e.g., black and white points) tended to travel towards the group mean upon second measurement, demonstrating potential movement due to RTM. RTM, regression to the mean.

contributed to the observed increase in hemoglobin. Finally, our study population had a high prevalence of genetic hemoglobin disorders (≥70%). These inherited conditions result in decreased or defective hemoglobin synthesis and function; thus, the hemoglobin response to iron supplementation may be inherently altered in this specific disease state (26). The placebo effect is unlikely to have played a significant role, given the design of the trial (e.g., double-blinded design) and the use of biochemical outcomes (21). Nonetheless, given our findings in the placebo group that the change in hemoglobin became larger and more statistically significant as baseline hemoglobin decreased (also evident in Figures 1 and 2), we are confident that RTM was present and likely played a central role in the increase in hemoglobin observed in the placebo group.

RTM estimates were calculated using within- and between-individual variances and mean baseline and endline hemoglobin values specific to our study sample. Large SDs for change in hemoglobin were observed at all cutoffs in both the placebo and iron groups (Table 3). We cannot ascertain why this occurred, but we speculate this may be due to the high prevalence of genetic hemoglobin disorders in this population (resulting in increased heterogeneity). In the original trial, hemoglobin genotyping was ascertained for all participants, which included hemoglobin electrophoresis for hemoglobin E (and other structural variants) and a StripAssay kit for α-globin gene deletions and point mutations. A high prevalence of genetic hemoglobin disorders (most commonly, the hemoglobin E homozygous variant or α-thalassemia) was found, which is typical for the Southeast Asia region. As shown previously in other analyses, the proportion of hemoglobin responders to iron supplementation is significantly altered by the presence of a genetic hemoglobin disorder (21). However, rates of hemoglobin disorders were similar between the placebo and iron groups (Table 1; 79% and 70%, respectively); thus, we expect that any contribution that this had to the RTM estimation would be consistent across the full cohort. Nonetheless, given the large variation in hemoglobin concentrations within this study, the RTM-free effect estimate would inevitably

TABLE 2 Baseline and endline hemoglobin concentrations and overall change in hemoglobin concentration after 12 wk among women in the placebo group[1]

Placebo[2]	*n*	Baseline Hb, g/L	Endline Hb, g/L	Change in Hb, g/L	*P*[3]
Hb cutoff					
<110 g/L	49	101.5 ± 7.6	103.4 ± 10.4	1.9 ± 6.1	0.016*
<115 g/L	72	104.9 ± 8.0	106.2 ± 9.9	1.3 ± 5.9	0.033*
<120 g/L	95	107.8 ± 8.8	108.8 ± 10.1	0.9 ± 5.8	0.060
No cutoff	185	117.1 ± 12.1	116.6 ± 11.8	−0.5 ± 5.8	0.881

[1]Values are means ± SDs. *Significant, $P < 0.05$. Hb, hemoglobin.
[2]$n = 185$ women were allocated to the placebo group; each "Hb cutoff" represents women with a baseline hemoglobin below that value; "no cutoff" reflects the complete placebo group.
[3]*P* values obtained using a paired *t* test.

vary across individuals within a defined cutoff group, and therefore RTM estimates should not be applied on an individual scale. The inclusion of 95% CIs helps to account for the uncertainty of the RTM estimate.

Calculation of the RTM-free effect can be used to disentangle whether there is true effect modification of an intervention based on baseline status, or if the increased change in those with more extreme baseline values is simply due to RTM (4). We found that, among women who received iron, an increased change in hemoglobin from baseline to endline was observed in those with lower baseline hemoglobin values, which persisted after applying the RTM estimate when calculating the RTM-free effect. From a physiological perspective, this is feasible. Absorption of iron is notably influenced by iron status, with an increased absorption among individuals with lower iron stores (20, 27). Iron is essential for hemoglobin synthesis; thus, hemoglobin is often used as a marker of iron status. A systematic review including 41 randomized controlled trials that enrolled a healthy adult population (men and women aged ≥18 y) investigated the effect of iron supplementation on iron status (28). Iron was shown to have a significantly greater effect on increasing hemoglobin in those with iron deficiency anemia (hemoglobin <120 g/L and ferritin <15 μg/L) as compared with those with solely iron deficiency or normal iron status (28).

Strengths of this study include the large sample size and rigorous design of the original clinical trial, including random allocation to treatment groups, double-blinding, and use of gold-standard methods for hemoglobin measurement (automated hematology analyzer). Calculation of RTM may be limited by the use of slightly skewed data, as this has been shown to underestimate the true RTM effect (22). Another possible limitation may include the generalizability of our results to populations outside of Southeast Asia (given the high prevalence of genetic hemoglobin disorders in this region, as previously discussed), and given that the study aimed to recruit only anemic women (the mean baseline hemoglobin of all participants was <120 g/L). We only calculated RTM in those with hemoglobin below clinically relevant cutoffs (determined "a priori"); however, it appears that participants on both ends of the distribution in the placebo group regressed towards the group baseline mean of 117 g/L (Figure 2). As such, calculation of the RTM estimate for those at the cutoff for anemia (hemoglobin <120 g/L) should be interpreted with some caution, as this cutoff was higher than the group mean (117 g/L) and may include individuals regressing both upwards and downwards towards the group mean. In addition, as RTM estimates are calculated using study-specific mean and variance values, estimates may differ in a study of the general population with a higher population mean, where participants are not recruited based on hemoglobin status.

TABLE 3 Change in hemoglobin after 12 wk and RTM effects among women in the iron and placebo groups[1]

	n	Change in Hb, g/L	RTM estimate, g/L (95% CI)	RTM-free effect,[2] g/L (95% CI)	Percentage of treatment effect accounted for by RTM[3]
Placebo[4]	185				
Hb cutoff					
<110 g/L	49	1.9 ± 6.1	1.8 (1.3, 2.2)	—	—
<115 g/L	72	1.3 ± 5.8	1.3 (1.0, 1.7)	—	—
<120 g/L	95	0.9 ± 5.8	1.0 (0.7, 1.2)	—	—
Iron group[4]	191				
Hb cutoff					
<110 g/L	56	15.4 ± 17.7	—	13.6 (13.2, 14.0)	11
<115 g/L	81	12.2 ± 16.2	—	10.9 (10.5, 11.2)	10
<120 g/L	115	9.7 ± 14.6	—	8.7 (8.5, 9.0)	10

[1]Values are means ± SDs unless otherwise indicated. Hb, hemoglobin; RTM, regression to the mean.
[2]RTM-free effect = mean change in Hb in iron group − RTM estimate, and 95% CI.
[3]The percentage of the treatment effect (%) in the iron group that may be accounted for by RTM as calculated as: RTM estimate/mean change in Hb in the iron group.
[4]$n = 185$ women were allocated to the placebo group, $n = 191$ women were allocated to the iron group.

In conclusion, our findings indicate that RTM was present and did impact the hemoglobin measurements in this iron supplementation trial. The inclusion of a placebo group in the original trial design appropriately controlled for this RTM effect in the statistical analyses between groups and reflected accuracy in the published results (11). However, our post hoc calculation of the RTM in this study highlights the importance of the inclusion of a placebo group in future iron supplementation trials (especially trials that aim to recruit anemic individuals) to avoid misinterpretation of results. In order to accurately estimate within-group treatment effects, one should calculate RTM within the placebo group, and apply it to the observed change in the intervention group, in order to ascertain the RTM-free effect. We estimate that, for those who were anemic at baseline, the RTM-free effect of iron supplementation on increasing hemoglobin concentrations after 12 wk is ∼10% lower than the observed change in hemoglobin. We are unable to determine what effect this has from a clinical perspective (e.g., to estimate the number of individuals who would move from being classified as anemic to nonanemic following application of RTM) given that RTM estimates are applied to the full group and not on an individual scale. Thus, whether a 10% reduction in the treatment effect of iron is clinically meaningful to policymakers may depend on several other factors and warrants further investigation. Our aim is that these findings will help inform best practice in the field of nutritional anemia and highlight the importance of appropriate trial design, including use of a placebo group and randomization of participants to study groups, in order to account for the influence of RTM. In cases where inclusion of a placebo group is not possible, one can minimize the effect of RTM in the design phase (through use of repeated baseline measurements) or in the analysis phase (using modified calculations for RTM estimation or ANCOVA) (23).

Acknowledgments

We thank Dusty Turner (Department of Mathematical Sciences, US Military Academy, West Point, NY) for guidance on the use of the RTM formulas and statistical code and Lang Wu (Department of Statistics, University of British Columbia) for providing guidance on the statistical derivations. The authors' responsibilities were as follows—CDK: conceived the project, conducted research (conducted data collection in the original trial), and had primary responsibility for final content; KMC: designed and spearheaded the research (development of overall research plan and research oversight); KMC and LXP: analyzed data and performed statistical analyses; KMC and BAW: drafted the manuscript; KMC, LXP, CDK, JAJF, and KLIS: developed the figures and tables; and all authors: contributed to revising content until the final version and read and approved the final manuscript.

References

1. Stigler SM. Regression towards the mean, historically considered. Stat Methods Med Res 1997;6:103–14.
2. Thomas DM, Clark N, Turner D, Siu C, Halliday TM, Hannon BA, Kahathuduwa CN, Kroeger CM, Zoh R, Allison DB. Best (but oft-forgotten) practices: identifying and accounting for regression to the mean in nutrition and obesity research. Am J Clin Nutr 2020;111:256–65.
3. Skinner AC, Goldsby TU, Allison DB. Regression to the mean: a commonly overlooked and misunderstood factor leading to unjustified conclusions in pediatric obesity research. Childhood Obesity 2016;12:155–8.
4. Allison DB, Loebel AD, Lombardo I, Romano SJ, Siu CO. Understanding the relationship between baseline BMI and subsequent weight change in antipsychotic trials: Effect modification or regression to the mean? Psychiatry Res 2009;170:172–6.
5. Morton V, Torgerson D. Effect of regression to the mean on decision making in health care. BMJ 2003;326:1083.
6. Kahathuduwa CN, Thomas DM, Siu CO, Allison DB. Unaccounted for regression to the mean renders conclusion of article titled "Uric acid lowering in relation to HbA1c reductions with the SGLT 2 inhibitor tofogliflozin" unsubstantiated. Diabetes Obes Metab 2018;20:2039–40.
7. Kroeger CM, Thomas DM, Siu CO, Allison DB. Overlooking regression to the mean leads to unwarranted interpretations: Letter concerning, "Do obese and extremely obese patients lose weight after lumbar spine fusions? Analysis of a cohort of 7303 patients from the Kaiser National Spine Registry. Spine 2018;43:E492–3.
8. Cummings S, Palermo L, Browner W, Marcus R, Wallace R, Pearson J, Blackwell T, Eckert S, Black D. Monitoring osteoporosis therapy with bone densitometry: misleading changes and regression to the mean. JAMA 2000;283:1318.
9. Asmar R, Safar M, Queneau P. Evaluation of the placebo effect and reproducibility of blood pressure measurement in hypertension. Am J Hypertens 2001;14:546–52.
10. Salam A, Atkins E, Sundström J, Hirakawa Y, Ettehad D, Emdin C, Neal B, Woodward M, Chalmers J, Berge E, et al. Effects of blood pressure lowering on cardiovascular events, in the context of regression to the mean: a systematic review of randomized trials. J Hypertens 2019;37:16–23.
11. Denke MA, Frantz ID. Response to a cholesterol-lowering diet: efficacy is greater in hypercholesterolemic subjects even after adjustment for regression to the mean. Am J Med 1993;94:626–31.
12. Forrow L, Calkins DR, Allshouse K, Horowitz G, Delbanco TL. Evaluating cholesterol screening: the importance of controlling for regression to the mean. Arch Intern Med 1995;155:2177–84.
13. Neogi SB, Devasenapathy N, Singh R, Bhushan H, Shah D, Divakar H, Zodpey S, Malik S, Nanda S, Mittal P, et al. Safety and effectiveness of intravenous iron sucrose versus standard oral iron therapy in pregnant women with moderate-to-severe anaemia in India: a multicentre, open-label, phase 3, randomised, controlled trial. Lancet Glob Health 2019;7:e1706–16.
14. Kaundal R, Bhatia P, Jain A, Jain A, Nampoothiri RV, Mishra K, Jandial A, Goni D, Sandal R, Jindal N, et al. Randomized controlled trial of twice-daily versus alternate-day oral iron therapy in the treatment of iron-deficiency anemia. Ann Hematol 2020;99:57–63.
15. Stoffel NU, Zeder C, Brittenham GM, Moretti D, Zimmermann MB. Iron absorption from supplements is greater with alternate day than with consecutive day dosing in iron-deficient anemic women. Haematologica 2020;105:1232–9.
16. Abbas AM, Abdelbadee SA, Alanwar A, Mostafa S. Efficacy of ferrous bisglycinate versus ferrous glycine sulfate in the treatment of iron deficiency anemia with pregnancy: a randomized double-blind clinical trial. J Matern Neonatal Med 2019;32:4139–45.
17. Moriarty-Craige SE, Ramakrishnan U, Neufeld L, Rivera J, Martorell R. Multivitamin-mineral supplementation is not as efficacious as is iron supplementation in improving hemoglobin concentrations in nonpregnant anemic women living in Mexico. Am J Clin Nutr 2004;80:1308–11.
18. World Health Organization. Haemoglobin concentrations for the diagnosis of anaemia and assessment of severity. Geneva (Switzerland): World Health Organization; 2011.
19. World Health Organization. Guideline: daily iron supplementation in adult women and adolescent girls. Geneva (Switzerland): World Health Organization; 2016.
20. Gibson RS. The role of diet- and host-related factors in nutrient bioavailability and thus in nutrient-based dietary requirement estimates. Food Nutr Bull 2007;28:S77–100.

21. Karakochuk CD, Barker MK, Whitfield KC, Barr SI, Vercauteren SM, Devlin AM, Hutcheon JA, Houghton LA, Prak S, Hou K, et al. The effect of oral iron with or without multiple micronutrients on hemoglobin concentration and hemoglobin response among nonpregnant Cambodian women of reproductive age: a 2 x 2 factorial, double-blind, randomized controlled supplementation trial. Am J Clin Nutr 2017;106:233–44.
22. Linden A. Assessing regression to the mean effects in health care initiatives. BMC Med Res Methodol 2013;13:1–7.
23. Barnett AG, van der Pols JC, Dobson AJ. Regression to the mean: what it is and how to deal with it. Int J Epidemiol 2004;34:215–20.
24. Davis CE. The effect of regression to the mean in epidemiologic and clinical studies. Am J Epidemiol 1976;104:493–8.
25. Thurnham DI, McCabe LD, Haldar S, Wieringa FT, Northrop-Clewes CA, McCabe GP. Adjusting plasma ferritin concentrations to remove the effects of subclinical inflammation in the assessment of iron deficiency: a meta-analysis. Am J Clin Nutr 2010;92:546–55.
26. Bain B. Haemoglobinopathy diagnosis. 2nd ed. Oxford, United Kingdom: Blackwell Publishing Ltd.; 2006.
27. Miret S, Simpson RJ, Mckie AT. Physiology and molecular biology of dietary iron absorption. Annu Rev Nutr 2003;23:283–301.
28. Casgrain A, Collings R, Harvey LJ, Hooper L, Fairweather-tait SJ. Effect of iron intake on iron status: a systematic review and meta-analysis of randomized controlled trials. Am J Clin Nutr 2012;96: 768–80.