ABSTRACT. Objective. To determine the long-term effects of iron deficiency in infancy.

Design. Longitudinal follow-up study of children who had been tested and treated for iron deficiency as infants.

Setting. Periurban community near San Jose, Costa Rica.

Participants. Of the original 191 participants, 87% were reevaluated at 11 to 14 years old (average age: 12.3 years). The children were free of iron deficiency and growing normally by US standards. Those who had chronic, severe iron deficiency in infancy \( (n = 48) \) were compared with those who had good iron status before and/or after iron therapy in infancy \( (n = 114) \).

Outcome Measures. Comprehensive set of cognitive, socioemotional, and motor tests and measures of school functioning.

Results. Children who had severe, chronic iron deficiency in infancy scored lower on measures of mental and motor functioning. After control for background factors, differences remained statistically significant in arithmetic achievement and written expression, motor functioning, and some specific cognitive processes (spatial memory, selective recall, and tachistoscopic threshold). More of the formerly iron-deficient children had repeated a grade and/or been referred for special services or tutoring. Their parents and teachers rated their behavior as more problematic in several areas, agreeing in increased concerns about anxiety/depression, social problems, and attention problems.


ABBREVIATIONS. Hb, hemoglobin; HOME, Home Observation for Measurement of the Environment; SD, standard deviation; CBCL, Child Behavior Checklist; TRF, Teacher Report Form; CESD, Center of Epidemiologic Studies Depression Scale; SE, standard error.

From the Center for Human Growth and Development and Departments of *Pediatrics and Communicable Diseases and †Psychology, University of Michigan, Ann Arbor, Michigan; ‡Hospital Nacional de Niños, San Jose, Costa Rica; ¶Department of Pediatrics and Communicable Diseases, University of Michigan, Ann Arbor, Michigan; and †Department of Psychiatry, MetroHealth Medical Center, Case Western Reserve University School of Medicine, Cleveland, Ohio.

Preliminary results were presented at the Academic Pediatric Societies Meeting; May 5, 1997; Washington, DC.

Received for publication Jun 3, 1999; accepted Nov 9, 1999.

Reprint requests to (B.L.) Center for Human Growth and Development, 300 N Insalls, University of Michigan, Ann Arbor, MI 48109-0406. E-mail: blozoff@umich.edu

PEDIATRICS (ISSN 0031 4005). Copyright © 2000 by the American Academy of Pediatrics.

http://www.pediatrics.org/cgi/content/full/105/4/e51

METHODS

Sample

This study reassessed a group of Costa Rican children who participated in earlier phases of the project as infants and at 5 years of age. All aspects of the research have been conducted in collaboration with the Hospital Nacional de Niños, San Jose, Costa Rica.
Rica and supported by the National Institute of Child Health and Human Development. The original study involved 191 infants from the urban community of Hatillo, located at an elevation of 1100 m near San Jose, the capital of Costa Rica. The community was predominantly working class, and parents averaged 8 to 10 years of education. Enrollment entailed door-to-door screening of the entire community and included all 12- to 23-month-old infants who had been born with birth weight $\geq 2.5$ kg, of singleton, uncomplicated births, who were free of acute or chronic medical problems, and who had normal physical examinations. These healthy infants had relatively low lead levels, no evidence of growth failure or other nutrient deficiencies, and were free of parasites except in 5%. Iron status in infancy, determined by venous levels of hemoglobin (Hb), transferrin saturation, erythrocyte protoporphyrin, and serum ferritin, varied from iron sufficiency to moderate iron deficiency anemia (Hb $\leq 100$ g/L, low ferritin, and either high erythrocyte protoporphyrin or low transferrin saturation). Comprehensive information was collected about the child and family, including demography, birth history, nutrition, socioeconomic status, stimulation in the home, and parental IQ. Developmental tests were administered before and both 1 week and 3 months after intramuscular or closely supervised oral administration of iron with appropriate placebo controls. Details of the original study have been published previously.10

Infants with moderate iron deficiency anemia tested lower than the rest of the sample in mental and motor functioning. Those with mild anemia showed lower motor scores.10 There were other behavioral differences, such as increased contact with the mother, wariness or hesitance, fewer attempts at test items, and others.18 Hematologic response to iron therapy was excellent with an average Hb increase of 37 g/L among anemic infants. All iron-deficient anemic infants corrected their anemia with 3 months of therapy, although most infants still had biochemical alterations, such as elevated erythrocyte protoporphyrin values. As a group, formerly anemic infants still had lower developmental test scores after treatment. The minority who became iron-sufficient showed improvement in motor scores and no decline in mental scores.10

For the 5-year follow-up study, 85% of the original cohort participated in a comprehensive psychoeducational assessment.10 Children who had moderate iron deficiency anemia as infants continued to test lower in mental and motor functioning, as did children with higher hemoglobin levels who still had some biochemical evidence of iron deficiency after iron therapy in infancy. Both groups had experienced more severe and chronic iron deficiency in infancy, evidenced by lower initial hemoglobin levels and higher initial erythrocyte protoporphyrin values.20 However, both groups had responded to iron with an average hemoglobin increase greater than the gold standard of 10 g/L or more.27

Of the original infant cohort, 167 (87%) participated in the reevaluation in early adolescence. Included were 22 children who could not be located for the 5-year evaluation and 3 children who were lost during the infancy study but located for testing during both follow-up studies. Eighteen children who were evaluated at 5 years could not be located. Children who were tested at this follow-up did not differ from those who could not be located in background characteristics, iron status in infancy, or initial developmental test scores. The average age at testing was 12.3 years, with a range of 10.9 to 13.7 years.

Signed informed consent was obtained by the project pediatrician. The follow-up protocol was approved by the institutional review boards of the University of Michigan, Ann Arbor; the Hospital Nacional de Niños, Costa Rica; and the Office of Protection of Research Risks, National Institutes of Health.

### Procedure

A complete physical examination was part of the early adolescent assessment, with pubertal development determined by a standard protocol.20,29 Venipuncture blood specimens were collected to determine current iron status and stress-responsive hormones.30 Anemia was defined as a hemoglobin $< 118$ g/L for children of both genders who were $< 12$ years old, $< 119$ g/L for girls 12 years old or more, and $< 126$ g/L for boys 12 years old or more. Iron deficiency was defined as 2 or more measures of iron status in the deficient range—serum ferritin $< 12$ mg/L, transferrin saturation $< 14$%, mean cell volume $< 78$ fl. With the exception of mean cell volume,24 these hematology cutoffs were based on the National Health and Nutrition Examination Survey III.4 Because hemoglobin values at this altitude would be slightly higher than those at sea level (2–4 g/L),29,30 we also determined whether any iron-deficient child was anemic using these higher cutoffs.

The rest of the evaluation consisted of a comprehensive psychoeducational and behavioral test battery, including school functioning, socioeconomic status, stimulation in the home,25 and either high erythrocyte protoporphyrin or low transferrin saturation. Because hemoglobin values at this altitude would be slightly higher than those at sea level (2–4 g/L),29,30 we also determined whether any iron-deficient child was anemic using these higher cutoffs.

The rest of the evaluation consisted of a comprehensive psychoeducational and behavioral test battery, including school functioning, socioeconomic status, stimulation in the home,25 and either high erythrocyte protoporphyrin or low transferrin saturation. Because hemoglobin values at this altitude would be slightly higher than those at sea level (2–4 g/L),29,30 we also determined whether any iron-deficient child was anemic using these higher cutoffs.

### Cognitive and Motor Testing

The composition of the test battery is shown in Table 1. The measures were selected on conceptual grounds to be sensitive to

<table>
<thead>
<tr>
<th>TABLE 1. Cognitive and Motor Test Battery</th>
</tr>
</thead>
<tbody>
<tr>
<td>Test</td>
</tr>
<tr>
<td>------------------------------------------</td>
</tr>
<tr>
<td>Overall functioning</td>
</tr>
<tr>
<td>Wechsler Intelligence Scale for Children-Revised78</td>
</tr>
<tr>
<td>Wide Range Achievement Test-Revised79</td>
</tr>
<tr>
<td>Directed Writing Task</td>
</tr>
<tr>
<td>Bender Visual–Motor Gestalt Test81</td>
</tr>
<tr>
<td>Bruninks-Oseretsky Test of Motor Proficiency82</td>
</tr>
<tr>
<td>Specific cognitive processes</td>
</tr>
<tr>
<td>Central/Incidental Serial Recall Test83</td>
</tr>
<tr>
<td>Attentional Capacity Test84</td>
</tr>
<tr>
<td>Underlining Test85</td>
</tr>
<tr>
<td>K-ABC Spatial Memory44</td>
</tr>
<tr>
<td>Tactual Performance Task86</td>
</tr>
<tr>
<td>Cognitive Abilities Tests86</td>
</tr>
</tbody>
</table>
processes unfolding in early adolescence, achieve a balance be-
tween standardized tests and process-oriented measures, and pur-
sue results of previous research on early nutritional insults. Prac-
tical considerations were psychometric soundness, availability in
Spanish, and common use in Costa Rica. Each child’s assessment
was conducted during 1 morning and 1 afternoon of testing, with
frequent breaks and a snack. Two thoroughly trained Costa Rican
psychologists, who were unaware of the children’s iron status or
response to treatment in infancy, administered the test battery.
Morning test sessions were conducted by 1 of the psychologists,
while afternoon sessions were conducted by the other. Thus, all
children received tests at comparable times of day, and the same
tester administered his or her subset of tests to all children.

The follow-up assessment used several measures that were
developed and standardized in the United States. Although ap-
plying such tests in another culture is always of concern, the
overall means for this group of Costa Rican children were gener-
cally close to US norms, as they had been during the infant and
5-year assessments. For instance, IQ scores at 11 to 14 years old
averaged 102 ± 15 standard deviation (SD), close to the norm for
this test (100 ± 15 SD), and motor test scores averaged 46 ± 12 SD,
also similar to the test norm of 50 ± 10 SD. These observations
suggest that the measures were not inappropriate for use in Costa
Rica. Furthermore, the stringent entrance criteria in this study
were effective in identifying a group of children who continue to
test in the normal range by US standards. However, a Costa Rican
comparison group has always been an essential component of the
study. The purpose is to compare formerly iron-deficient children
with their Costa Rican peers who had better iron status in infancy,
rather than to compare them with US children.

Behavior
The Spanish versions of the Child Behavior Checklist (CBCL) and
the Teacher Report Form (TRF) were administered to provide
independent but parallel assessments of the children’s behavior at
home and school.36 The children’s behavior was rated during
several stressors (physical examination, blood drawing, and hav-
ing dental examination).36 In addition, the testing session in-
cluded a brief, socially stressful situation that might bring out
wariness or hesitance in the young adolescent.37

Scholastic Achievement
School records and a teacher questionnaire were used to deter-
mine grade retention, requests for special testing or tutoring, and
placement in special classrooms.

Statistical Analysis
The approach to analyzing the early adolescent data was based on
the results of the 5-year follow-up. Preliminary analyses of the
early adolescent data showed a similar pattern to that observed at
5 years old: similar test scores among children who had moderate
iron deficiency anemia as infants and those with higher hemoglo-
bin levels who had biochemical evidence of iron deficiency after
treatment, even though they had responded with an increase in
hemoglobin.35 Therefore, all further analyses combined them to
form a severe, chronic iron-deficient group (henceforward re-
ferr ed to as the iron-deficient group; n = 48) for comparison with
children who had good iron status before and/or after treatment
in infancy (combined to form a good-iron-status group; n = 114).
Data for 5 additional adolescents could not be included in these
analyses. They had missed the blood sample after 3 months of

treatment in infancy, and hence, their response to therapy was
unknown and could not be inferred with confidence.

Analysis of covariance was the primary statistical approach.
For standardized tests that take age into account, the standard
score was analyzed. For other psychoeducational measures, we
included age as a covariate. In addition, we considered the effects
of mediating and masking variables. We expected that children
who had severe, chronic iron deficiency in infancy might grow up
in less advantaged family environments, even within a relatively
homogeneous community.36 To provide additional information on
the relation of other factors to psychoeducational outcome in
adolescence, we first examined zero-order correlations between
potential control variables and test scores. In general, test scores in
early adolescence were unrelated to characteristics in infancy
(birth weight, feeding, etc.) but correlated with gender, mother’s
IQ or education, and/or HOME scores, and the iron-deficient and
good-iron-status groups differed on these factors.10,36 All analyses
of cognitive and motor differences were repeated with these co-
vat iates.

Age, gender, mother’s IQ, and HOME scores also correlated with
CBCL T scores (parent report). In addition, pubertal develop-
ment, current size, and maternal depression, which were gen-
erally unrelated to cognitive/motor test scores, correlated with
parental ratings of child behavior. Pubertal development and
height-for-age percentiles were added to the covariate set for CBCL
analyses.

Considering the role of maternal depression raised different
issues, however. Our measures of maternal depression were ob-
tained when the children were 5 years old, rather than concurrently
with the early adolescent behavior ratings, and depression data
were not available for ~20% of mothers. Despite the missing
information on maternal depression and the earlier time of data
collection, we believed it was important to take this factor into
account. Depressed mothers tend to report more behavior prob-
lems in their children, which may be an accurate reflection of the
child’s behavior or a distortion related to the mother’s distress,
etc.38 Although there were no statistically significant differences
between groups in mean scores on the Center of Epidemiologic
Studies Depression Scale (CESD), the proportion with scores
above the clinical cutoff (CESD ≥16), or the proportion who met
criteria for a lifetime diagnosis of major depression on the Diag-
nostic Interview Schedule,39 we conducted additional covariate
analyses for the CBCL including maternal depression (CESD) as a
covariate. In considering moderating and masking variables for
the TRF data (the teacher equivalent of the CBCL), we included
child characteristics that might affect teacher perceptions and/or
child behavior in the classroom—age, gender, height-for-age per-
centile, and pubertal development.

Results for continuous variables are reported with and without
adjustment for the relevant set of background variables. Given the
hypothesis-testing nature of our analyses and the recent argument
that routine adjustment for multiple comparisons is not in the best
interest of good empirical science,42 a significance level of .05 is
used throughout. All analyses were performed with standard
statistical packages (SAS, Cary, NC; SPSS, Chicago, IL; and Systat,
Chicago, IL).

RESULTS

Growth and Nutrition (Table 2)
A greater proportion of the iron-deficient group was male (71% vs 49%; χ² = 6.45; P = .01). Children who had been iron-deficient in infancy were some-
what younger, on average, than those in the good-
iron-status group (12.1 years old ± .1 standard error [SE] vs 12.4 years old ± .1 SE; t(160) = 3.16; P < .01). The children’s growth and iron status were excellent at the transition to adolescence, as they had been at
the 5-year follow-up.20 With the exception of boys’
height, growth averaged between the 40th and 50th
percentile by US standards for age and gender. No
child had iron deficiency anemia (with or without
adjustment of hemoglobin cutoffs for altitude), and
only 2% (n = 3) had biochemical evidence of iron
deficiency.

Standardized Tests (Table 3)
Compared with children who had good iron status in
infancy, the iron-deficient group showed lower Verbal and Full-Scale IQ scores and lower achievement


http://www.pediatrics.org/cgi/content/full/105/4/e51
TABLE 2. Nutritional Status in Early Adolescence*

<table>
<thead>
<tr>
<th>(n)</th>
<th>Severe, Chronic Iron Deficiency in Infancy (48)</th>
<th>Good Iron Status in Infancy (114)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Age (y)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Girls</td>
<td>12.1 ± .2</td>
<td>12.3 ± .1</td>
</tr>
<tr>
<td>Boys†</td>
<td>12.1 ± .1</td>
<td>12.6 ± .1</td>
</tr>
<tr>
<td>Weight-for-age percentile</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Girls</td>
<td>49.8 ± 9.9</td>
<td>45.3 ± 4.2</td>
</tr>
<tr>
<td>Boys</td>
<td>43.7 ± 5.5</td>
<td>39.7 ± 4.2</td>
</tr>
<tr>
<td>Height-for-age percentile</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Girls</td>
<td>46.1 ± 7.7</td>
<td>39.2 ± 4.3</td>
</tr>
<tr>
<td>Boys</td>
<td>36.6 ± 4.8</td>
<td>31.9 ± 3.3</td>
</tr>
<tr>
<td>Hb (g/L)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Girls</td>
<td>140.5 ± 3.0</td>
<td>142.6 ± 1.5</td>
</tr>
<tr>
<td>Boys</td>
<td>140.2 ± 1.9</td>
<td>142.1 ± 1.4</td>
</tr>
<tr>
<td>Transferrin saturation (%)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Girls</td>
<td>22.1 ± 2.1</td>
<td>25.3 ± 1.1</td>
</tr>
<tr>
<td>Boys</td>
<td>22.2 ± 1.1</td>
<td>24.5 ± 0.9</td>
</tr>
<tr>
<td>Mean cell volume (fl)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Girls</td>
<td>84.5 ± 1.2</td>
<td>85.0 ± 0.6</td>
</tr>
<tr>
<td>Boys</td>
<td>83.4 ± 7</td>
<td>84.2 ± 6</td>
</tr>
<tr>
<td>Serum ferritin (μg/L)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Girls</td>
<td>22.1 ± 5.1</td>
<td>28.0 ± 2.5</td>
</tr>
<tr>
<td>Boys</td>
<td>25.4 ± 3.0</td>
<td>30.3 ± 2.3</td>
</tr>
</tbody>
</table>

Values are means ± SE. Hematologic analyses include age as a covariate, and the values shown are age-adjusted. Tests of statistical significance of the differences between the iron-deficient and good-iron-status groups are based on analysis of variance or covariance.  
† P < .001.

justed values, became statistically significant after control for background factors, primarily attributable to the effect of gender. Although differences in IQ were no longer statistically significant after control for background variables, significant or suggestive differences remained for the subtests of Similarities (F(1,157) = 4.75; P = .03) and Information (F(1,157) = 3.16; P = .08), which assess the important areas of abstraction and informal learning.

Figure 1 shows the magnitude of differences between the iron-deficient and good-iron-status groups with and without adjustment for background variables. For measures with statistically significant differences, children who had been iron-deficient tested .4 to .7 SD lower than their peers who had good iron status as infants. Differences in effect size did not indicate a clear pattern of more marked difficulty in tasks requiring visual–motor and motor skills.

Specific Cognitive Functions

Formerly iron-deficient children, regardless of age, did more poorly on the K-ABC Spatial Memory Task,43 which involves visual-perceptual vigilance, attentiveness, and visual-spatial memory. The raw score (with age as a covariate) of the formerly iron-deficient group averaged 12.8 ± 4 SE compared with 13.9 ± .4 SE in the good-iron-status group (F(2,158) = 4.86; P < .03). The findings were virtually identical after additional statistical control for gender, mother’s IQ, and HOME scores.

Children who were either iron-deficient or had good iron status in infancy also seemed to differ in some cognitive transitions that occur in early adolescence. For instance, US children show a shift in selectivity at about this age on the Central/Incidental Task.44,45 This short-term memory task is designed to measure skill in attending selectively to task-relevant (central) information while inhibiting response to task-irrelevant (incidental) information. A measure of processing efficiency is derived by dividing the number of correct responses for task-relevant information (central) by the total number of responses (central + incidental). Efficiency of processing thus captures the relative recall of central and incidental information and reflects selectivity. Costa Rican children who had good iron status in infancy showed an increase in selectivity with age, like that observed in US children, but formerly iron-deficient children did not. Specifically, older children who had good iron status as infants showed improved selective recall of central (task-relevant) information compared with children who were younger (43.9% ± 2.2% SE vs 39.3% ± 2.0% SE). No such shift was observed in the formerly iron-deficient group. Formerly iron-deficient older children (>12.5 years old) did not do better in selective recall than younger children in the iron-deficient group (32.2% ± 5.0% SE vs 40.4% ± 2.7% SE), and they showed less selective recall than older children in the good-iron-status group, a significant iron status × age group interaction (F(1,155) = 4.03; P < .05). The results were the same with and without covariates.

The other statistically significant difference in specific cognitive functions was on a subtest of the Cognitive Abilities Tests.46 Formerly iron-deficient children were slower on the Tachistoscopic Threshold, which assesses perceptual apprehension or the minimum time required to determine whether 2 stimuli are the same or different. The median threshold (in milliseconds) for iron-deficient children was longer (208 ± 19 SE vs 158 ± 12 SE; F(5,141) = 4.91; P = .03). This difference still showed a suggestive trend after control for gender, age, maternal IQ, and HOME scores (202 ± 20 SE vs 160 ± 12 SE; F(5,141) = 3.06; P = .08).

Scholastic Achievement

All children were enrolled in school. However, a greater proportion of the formerly iron-deficient group had repeated a grade (26% vs 12% of the good-iron-status group; χ² = 4.33; P = .04). The prevalence of grade repetition in the good-iron-status group is similar to that in Costa Rican school children as a whole.47 More of the iron-deficient group had been referred for special services or tutoring (21% vs 7%; χ² = 5.81; P = .02). There were no differences in the proportion currently in remedial or special education classes.

Behavior Problems

The CBCL and TRF data were analyzed as categorical variables, comparing the proportion of children in the clinical range on the various scales, and as continuous variables, comparing mean T scores. Parents of children in the iron-deficient group reported more concerns in several areas, often in the clinical range (Table 4). A significantly greater proportion of the formerly iron-deficient group was...
above the US clinical cutoff (T score >70) for Somatic Complaints, Anxious/Depressed, Thought Problems, Delinquent Behavior, and Aggressive Behavior. Suggestive differences were observed for Social Problems (P < .10) and Attention Problems (P = .06). With regard to CBCL summary measures, a greater proportion of the formerly iron-deficient group was above the US clinical cutoff (T score >63) for Internalizing Problems. Because these cutoffs are derived from the distributions of T scores in US standardization samples, we also determined cutoffs based on the Costa Rica distributions for any scale for which the US clinical cutoff placed an excessively large proportion (≥10%) of the good-iron-status group (our reference) in the clinical range. This was the case for Somatic Complaints and for the 3 summary measures (Internalizing, Externalizing, and Total Problems). With a Costa Rica cutoff of 2 SD above the mean T score for the good-iron-status group, a greater proportion of the iron-deficient group was in the clinical range for each of the summary measures.

In analyses of CBCL data as continuous variables, differences in the mean T scores were statistically significant for Anxious/Depressed, Social Problems, Thought Problems, Attention Problems, and Delinquent Behavior and showed a suggestive trend for Somatic Complaints (P = .08). On CBCL summary measures, formerly iron-deficient children had significantly higher T scores for Internalizing Problems, Externalizing Problems, and Total Problems. After control for covariates (age, gender, HOME scale, maternal IQ, pubertal development, and height-for-age percentile), all of these differences remained statistically significant or suggestive with the exception of

### TABLE 3. Overall Mental and Motor Functioning at 11 to 14 Years of Age*

<table>
<thead>
<tr>
<th>Test</th>
<th>Severe, Chronic Iron Deficiency in Infancy (n = 48)</th>
<th>Good Iron Status in Infancy (n = 114)</th>
<th>Significant Background Factors</th>
</tr>
</thead>
<tbody>
<tr>
<td>Wechsler Intelligence Scale for Children-Revised</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Verbal IQ</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Unadjusted‡</td>
<td>99.5 ± 2.1</td>
<td>105.5 ± 1.4</td>
<td>Gender, mother’s IQ, HOME</td>
</tr>
<tr>
<td>Adjusted</td>
<td>101.8 ± 2.0</td>
<td>104.6 ± 1.3</td>
<td></td>
</tr>
<tr>
<td>Performance IQ</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Unadjusted‡</td>
<td>97.7 ± 2.2</td>
<td>100.2 ± 1.4</td>
<td>Gender, HOME</td>
</tr>
<tr>
<td>Adjusted</td>
<td>99.1 ± 2.1</td>
<td>99.7 ± 1.4</td>
<td></td>
</tr>
<tr>
<td>Full-Scale IQ</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Unadjusted‡</td>
<td>98.4 ± 2.1</td>
<td>103.2 ± 1.3</td>
<td>Gender, mother’s IQ, HOME</td>
</tr>
<tr>
<td>Adjusted</td>
<td>100.4 ± 1.9</td>
<td>102.3 ± 1.2</td>
<td></td>
</tr>
<tr>
<td>Wide Range Achievement Test-Revised</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Arithmetic</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Unadjusted‡</td>
<td>86.9 ± 2.2</td>
<td>96.5 ± 1.4</td>
<td>HOME</td>
</tr>
<tr>
<td>Adjusted</td>
<td>88.8 ± 2.2</td>
<td>95.7 ± 1.4</td>
<td></td>
</tr>
<tr>
<td>Reading</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Unadjusted‡</td>
<td>120.1 ± 2.3</td>
<td>127.6 ± 1.5</td>
<td></td>
</tr>
<tr>
<td>Adjusted‡</td>
<td>121.6 ± 2.4</td>
<td>126.9 ± 1.5</td>
<td></td>
</tr>
<tr>
<td>Directed Writing Task</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Unadjusted‡</td>
<td>91.7 ± 1.9</td>
<td>99.2 ± 1.2</td>
<td>HOME</td>
</tr>
<tr>
<td>Adjusted‡</td>
<td>93.2 ± 1.9</td>
<td>98.6 ± 1.2</td>
<td></td>
</tr>
<tr>
<td>Bender Visual–Motor Gestalt Test</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Unadjusted‡</td>
<td>2.1 ± .3</td>
<td>2.2 ± .2</td>
<td>HOME</td>
</tr>
<tr>
<td>Adjusted‡</td>
<td>2.0 ± .3</td>
<td>2.2 ± .2</td>
<td></td>
</tr>
<tr>
<td>Bruininks-Oseretsky Test of Motor Proficiency, Short Form</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Unadjusted‡</td>
<td>44.0 ± 1.8</td>
<td>47.4 ± 1.1</td>
<td></td>
</tr>
<tr>
<td>Adjusted‡</td>
<td>42.4 ± 1.8</td>
<td>48.0 ± 1.1</td>
<td>Gender</td>
</tr>
</tbody>
</table>

*Values are means ± SE with and without adjustment for background factors. Standard scores take age into account; age is included as a covariate for Bender raw scores. Adjusted means are derived from analysis of covariance, controlling for gender, HOME score, and mother’s IQ. A cumulative HOME index summed the scores obtained in infancy, school-age, and early adolescence, which were highly intercorrelated (r values >.70), and all related to test scores at 11 to 14 years. Mother’s IQ was the covariate, rather than mother’s education, because mother’s IQ generally showed higher correlations with adolescent outcome; the regression coefficient of maternal education on mother’s IQ was used to estimate IQ if the mother had not been tested (32 cases). Tests of statistical significance are based on analysis of variance or covariance.

† P < .05.
‡ P < .01.
§ P < .001.
∥ Suggestive trend; P < .10.

Fig 1. Standard test score differences at 11 to 14 years old. Effect size is the difference between group means divided by the SD of the reference group. The good-iron-status group is the reference in these analyses. Results are shown with and without adjustment for age, gender, mother’s IQ, and HOME score. Levels of statistical significance are given in Table 3.
CBCL Summary Measures

<table>
<thead>
<tr>
<th></th>
<th>Above US Clinical Cutoff</th>
<th>Above Study Reference Cutoff</th>
<th>T Score Unadjusted</th>
<th>T Score Adjusted</th>
<th>Significant Background Factors</th>
</tr>
</thead>
<tbody>
<tr>
<td>Internalizing Problems</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Iron-deficient</td>
<td>61%</td>
<td>7%</td>
<td>65.8 ± 1.5</td>
<td>65.1 ± 1.7</td>
<td>Pubertal development, HOME</td>
</tr>
<tr>
<td>Good-iron-status</td>
<td>42%†</td>
<td>0%†</td>
<td>60.7 ± 1.0†</td>
<td>61.2 ± 1.0†</td>
<td></td>
</tr>
<tr>
<td>Externalizing Problems</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Iron-deficient</td>
<td>37%</td>
<td>11%</td>
<td>60.3 ± 1.7</td>
<td>60.1 ± 1.7</td>
<td></td>
</tr>
<tr>
<td>Good-iron-status</td>
<td>27%</td>
<td>2%†</td>
<td>56.5 ± 1.0†</td>
<td>56.9 ± 1.0</td>
<td></td>
</tr>
<tr>
<td>Total Problems</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Iron-deficient</td>
<td>56%</td>
<td>11%</td>
<td>65.2 ± 1.6</td>
<td>65.0 ± 1.6</td>
<td>Pubertal development, HOME</td>
</tr>
<tr>
<td>Good-iron-status</td>
<td>43%</td>
<td>3%†</td>
<td>60.3 ± 0.9†</td>
<td>60.7 ± 1.0†</td>
<td></td>
</tr>
</tbody>
</table>

Individual CBCL Scales

Withdrawn

|                      |                         |                             |                    |                 |                               |
| Iron-deficient       | 9%                      |                             | 61.1 ± 1.2         | 60.0 ± 1.4      | HOME                          |
| Good-iron-status     | 9%                      |                             | 59.1 ± 8           | 59.5 ± 8       |                               |

Somatic Complaints

|                      |                         |                             |                    |                 |                               |
| Iron-deficient       | 41%†                    | 4%                          | 66.2 ± 1.7         | 66.2 ± 1.7      |                               |
| Good-iron-status     | 23%†                    | 5%                          | 63.0 ± 0.9         | 63.1 ± 1.0      |                               |

Anxious/Depressed

|                      |                         |                             |                    |                 |                               |
| Iron-deficient       | 25%                     |                             | 64.1 ± 1.7         | 64.0 ± 1.6      | Pubertal development          |
| Good-iron-status     | 10%‡                    |                             | 59.2 ± 0.9         | 59.6 ± 0.9      |                               |

Social Problems

|                      |                         |                             |                    |                 |                               |
| Iron-deficient       | 20%                     |                             | 63.8 ± 1.5         | 64.1 ± 1.5      |                               |
| Good-iron-status     | 9%‡                     |                             | 60.0 ± 8           | 60.1 ± 9†      |                               |

Thought Problems

|                      |                         |                             |                    |                 |                               |
| Iron-deficient       | 13%                     |                             | 60.3 ± 1.5         | 59.9 ± 1.4      |                               |
| Good-iron-status     | 1%‡                     |                             | 56.6 ± 0.7         | 57.0 ± 0.8      |                               |

Attention Problems

|                      |                         |                             |                    |                 |                               |
| Iron-deficient       | 15%                     |                             | 62.2 ± 1.6         | 62.0 ± 1.4      | Pubertal development          |
| Good-iron-status     | 5%‡                     |                             | 57.9 ± 0.7         | 58.3 ± 0.8      |                               |

Delinquent Behavior

|                      |                         |                             |                    |                 |                               |
| Iron-deficient       | 17%                     |                             | 60.8 ± 1.4         | 60.2 ± 1.2      | HOME, mother’s IQ             |
| Good-iron-status     | 2%§                     |                             | 57.0 ± 0.6         | 57.4 ± 0.7      |                               |

Aggressive Behavior

|                      |                         |                             |                    |                 |                               |
| Iron-deficient       | 20%                     |                             | 61.4 ± 1.6         | 61.8 ± 1.5      | Pubertal development, HOME, age|
| Good-iron-status     | 7%‡                     |                             | 58.4 ± 8           | 58.8 ± 9†      |                               |

* T score values are means ± standard error with and without adjustment for background factors. Adjusted means are derived from analysis of covariance, controlling for age, gender, HOME score, mother’s IQ, pubertal development, and height-for-age percentile. The degree of pubic hair development was used in these analyses. This measure could be applied to the entire sample, because it was highly correlated with axillary hair for both boys and girls, breast development and menstruation for the girls, and testicular development for the boys (r values ranging from .73–.88). Because height- and weight-for-age percentiles were closely correlated (r = .67), height-for-age percentile was chosen as the covariate. Tests of statistical significance are based on analysis of variance or covariance for T scores and χ² test or Fisher’s exact test for categorical analyses.

† P < .05.
‡ P < .01.
§ P < .001.
|| P < .10 (suggestive trend).

Somatic Complaints and Externalizing Problems. Aggressive Behavior indicated a suggestive trend (P = .09). Figure 2A shows the behavior problem profiles derived from CBCL T scores adjusted for covariates. Significant differences were similar in additional covariate analyses including maternal depression (CESD), except for Attention Problems, Delinquent Behavior, and Thought Problems, which became suggestive trends (P = .06 – .07).

TRFs were available for 143 children. A higher proportion of the formerly iron-deficient group was above the US clinical cutoff on the Withdrawn subscale (15% vs 4% in the good-iron-status group; Fisher’s exact test; P = .03). In analyses of T scores without adjustment for child characteristics, there were no statistically significant differences. However, when the effects of age, gender, height-for-age percentile, and pubertal progression were considered in analyses of covariance, teachers reported more concerns about the behavior of formerly iron-deficient children in several areas, much like parental observations. Specifically, adjusted T scores were significantly higher in the iron-deficient group for Withdrawn, Anxiety/Depression, Social Problems, Attention Problems, Internalizing Problems, and Total Problems and showed a suggestive trend for Delinquent Behavior (P = .06). Figure 2B shows the behavior problem profiles based on TRF data with covariate adjustment.

No statistically significant differences were observed in ratings of the children’s behavior during the physical examination and blood sampling or the socially stressful situation.

DISCUSSION

In this study, we reevaluated a group of 11- to 14-year-old Costa Rican children who had been tested and treated for iron deficiency as infants. The results confirmed most, but not all, of our hypotheses. The children who had severe, chronic iron defi-
ciency in infancy scored lower on measures of overall mental and motor functioning. More of them had repeated a grade and/or been referred for special services or tutoring. They showed a delay or disruption in a shift in cognitive processing expected in early adolescence. However, a differential effect on visual–motor and motor skills was not clearly observed, although the iron-deficient group did have difficulty on a motor test and on tasks involving visual–spatial memory and selective recall for visual stimuli. Parents and teachers had more concerns about the behavior of formerly iron-deficient children in several areas. Both parents and teachers reported that the iron-deficient group showed more anxiety/depression, social problems, and attention problems, with a corresponding increase in the summary measures of internalizing problems and total problems.

It is not obvious how an early biologic insult like iron deficiency could produce such long-lasting cognitive, motor, and emotional differences. The adaptability of the human infant and the plasticity of the brain might mitigate against such outcomes. Long-lasting differences are perhaps especially challenging to understand in the case of early nutritional deficiencies. These insults do not involve neurotoxins or brain lesions, and peripheral manifestations of the nutrient deficiency can generally be treated effectively.

There are several possible explanations for the long-lasting differences we observed. The differences might not be attributable to iron deficiency at all but to some associated factor, such as other unidentified nutrient deficiencies, parental intellectual and emotional limitations, or environmental disadvantage. Although we found no evidence of other nutritional deficiencies, there is little doubt that iron deficiency affects disadvantaged infants disproportionately (see reviews). In this study, we conducted a relatively comprehensive assessment of home and family and also found a number of disadvantages among the group with iron deficiency in

---

**Fig 2. Behavior problem profiles. Values are mean T scores ± SE adjusted for background factors. Panel A compares the iron-deficient and good-iron-status groups with respect to CBCL (parent report) T scores adjusted for age, gender, mother’s IQ, HOME score, pubertal development, and height-for-age percentile. Panel B shows TRF T scores adjusted for age, gender, pubertal development, and height-for-age percentile. The statistical significance of differences between the study groups was determined in analyses of covariance. ∗P < .05; †P < .10.**

---

http://www.pediatrics.org/cgi/content/full/105/4/e51

Downloaded from by guest on April 7, 2017
infancy. Statistical control for differences in these factors did not eliminate the significant effect of chronic and severe iron deficiency on early adolescent behavior and development. Nonetheless, unmeasured differences in the stimulation and care the formerly iron-deficient children received could account for their poorer outcomes.

An alternate explanation is that nervous system effects of iron deficiency during brain development account for later differences in behavior and development. Although animal studies show that dietary iron deficiency during maximal brain growth leads to a deficit in brain iron that is not reversed with treatment, we have concluded in the past that brain–behavior connections were too poorly understood in the human to attribute poorer outcome to central nervous system effects of iron deficiency. However, evidence for iron’s role in the developing brain has steadily accumulated, inviting cautious interpretation of some of our findings. In discussing the results, we will first relate them to previous studies and then speculate about the connections to recent basic research on iron and the brain.

Relation to Previous Studies

Follow-up studies in Israel, Chile, and France found that 4- to 8-year-old children who had been anemic as infants or toddlers tested lower than peers several years after iron treatment. Two projects have followed children at older ages. In Chile, Rivera and Walter reevaluated a subset of their original infant cohort at 10 years of age, and preliminary analyses show that the formerly anemic group had poorer school functioning and lower achievement test scores. In a population-based study of Women, Infants, and Children (WIC) participants in Florida, Hurtado and associates recently reported an inverse relationship between hemoglobin level in infancy and the risk of mild/moderate mental retardation at 10 years of age. We were unable to find data from other studies on grade repetition or specific cognitive functions—some of our major findings. Nonetheless, the consistency of results from studies in 5 different countries is striking. The findings show that anemia presumably attributable to iron deficiency or severe, chronic iron deficiency in infancy identifies children with poorer overall cognitive functioning and lower school achievement test scores years later.

There is little information on behavior problems in other follow-up studies of children who had iron deficiency and/or anemia as infants. However, the increase in behavior problems was marked in our project. Whether the behavior problems we observed were a response to cognitive limitations and accompanying frustrations or direct effects of the infancy condition is unknown. Observations of affective changes during the period of iron deficiency in infancy (wariness, unhappiness, and hesitance in an unfamiliar setting) suggest the latter possibility. Behavioral differences that begin in infancy could contribute to poorer developmental outcome later on. This line of reasoning is based on a transactional model of development. That is, attributes of the infant affect caregivers, whose behavior in turn influences the infant, etc. Such a behavior pattern suggests that infants with chronic and severe iron deficiency might be functionally isolated, seeking and/or receiving less stimulation from the physical and social environment, with adverse effects on development (see review). Behavior problems after the period of deficiency, like those observed in this follow-up, might continue to interfere with the children’s ability to learn from school and environment. Thus, behavioral differences could contribute to lower achievement test scores and other outcomes that depend on formal and informal learning.

Speculation

New basic science research may help make sense of some of our findings. For instance, recent rodent studies document a differential impact of iron deficiency on regions of the brain that are involved with memory, particularly the hippocampal formation. In addition, rats who were iron-deficient in early development showed lasting difficulty with spatial navigation, a task that involves hippocampal function. There is also evidence that early nutrient deficiencies in humans affect memory functions at later ages or developmental periods. For instance, 8-month-old infants thought to have prenatal iron deficiency had impaired recognition memory, a function that seems to depend on the hippocampus, and Indonesian children who received a nutritional supplement that included iron before 18 months of age showed improved memory at 8 years of age.

Older research also pointed to iron’s role in central nervous system neurotransmitter function, especially involving the dopaminergic system. Studies conducted in the 1970s implicated the D2 receptor (see reviews). After nearly 20 years of little or no new work, there now are modern studies in this area, confirming dopaminergic alterations in iron deficiency, particularly in the striatum (caudate/putamen). One neural system in which dopamine plays a particularly prominent role is the prefrontal–striatal system and recent advances in cognitive neuroscience have identified behavioral tasks that specifically involve the prefrontal–striatal and hippocampal systems.

Of particular interest in this follow-up are the findings that the formerly iron-deficient group showed poorer visual–spatial/working memory and a delay in developing the ability to attend selectively and inhibit attention to the irrelevant. These results seem to fit with new understanding of iron’s role in prefrontal–striatal and hippocampal systems. However, because of the necessity of inferring the neural bases of the particular tasks we used, our results are best used to generate hypotheses for systematic investigation in future studies.

The poorer motor test scores of formerly iron-deficient children may also be considered in light of new understanding of iron and the brain. Although it has been known for some time that high iron concentrations are found in brain regions involved with motor function and coordination (basal ganglia, cerebellum, etc), iron’s essential role in myelin for-
mation and maintenance has been described only more recently.72–76 Perhaps the most direct evidence for an effect in human infants that could be attributable to impaired myelination comes from a recent study showing that young infants with iron deficiency anemia had slower nerve conduction through the auditory pathway, even after effective iron therapy.77 However, explaining how delayed myelination could produce long-lasting motor changes remains a challenge. Furthermore, in contrast to plausible central nervous system mechanisms to account for the motor and cognitive differences we observed, there is little basis for hypothesizing the neural underpinnings of the increase in behavior problems in formerly iron-deficient children.

Summary and Implications

The explanations for long-lasting ill effects of early iron deficiency considered above are not mutually exclusive. Different mechanisms may be involved, perhaps depending on the domain in question. For instance, a deficit in visual–spatial/working memory might be a direct result of lower iron in certain brain regions, and long-lasting motor differences might result from delayed myelination. Indirect effects, which might still begin with central nervous system changes, could also be postulated. For instance, lowered brain iron, altered neurotransmitter function, or impaired myelination during infancy could disrupt the process of laying down the neural bases for some cognitive, socioemotional, and motor fundamentals. This disruption could get iron-deficient children onto developmental trajectories that differ from their peers with better iron status. Or nonspecific behavioral or cognitive differences in infancy might produce long-lasting effects through their impact on caregivers in a transactional fashion. In contrast, poorer school achievement and grade repetition might be influenced heavily by family disadvantages and limitations. It seems likely that several of the above mechanisms combine to produce poorer outcome in the various domains.18

Based on experience in the United States, differences in achievement test scores of the magnitude we observed and the increase in grade repetition are likely to affect the educational and career paths of these children. Similarly, their high behavior problem scores are likely to correlate with poorer mental health and social functioning later on. Continued follow-up of this cohort will determine how these differences affect such outcomes as school drop out, occupational and educational trajectories after high school, mental illness, and impaired social relations.

In this study, the sample was carefully selected to exclude children with conditions in infancy that could adversely affect development, such as low birth weight or illness, and the children have been growing normally by US standards. Thus, the results are most generalizable to full-term, healthy infants who are free of generalized undernutrition. These conditions apply to many children in industrialized countries, but most infants in developing countries do not have as optimal growth and health as the sample here. The outcome of chronic, severe iron deficiency under less healthy conditions is as yet unknown, but it might be even more adverse.

CONCLUSION

In summary, the results show that children who had severe, chronic iron deficiency in infancy continue at behavioral and developmental disadvantage relative to peers >10 years after treatment. Despite limitations in our ability to attribute causality and to understand the underlying mechanisms, the research has important practical implications. Children with severe, chronic iron deficiency may require special intervention in infancy in addition to iron therapy. They may also benefit from additional interventions in the school years. Even more importantly, preventing iron deficiency might help foster the behavior and development of disadvantaged infants throughout the world.

ACKNOWLEDGMENTS

This work was supported by Grant R01 HD31606 from the National Institutes of Health.

We gratefully acknowledge the dedication of study participants and their families and the meticulous evaluations by project psychologists Marjorie Barquero and Willy Sommaribas and project pediatrician Nereida Arjona. Ernesto Pollitt encouraged us to puzzle about the explanation for long-lasting effects. Charles A. Nelson, Edward Smith, and John Jonides helped in considering the neural correlates of some specific cognitive functions. We appreciate additional assistance in statistical analyses, graphical presentation, and manuscript preparation from Stephen Peck, Gloria Wheatcroft, Maria Perez, and Gail Sonnett.

REFERENCES

15. Honig AS, Oski FA. S olemnity: a clinical risk index for iron deficient
infants. Early Child Dev Care. 1984;16(9):84.


Poorer Behavioral and Developmental Outcome More Than 10 Years After Treatment for Iron Deficiency in Infancy
Betsy Lozoff, Elias Jimenez, John Hagen, Eileen Mollen and Abraham W. Wolf

Pediatrics 2000;105;e51

The online version of this article, along with updated information and services, is located on the World Wide Web at:
/content/105/4/e51.full.html