Iron-Deficiency Anemia: Reexamining the Nature and Magnitude of the Public Health Problem

A Review of Studies on the Effect of Iron Deficiency on Cognitive Development in Children[1,2]

Sally Grantham-McGregor[3] and Cornelius Ani

Centre for International Child Health, Institute of Child Health, University College, London, UK

ABSTRACT   Studies on the effect of iron deficiency on children's cognition and behavior are selectively reviewed, looking for evidence of a causal relationship. Most correlational studies have found associations between iron-deficiency anemia and poor cognitive and motor development and behavioral problems. Longitudinal studies consistently indicate that children anemic in infancy continue to have poorer cognition, school achievement, and more behavior problems into middle childhood. However, the possible confounding effects of poor socioeconomic backgrounds prevent causal inferences from being made. In anemic children <2 y old, short-term trials of iron treatment have generally failed to benefit development. Most longer trials lacked randomized placebo groups and failed to produce benefits. Only one small randomized controlled trial (RCT) has shown clear benefits. It therefore remains uncertain whether the poor development of iron-deficient infants is due to poor social backgrounds or irreversible damage or is remediable with iron treatment. Similarly, the few preventive trials have had design problems or produced no or questionable benefits only. For children >2 y old, the evidence from RCT is reasonably convincing but not conclusive. RCT of iron treatment are warranted especially in younger children.   J. Nutr. 131: 649S–668S, 2001.

KEY WORDS: • iron deficiency • children • cognition • behavior • development

Over the past three decades, there have been a considerable number of studies on the relationship between iron status and cognition, and behavior, but the topic remains controversial (Logan 1999). Many professionals are lobbying to promote fortification or supplementation programs, claiming conclusive evidence of a causal relationship between iron deficiency and poor cognitive development, whereas others consider that there is no clear evidence (Morley et al. 1999). In this review, we examine studies in humans (mostly children), looking for evidence of a causal relationship between iron status and cognition and behavior. Throughout the review, we refer to iron-deficiency anemia as anemia.

DEMONSTRATING CAUSAL RELATIONSHIPS

To demonstrate a causal effect, several conditions have to be fulfilled. Iron deficiency alone has to be shown to cause a change in development. There should also be a biologically plausible mechanism linking iron deficiency to development.

Demonstrating significant associations between anemia and poor development in correlational or case-control studies is helpful in identifying at-risk populations but cannot establish cause-and-effect relationships. They provide no information as to the timing of any relationship and it is possible that poor development precedes iron deficiency. In addition, there is considerable evidence that anemia is associated with a large number of socioeconomic and biomedical disadvantages that can themselves affect children's development. Some of the factors found to be associated with both anemia and poor cognitive development are low socioeconomic status (Owen et al. 1971); poverty (Czajka-Narins et al. 1978); lack of stimulation in the home (de Andraca et al. 1990), including lack of maternal warmth; poor maternal education (de Andraca et al. 1990, Idjradinata and Pollitt 1993) and intelligence quotient (IQ)[4] (Lozoff et al. 1991); maternal depression (de Andraca et al. 1990); more absent fathers; low birth weight (<2.5 kg) and early weaning (Lozoff et al. 1991); parasitic infection (Ramdath et al. 1995); elevated blood lead levels; and undernutrition. It is highly unlikely that all of these factors are controlled for in one study, and there are probably many other confounding factors.

[4] Abbreviations: CNS, central nervous system; DBRCT, double-blind randomized controlled trials; IQ, intelligence quotient; MDI, mental development index; PDI, psychomotor development index; RCT, randomized controlled trials.

Longitudinal observational studies give additional useful information about the long-term prognosis of children with anemia and the types of deficits at different stages of development. However, they also cannot provide evidence of a causal relationship, but finding reasonably consistent associations between anemia and cognition—after controlling for the most obvious confounders—is a first step toward making causal inferences.

The most accurate way of pinpointing iron deficiency as a cause of poor development is to conduct a double-blind, randomized, controlled trial and demonstrate that producing or preventing anemia changes children's development. Obviously, one has to use animal models for producing iron deficiency, but preventive trials beginning with nonanemic children are possible to conduct. Unfortunately, preventive trials are extremely difficult and expensive to run. They need large samples to have adequate statistical power, even in populations in which the prevalence of anemia is high. The samples must be followed for some time and it is essential that they remain intact.

Randomized controlled therapeutic trials in which iron is given to anemic children can demonstrate whether a developmental deficit is remediable with iron treatment. They are equally rigorous and are easier to conduct than preventive trials because much smaller numbers are needed. Unlike preventive trials in which a substantial proportion of the placebo children are not anemic and not all treated children are expected to benefit, in therapeutic trials, it is reasonable to expect all treated children to benefit from iron. However, failure of response to treatment does not necessarily negate the presence of a causal relationship because it is possible that the developmental deficit is irremediable at least in the short term.

MECHANISMS

Several mechanisms linking anemia to altered cognition are possible. The most direct one is that changes that affect development occur to the structure and function of the central nervous system (CNS). There is substantial evidence of such changes from animal research; these studies are being examined in other papers in this supplement and will not be discussed here.

Evidence from children of changes to the CNS is limited. However, investigators recently studied auditory brain stem responses in children with anemia (Roncagliolo et al. 1998). These responses provide a measure of the activation of the auditory pathway from the distal part of the acoustic nerve to the lateral lemniscus, and the central conduction time is an indicator of CNS development. The central conduction time was found to be prolonged in 6-mo-old children ($n = 29$) with anemia compared with nonanemic children ($n = 26$). Furthermore, they did not improve with correction of anemia and the difference was greater 6 and 12 mo later. The investigators speculated that the prolonged central conduction time was due to changes in myelination that have been reported in iron-deficient animals (Yu et al. 1986). Recent work has shown that formerly anemic children also have longer latencies in visual evoked potentials (B. Lozoff, personal communication). None of the above studies controlled for social background and it is possible that deprivation could affect brain development.

Another hypothesis linking anemia to poor development is functional isolation, which was originally conceived to explain poor development in children with protein-energy malnutrition (Levitsky and Strupp 1995). Anemic children explore and move around their environment less than nonanemic children, and they induce less stimulating behavior in their

caretakers. These behaviors and the caretakers' response are thought to delay the acquisition of new skills.

There are many reports of clinical impressions of anemic children being fearful. More systematic observations have been made comparing anemic with nonanemic children during testing with the Bayley Scales using the Bayley infant behavior ratings. These studies have found that anemic children tend to be more fearful (Lozoff et al. 1982a and 1996), withdrawn, tense, unreactive to usual stimuli (Lozoff et al. 1982a), more solemn, less involved (Honig and Oski 1984) and more unhappy (Lozoff et al. 1996, Walter et al. 1983).

Surprisingly, there have been few observations in nontest situations. In a study in the United States (Johnson and McGowan 1983), anemic children were observed in a standard situation that included two set tasks and free play. They were not different from nonanemic children in activity, reactivity, emotional tone, or attention span. However the observation period was extremely short (total of 16 min) and unlikely to provide a representative sample of behavior. Also, the groups showed no significant difference between their scores on the Bayley Scales, which is unusual. In another observation study of a very short free-play session (Lozoff et al. 1986), children stayed closer to their mothers and this was attributed to both the mothers' and children's behavior.

In a more extensive study, behavior observations were made in a 15-min free-play situation and throughout developmental testing. Children with anemia stayed closer to their caretakers, showed less pleasure, and were more wary, hesitant and easily tired. An average of 14 spot observations were also made during home visits, and anemic children were more likely to be asleep, irritable, doing nothing, being carried or in bed, and less likely to be on the patio or playing interactively with objects. During the Bayley test session, the anemic children made fewer attempts at test tasks, were less playful and had poorer attention than nonanemic children (Lozoff et al. 1998). These types of behaviors persisted after treatment. Most interestingly, mothers of anemic children were rated as being less affectionate, and even the testers behaved differently with the children, giving them fewer tasks and making fewer attempts to elicit responses.

It is of course possible that these behaviors could be due to deprived environments. In Jamaica, similar behavior was found in undernourished children and the behavior was changed with stimulation alone, without changing the children's nutritional status (Grantham-McGregor et al. 1989).

One study linked children's behavior during assessment on the Bayley Scales to their test scores (Lozoff et al. 1985). The children with abnormal ratings were more likely to have lower Bayley scores. The investigators hypothesized that the anemic children's lower scores were mediated through behavior disturbances. There are therefore several biologically plausible ways, demonstrated in both animal and human research, in which iron deficiency could affect child development.

REVIEW OF STUDIES

Several recent comprehensive reviews exist (Lansdown and Wharton 1995, Lozoff 1998, Watkins and Pollitt 1998). In this review, we have chosen to consider critically selected important studies, focusing on evidence of causality. We discuss the studies grouped by study design and further divided by subjects' age. The categories are correlational and case-control studies, longitudinal observation studies, therapeutic treatment trials and preventive treatment trials. Within these categories, the internal validity of the studies is examined. The definition of iron-deficiency anemia has been problematical in the past; this

topic is discussed in another paper in this supplement and will not be dealt with here.

Correlational and case-control studies.

Beginning as early as 1919, many investigators found significant concurrent associations between hemoglobin concentrations and measures of cognitive development or school achievement (Agarwal et al. 1987, Clarke et al. 1991, Florencio 1988, Grindulis et al. 1986, Popkin and Lim-Ybanez 1982, Waite and Neilson 1919, Walker et al. 1998, Webb and Oski 1973). In addition, baseline differences in developmental levels, cognition or school achievements were found between nonanemic and anemic groups in treatment trials. For example, in trials concerning children <2 y old, in six of seven studies with nonanemic and anemic children (Idjradinata and Pollitt 1993, Lozoff et al. 1982b, 1987 and 1996, Walter et al. 1983 and 1989), the anemic groups had significantly lower scores on the mental development index (MDI) of the Bayley Scales. There were only eight nonanemic children in the seventh study (Driva et al. 1985). Four of the studies also showed differences in the psychomotor development index (PDI) (Idjradinata and Pollitt 1993, Lozoff et al. 1982b and 1987, Walter et al. 1989). Most of these studies had some control for social background and biomedical conditions, but few had extensive controls for both socioeconomic and biomedical conditions. Although most studies found associations between anemia and a developmental outcome, a puzzling minority of the studies failed to find significant associations (Deinard et al. 1981 and 1986, Huda et al. 1999, Johnson and McGowan 1983, Moock and Leslie 1986). Small sample sizes may explain some failures to find associations. Also, some of the studies did not have measures of iron status other than hemoglobin levels and it is possible, but not very likely, that iron deficiency was not the commonest cause of anemia.

Although correlational studies offer the opportunity to look for possible interactions between anemia and socioeconomic or biomedical conditions, few investigators have attempted to do so and most had sample sizes insufficient for doing so. Many studies of protein-energy malnutrition and low birth weight have shown that these conditions interact with social background and other biomedical conditions in their effect on child development (Grantham-McGregor et al. 1998 and 1999, Pollitt et al. 1993). It is likely that analogous relationships exist with anemia. For example, anemia may have different effects in low-birth-weight infants than in normal-birth-weight infants. Most investigators have gone to great lengths to exclude high risk infants; thus, information on these types of questions is scarce.

Some investigators have examined the relationships between severity of anemia and developmental decline. In one study (Lozoff et al. 1987), a decline in concurrent motor development was found at hemoglobin values <105 g/L, whereas a decline in mental development appeared at values <100 g/L. In contrast, Walter and colleagues (1989) compared the development of children with hemoglobin concentrations <100 g/L with those with concentrations between 105 and 109 g/L and >109 g/L. The three groups were significantly different from each other in both their motor and mental developmental indices, which were in the same ranking order as their hemoglobin concentrations. Therefore it appears that the level of anemia associated with declining development varies in different populations

Longitudinal observation studies

We identified seven studies in which hemoglobin levels in early childhood were linked to cognitive development or school achievement in later childhood (Cantwell 1974, de Andraca et al. 1990, Dommergues et al. 1989, Hurtado et al. 1999, Lozoff et al. 1991 and 2000, Palti et al. 1983 and 1985, Wasserman et al. 1992 and 1994). The ages and size of the samples, outcome measures and findings are given in **Table 1**.

Samples. The initial age was generally <2 y except for one study (Hurtado et al. 1999), which involved records of children who entered a nutrition program between birth and 5 y. The oldest children to be followed up were between 11 and 14 y old, whereas the youngest were 4 y old. The sample sizes tended to be small, ranging from 20 to 41 anemic children, except for the study of public records (Hurtado et al. 1999). The definition of the original episode of anemia varied and iron status was not clear in most studies. Three studies comprised one cohort, and hemoglobin levels as continuous variables (regardless of iron status) were related to later cognitive development or school achievement outcomes (Hurtado et al. 1999, Palti et al. 1983 and 1985, Wasserman et al. 1992 and 1994). In the other four studies, the criterion for anemia was not clear in one study (de Andraca et al. 1990); it was <105 g/L in another (Lozoff et al. 1991), <110 g/L in the third (Dommergues et al. 1989), and between 61 and 95 g/L in the fourth (Cantwell 1974).

Most importantly, all studies found that formerly anemic children continued to be at a developmental disadvantage at one or more of the follow-up assessments. All but one study (Cantwell 1974) reported controlling for some social background variables, gender and birth weight. Although the size and number of differences were reduced when scores were adjusted, some tests remained significant in all but one study. In that study (Wasserman et al. 1994), the final examination was at 4 y, and previous hemoglobin levels were only inconsistently related to IQ (negative at one age and positive at another). However, early hemoglobin levels had been related to development at 24 mo (Wasserman et al. 1992).

Specific functions affected. All but one study (Hurtado et al. 1999) had a global measure of development, either an infant developmental assessment or an IQ test, and these were poorer in anemic children. Specific cognitive functions were assessed in only two studies. In Costa Rica and Chile (de Andraca et al. 1990, Lozoff et al. 1991 and 2000), children were given a comprehensive battery of tests at ~5 y of age. In both studies, the formerly anemic children had deficits, which were not identical, across a wide range of functions. Preschool skills, fine and gross motor skills, and visual-motor integration were affected in both studies, whereas language and global IQ were affected in the Chilean sample (de Andraca et al. 1990) and only performance IQ in Costa Rica (Lozoff et al. 1991). Children in Costa Rica were reassessed between ages 11 and 14 y for an even wider range of functions (Lozoff et al. 2000). The anemic children's performance was worse in practically all tested functions, but they came from more deprived environments than did the nonanemic children. After many covariates were controlled for, the differences were reduced to writing, reading and arithmetic; motor skills; spatial memory; and, in the older children only, selective attention. The children's behavior was also assessed by teacher and parent reports. The anemic children were reported to have higher scores in anxiety and depression, social and attentional problems, and total problems after covariates were controlled for. The only other study reporting behavior (Cantwell 1974) reported that ane-

TABLE 1

Longitudinal observational studies[1]

Author, country, date	Sample n	Period of follow-up	Exclusions	Study design	Outcome measures	Covariates	Dropout	Findings	Remarks
Cantwell 1974 USA	61 full-term neonates from comparable socioeconomic groups: 29 given IM iron in neonatal period 32 infants developed Fe deficiency anemia between 6–18 mo (Hb 6.1–9.5 g%) without PEM. No details of iron status	Birth to 7 y	Preterm	29 of 61 infants received Fe injections (method of assignment not given) and were not anemic (Hb 11.5–12.9) 32 infants developed Fe-deficiency anemia Examined at 6–7 y by examiners blind to the groups.	Neurological examination and Stanford Binet IQ.	None reported	Not reported	The formerly anemic group had a higher incidence of "soft signs," e.g., clumsiness with balancing on one foot, in tandem walking, and repetitive hand and foot movement and were more inattentive and hyperactive than the non-anemic group. IQ scores averaged 98 in the nonanemic and 92 in the anemic. No significance levels reported	Details not reported No covariate control or statistics reported
Palti et al. 1983 Israel	Routine health service screen for Hb at 9 mo Tested at 2 y = 873 At 3 y = 388 At 5 y = 239 Hb only measure of iron status	9–10 mo to 5 y	Not given	Follow-up of all children from 9–10 mo to 2, 3 and 5 y. All with Hb < 110 g/L treated with iron at 9 mo for 3 mo At 5 y took a random sample of remaining children	Brunet Lezine test at 2 y. MILI test, an Israeli intelligence test at 3 y Wechsler with Israeli adaptation (WPSSI) at 5 y	Covariates: maternal education, father's occupation, BW, sex.	485 from 2–3 y. 149 from 3–5 y.	When controlling for covariates: Hb at 9 mo not significantly associated with DQ at 2 y ($P = 0.105$) & at 3 y ($P = 0.07$) but at 5 y had a significant effect on IQ ($P = 0.02$) At 5 y an increase of 10 g/L of Hb associated with 1.75 change in IQ points	At all ages mothers' education had the most significant effect on DQ
Palti et al. 1985 same as above	Attending local school IDA at 9 mo (Hb < 105 g/L) = 20 Non-A at 9 mo (Hb > 115 g/L) = 56 No other iron status data	To second grade	Not attending local school	Two groups selected for followed up in second grade	Teachers ratings on learning in reading, writing, spelling, arithmetic, general knowledge, positive and negative task orientation, mood	Covariates: maternal education IQ, age, birth order, sex, HOME, birth weight, lead levels		Controlling for covariates: IDA group significantly poorer in learning ($P = 0.04$) and + ve task orientation, − ve task orientation and mood not significant	Very selected population
Dommergues et al. 1989, France	147 children in 2 well-baby clinics, IDA = 22% at 10 mo, 9% at 2 y, 2% at 4 y, IDA = Hb < 110 g/L	10 mo to 4 y	Not given	Follow-up at 10 mo, 2 and 4 y	Brunet-Lezine test	Parental education, parity, ethnicity		10 mo, no association with iron status 24 mo, MCHC significantly related to DQ, motor and social quotient 48 mo, MCHC at 2 y significantly related to DQ at 4 y	
De Andraca et al. 1990 Chile Walter 1993	Total = 77 Formerly anemic = 41. Formerly non-A = 29 All anemic treated at 12 mo Criteria not given, presumably the same as {Walter 1989}	Birth to 5–6 y	BW < 2500 g, chronic ill health, intermediate levels of anemia. First graders (unequal psychological evaluation) Incomplete data.	Part of a randomized trial of iron fortification in early infancy, at 1 y, 25% of the nonfortified group had anemia. The anemic children all received 3 mo of Fe treatment Selected children re-examined at 5–6 y of age	Stanford Binet IQ, Illinois Psycholinguistic Abilities Test, Psychoeducational Abilities Test, Bruininks-Osteretsky Test of Motor Function, Visual motor integration (VMI), neurological exam	Covariates: HOME, maternal depression and stress. Not clear if used in analysis	7 formal exclusions, ?others	Hb at 1 y = 10.1 ± 0.7 vs. 13.0 ± 0.8. Hb a 15 mo = 12.8 ± 0.7 vs. 13.0 ± 0.8. Current Hb level not given. Formerly anemic children performed significantly worse in IQ ($P = 0.02$), psychoeducational abilities ($P < 0.01$), VMI ($P < 0.01$), motor proficiency ($P < 0.01$), language abilities ($P < 0.01$). They were more neurologically immature $P < 0.01$. Their homes were sig. less stimulating and their mothers were more depressed and less affectionate. Walter 1991: controlling for 'neurological maturity, maternal education, HOME, breast feeding, nursery attendance, height/age etc', psychoeducational abilities, language abilities, fine motor abilities remained significantly different.	Hb at 5 y not given.
Lozoff et al. 1991, Costa Rica	163 of 191 children originally evaluated at 12–24 mo. 30 had moderate anemia = Hb ≤ 100 g/L, Ferritin ≤ 12 mcg, EP > 1.77 mcgmol or transferrin ≤ 10%. 133 comparison group	12 mo to 5 y	BW < 2.5 kg, multiple pregnancy, complicated births, acute or chronic medical problem	Follow-up at 5 y of Lozoff et al. 1987. The IDA group was initially treated for 3 mo to correct their anemia. Current evaluators blind to original iron status All now free of anemia 2nd analysis	Current iron status. WIPPSI, Woodcock Johnson psychoeducational battery, Goodenough-Harris Draw—a-man test, Beery Developmental test of Visual-Motor Integration, (VMI) Bruininks-Oseretsky Test of Motor Proficiency	Covariates: sex, birth weight, mothers IQ, height & education, breastfeeding, absence of father, HOME	28	No current difference in Hb and other measures of iron status. After controlling for covariates previously anemic gp had lower scores on performance IQ, quantitative and visual matching subtests of the Woodcock Johnson battery, the VMI and the Bruininks-Oseretsky motor test. In post hoc analyses, children who were non-anemic but continued to have iron deficiency after treatment also had significantly lower scores.	Good covariate control Verbal skills less affected

TABLE 1 (continued)

Longitudinal observational studies[1]

Author, country, date	Sample n	Period of follow-up	Exclusions	Study design	Outcome measures	Covariates	Dropout	Findings	Remarks
Lozoff et al. 2000, Costa Rica same as above	167 (87%) of original 191 infants Anemic group = 48 (moderate anemia as above + those who were non-anemic but Fe deficient before and after treatment) 114 non-A with good iron status before and/or after treatment	Retest at 11 to 14 y	Same as above	All now free of anemia 2nd assessment at 11–14 y, using new classification for anemic group	Current iron status. WISC, Tests of: Wide Range Achievement, Bender Gestalt visual motor, Bruininks-Oseretsky motor proficiency, Central/incidental learning, Attention capacity Underlining, K-ABC Spatial memory, Tactual performance, computerised cognitive abilities: learning, reaction time, stimulus descrimination, Sternberg search, tachistoscopic threshold, self-paced probe recall parent and teacher behavior check list,	Covariates: sex, mothers IQ, HOME	13%	After control for covariates: consistent tendency for anemic group to have lower scores in most tests but had significantly lower scores in writing and arithmetic, motor test, spatial memory, poorer selective attention to central task in older children only. Also anemic group had a consistent tendency for more behaviour problems both internalizing and externalizing. After covariate control: anemic group was significantly more anxious/depressed, had more attention problems, social problems and total problems More repeated grades and had been referred for special service (not clear if covariates controlled)	A large number of cognitive tests
Wasserman et al. 1992, Kosovo	Children whose mothers were followed up from pregnancy in two areas of Kosovo Mitrovica = lead exposed Pristina = nonlead exposed 541 agreed to participate 392 (208 + 184) seen at 24 mo	Birth to 24 mo	Major CNS defects, chromosomal abnormalities, multiple pregnancy	Follow-up two cohorts from birth measuring serum lead, iron status, and developmental indices. Related Hb at each age with DQ at 24 mo Anemic children treated	Bayley MDI at 6, 12, 18 and 24 mo with iron and lead status	Covariates: ethnic gp, HOME, birth order, BW, sex, maternal IQ, education and age, lead levels	149 of 541 lost at 24 mo	Hb at 6, 12 and 24 mo was not significantly associated with MDI at 24 mo but Hb at 18 mo was significant. Controlling for all covariates, in both Mitrovica and Pristina a change in Hb at 18 mo of 20 g/L was associated with a change of 3.4 MDI points (P = 0.02). Other indices of iron deficiency not associated with development.	Age probably confounded with severity of anemia. Suggests an independent effect of anemia at 18 mo.
Wasserman et al. 1994 Kosovo same as above	388 at 3 y 332 at 4 y To 4 y	same	Same as above	MSCA a test of intellectual function	Covariates: ethnic gp, HOME, birth order, BW, sex, maternal IQ, education, and age, lead levels	209 lost of 541 at 4 y	Controlling for all covariates inconsistent associations between IQ at 4 y and previous Hb measures: 6 mo Hb −ve effect (B = −1.771, P = 0.03), and 36 mo Hb (B = 1.478, P = 0.03)	The effect from 18 mo appears to have been transient. Good covariate control
Hurtado et al. 1999, USA	5411 born in 1979–1980 in Dade County, Florida, who participated in WIC program between birth and 5 y. Final sample = 3771 after exclusions Hb only measure of iron status	Birth to 10 y	736 children with physical or sensory handicaps. 60 for whom maternal educational data were missing. 848 for missing Hb.	Retrospective study linking population birth records to school records at 10 y, then linking to WIC records of entrants from birth to age 5 y Logistic regression to estimate association between anemia and mental retardation controlling for covariates	Special education placement, based on criteria used by Florida Department of Education for mild or moderate mental retardation	Covariates: birth weight, maternal education, sex, race-ethnicity, age, and age of child at entry into WIC		Risk of placement in special education increased by 1.28 for each decrement of Hb even after controlling for all other variables in the equation.	Measure of Fe status was Hb only.

[1] Abbreviations: IM, intramuscular; IQ, intelligence quotient; Hb, hemoglobin; PEM, protein-energy malnutrition; DQ, developmental quotient; IDA, iron-deficiency anemia; MCHC, mean corpuscular haemoglobin concentration; EP, erythrocyte protopophyrin; CNS, central nervous system; MDI, mental development index.

mic children were inattentive and hyperactive but gave no details.

In the four studies that assessed the children's achievement in preschool or school subjects or placement in special classes, all found that formerly anemic children were poorer (de Andraca et al. 1990, Hurtado et al. 1999, Lozoff et al. 2000, Palti et al. 1985). Two studies found that anemic children had minor neurological dysfunction at 5 (de Andraca et al. 1990) and 7 y of age (Cantwell 1974).

Conclusions from longitudinal studies. In conclusion, longitudinal studies indicate consistently that children who were anemic in early childhood continue to have poor cognitive and motor development and school achievement into middle childhood. There is some evidence of behavior prob-

lems and minor neurological dysfunction, but evidence is not sufficient for identifying specific cognitive deficits.

The main question is whether the control for social background was adequate. Some environmental variables may not have been controlled for in the final analyses because they were not significantly related to the outcome variable. However, with small samples, this does not necessarily mean that they were not related. It remains possible that environmental variables, measured and unmeasured, could partly or completely explain these findings.

Therapeutic treatment trials

Numerous reviews concern the requirements that must be fulfilled in a treatment trial before it is possible to make causal inferences (Fairchild et al. 1989); thus, we will not discuss them in detail. Briefly, the following must be in place: the definition of initial iron-deficiency anemia should be clear and include at least three criteria; the samples must be large enough to provide adequate power; there should be a randomized control group that receives a placebo; treatment should be effective in improving the iron status; both tester and subjects should be blinded; and the outcome measures should have satisfactory construct validity, be sensitive to the range of changes expected and be reliable over time and between observers. If the data are required for determination of policy decisions, then the measurements must have good face validity for the policy makers. It is probably not necessary to point out to a group of fellow researchers that it is not always possible to meet all of these criteria for logistical and ethical reasons, but they remain the yardstick by which to evaluate studies.

In **Tables 2** and **3** we have listed all of the treatment trials with iron-deficient children with and without anemia (children under and over age 2 y) that we located. We have attempted to focus on the most salient points that help determine the validity of the studies. We have identified the sample size and ages and type of exclusions, definition of initial iron deficiency, manner of group assignment, content and duration of treatment, and outcome variables. We have indicated whether a significant treatment effect was reported. Treatment effect was restricted to cases in which the change in development of the treated anemic or iron-deficient group was significantly different from the change in development of the placebo anemic or iron-deficient group. However, we have also indicated whether the placebo and treated groups were significantly different after treatment. For studies in which there was no randomized anemic placebo group or the analysis of differences in score change was not reported, we have not claimed a treatment effect. We have also noted whether a response in hemoglobin level was reported. We have divided the children <2 y old from those >2 y old because the findings tend to be different.

Therapeutic treatment trials in children <2 y old

We identified nine studies of iron treatment in anemic children <2 y old (Aukett et al. 1986, Driva et al. 1985, Harahap et al. 2000, Idjradinata and Pollitt 1993, Lozoff et al. 1982b, 1987 and 1996, Oski and Honig 1978, Walter et al. 1983 and 1989) and one study of nonanemic iron-deficient children (Oski et al. 1983). Two of the studies conducted more than one trial with the same sample (Lozoff et al. 1987, Walter et al. 1989), making a total of 12 trials (Table 2).

The definition of anemia varied from hemoglobin <110 g/L in six studies (7 trials), <105 g/L in two studies (3 trials) and <100 g/L in two studies. It may be relevant that two of the

most experienced investigators in this area, B. Lozoff and E. Pollitt, usually used the lower cut-off values (Idjradinata and Pollitt 1993, Lozoff et al. 1982b, 1987 and 1996).

All studies except two (Aukett et al. 1986, Driva et al. 1985) had criteria for other measures of iron deficiency, and all of the studies lasting longer than 2 wk provided evidence that the treatment given was satisfactorily delivered by demonstrating changes in hemoglobin levels.

All but one study (Aukett et al. 1986) used the Bayley Scales as an outcome variable, which facilitates comparison across studies. The Bayley is a global measure of development and gives both a PDI and MDI. Initial differences between nonanemic and anemic children were found in the MDI in six of seven studies reporting the comparison and in the PDI in five of the studies.

Short-term trials. Seven early studies were short, lasting <15 d (Oski and Honig 1978, Oski et al. 1983, Lozoff et al. 1982b and 1987, Walter et al. 1983 and 1989). Four of them were double-blind randomized controlled trials (DBRCT) (Lozoff et al. 1982b and 1987, Oski and Honig 1978, Walter et al. 1989) and one was a randomized controlled trial (RCT) but without a placebo group (Driva et al. 1985). None of these five reported a significant treatment benefit. Investigators in two of the studies claimed treatment benefits. Oski and Honig (1978) showed that more of the treated anemic group improved 10 MDI points than did the placebo group, but this was an exploratory post-hoc analysis. The other study claiming benefits (Driva et al. 1985) showed that in the 10 d immediately after an iron injection, children showed a significant benefit but not after that. However, the appropriate analysis of difference between the treated and placebo group in score changes was not reported. Further problems with this study are that hemoglobin was the only criterion for iron deficiency and no placebo was given.

The other two short-term trials had no randomized control groups. Treated anemic Chilean children (Walter et al. 1983) or nonanemic iron-deficient children from the United States (Oski et al. 1983) were compared with nonanemic iron-replete children. In Chile, the treated anemic children improved significantly more (10 MDI points) than did the nonanemic iron-replete children (−1 MDI point). In the American study (Oski et al. 1983), the nonanemic iron-deficient groups combined gained significantly more than did the nonanemic iron-replete and -depleted groups.

Longer-term trials. In all, there were six studies in which anemic subjects were treated for 2–6 mo (Aukett et al. 1986, Harahap et al. 2000, Idjradinata and Pollitt 1993, Lozoff et al. 1987 and 1996, Walter et al. 1989). Four of these trials used nonanemic iron-replete subjects as controls (Harahap et al. 2000, Lozoff et al. 1987 and 1998, Walter et al. 1989). These studies were based on the idea that the anemic group would have an initial deficit and should catch up with iron treatment. In three of the studies (Lozoff et al. 1987 and 1998, Walter et al. 1989), the anemic group failed to show an improvement significantly greater than the iron-replete group. However, in one study (Lozoff et al. 1987), the subset of children showing complete recovery in anemia and iron status caught up to the iron-replete group in MDI and PDI. In the fourth study (Harahap et al. 2000), anemic children initially had poorer motor development and caught up with the nonanemic children during the study.

Only two of the longer-term iron trials were DBRCT (Aukett et al. 1986, Idjradinata and Pollitt 1993). One of the latter trials (Aukett et al. 1986) failed to find a significant treatment effect on scores of the Denver screening test; however, the authors reported a post-hoc analysis showing that

TABLE 2

Therapeutic treatment trials in children < 2 y old[1]

Study	Sample	Age	Exclusions	Study design and treatment	Outcome measures	Findings	Drop-outs	Remarks
Oski et al. 1978 USA	24 IDA, Hb < 105 g/L, MCV < 74 serum Fe < 51 mcg, transferrin saturation < 13% 12 treated IDA 12 untreated IDA	9–26 mo	Intercurrent illness or chronic illness	DBRCT Treatment = IM Fe Placebo = IM saline Duration 5–8 d Dose = enough to raise Hb to 120 g/L & replenish stores	Bayley Scales and IBR	Treatment: Change in scores not significantly different between groups in MDI or PDI Fe treated group significantly increased in MDI (but the control did not). No significant change in PDI of either group. 8/12 treated group increased >10 points in MDI compared with 3/12 controls ($P < 0.05$) treated group improved more than controls in reactivity ($P < 0.05$), gross and fine motor ratings ($P < 0.01$) attention not significantly different, other ratings not reported.		Small groups Short duration
Lozoff et al. 1982a, Guatemala	Total 68 out of 206 screened IDA group = Hb ≤ 105 g/L + 2 of 3 of ferritin ≤ 12 μg/L, transferrin ≤ 10% free erythrocyte protopophyrin > 1 mg/L of packed cells Non-A group = Hb ≥ 120 g/L IDA treated = 15 IDA placebo = 13 non-A treated = 19 non-A placebo = 21	6–24 mo	Hb ≤ 6, acute or chronic illness, birth complications, prematurity, congenital anomalies, retardation, malnutrition, birth weight < 5 lb	DBRCT Both IDA and non-IDA randomly assigned to Rx or placebo Treatment = 5 mg/kg ferrous ascorbate Duration of Rx = 1 wk. Placebo = carrier	Bayley Infant Behaviour Record	Initial: significant differences: IDA group more withdrawn, fearful, tense, unreactive to usual stimuli compared with non-IDA group. Treatment: No significant treatment effect IDA group improved on all above measures with significant change in responsiveness and tension. Non-IDA group were unchanged in 3 scales and deteriorated in 3 others. Only post-treatment difference was IDA group remained more fearful $P = 0.06$ Iron status by end: Treated IDA group Hb increased + 8.7 g/L $P < 0.01$ Placebo treated Hb = −0.17	7 non-IDA children—3 because post ℜ Hb was ≤ 10.5 suggesting wrong group assignment	Short duration Small groups
Lozoff et al. 1982b, Guatemala	Same as above.	Same	Same	Same	Bayley Scales	Initial: MDIs of IDA vs. non-IDA = 86.6 ± 23 vs. 100.4 ± 15.5 $P < 0.0025$. PDI 85.6 ± 19.7 vs. 94.4 ± 14.3 $P < 0.025$ Treatment: no significant treatment effect, all groups improved in MDI Iron treated IDA did not improve more than placebo treated IDA or non-IDA groups		
Oski et al. 1983 USA Honig et al. 1984	From 284 screened children 38 = non-A (Hb > 110 g/L) grouped into 4: Fe replete = 10 Fe depleted = (ferritin < 12 μg/L) = 10 Fe deficient (b + EP > 0.3 mg/L) = 10 Fe deficient (c + MCV < 70 fL) = 8	9–12 mo	Prematurity, neonatal distress, congenital anomalies, chronic illness	No randomization All subjects received IM Fe Duration 1 wk	Bayley Scales MDI IBR	Initial: MDI of group d less than group b ($P = 0.01$) but a & b not different from c & d Less involved ($P < 0.5$), more solemn ($P < 0.054$), attention, goal directed, responsivity, irritability not different, Treatment: Mean increase in MDI of a and b = 93.7 ± 21.3 to 98.6 ± 23.4 (9.6 points), less than group c and d combined 84.6 ± 19.0 to 106.3 ± 15.3 (22 points, $P < 0.01$) Goal directness c + d improved more than a + b ($P < 0.05$), All groups had normal ferritin levels		Small groups Not randomized Short duration Some rating change differences not given
Walter et al. 1983 Chile (subsample of Walter et al. 1989)	The last 51 of 314 infants completing a preventive trial. Had stopped breast feeding by 3 mo. 37 of the 51 had complete final data. Three groups: IDA = 10 Non-A Fe deficient = 15, Fe replete = 12 Criteria: IDA = Hb 109–86 g/L and at least two abnormal Fe measures or a response in Hb or MCV to Fe Rx); Fe deficient = Hb ≥ 11.0 but one or more abnormal Fe measures; Fe replete = (Hb ≥ 110 g/L, MCV ≥ 70 fL, TS ≥ 10%, FEP < 100 mcg, ferritin ≥ 10 μg/L or MCV <	15 mo	BW < 2500 g, neonatal complications, chronic or congenital disorders, inadequate growth or development	Original cohort was randomized at 3 mo to Fe fortified or unfortified formula. At 15 mo iron status was evaluated and all received 3–4 mg/(kg · d) ferrous sulphate Duration = 15 d	Bayley Scales and IBR.	Initial: IDA had lower MDI (98 ± 11) than Fe replete group (113 ± 10, $P < 0.0025$), whereas non-A Fe deficient group was not different (108 ± 13) PDI showed no significant group differences IDA infants more unhappy than Fe replete group ($P < 0.05$) on IBR. Treatment: Improvement in MDI larger in IDA group (10 points) than Fe replete group (−1 point $P < 0.05$) At post test, non-A Fe replete group still had higher scores than IDA group ($P < 0.05$) Non-A Fe deficient group improved (5 points) not significantly different from iron replete group No significant PDI differences IDA improved in cooperativeness and attention on the IBR ($P < 0.05$), but treatment effect not reported		Small groups Not randomized Other ratings not reported Short duration

TABLE 2 (*continued*)

Therapeutic treatment trials in children < 2 y old[1]

Study	Sample	Age	Exclusions	Study design and treatment	Outcome measures	Findings	Drop-outs	Remarks
Walter et al. 1989, Chile	All participated in a preventive trial from 3 mo of age (Table 4) 196 infants at 12 mo IDA = 39 Non-A Fe deficient = 127 Fe replete = 30 Criteria: IDA = Hb < 11.0 + two or more abnormal Fe measures. Non-A Fe-deficient = Hb > 11.0 but not Fe replete Fe replete = (Hb ≥ 110 g/L, MCV ≥ 70 fL, Fe/iron binding capacity ≥ 10%, FEP < 100 mcg, ferritin ≥ 10 μg/L + <10 g/L increase in Hb after Fe treatment.	12 mo	Same as above	Study 2: DBRCT All groups randomized to placebo or 15 mg of Fe^{2+} Duration = 10 d. Pre and post-testing Study 3: All subjects from study 2 given 15 mg of Fe^{2+} until 15 mo, pre- and post-testing Duration = 3 mo	Bayley MDI, PDI, IBR.	Initial: IDA infants had lower MDI (96.4 ± 1.3) than Fe replete (102.1 ± 1.8 $P < 0.01$) and non-A Fe deficient group (103.4 ± 0.8 $P < 0.01$). IDA scored lower on PDI (90.0 ± 2.0) than Fe replete group (101.2 ± 2.1 $P < 0.01$) and non-A iron deficient group (98.7 ± 1.0 $P < 0.0001$). PDI and MDI showed a sigmoid curve relationship with Hb levels with the intermediate point (Hb 10.5–10.9) sig. diff ($P < 0.05$) from both extremes (i.e. < 10.5 or > 11.0). Treatment: 2nd study No significant treatment effect from 10 d of iron. Treatment 3rd study No significant improvement in any group after 3 mo Anemia corrected in all IDA children	0	Study 2: Short duration Small groups Study 3: Not randomized
Driva et al. 1985 Greece	48 babies in an institution Non-A (Hb > 110 g/L) = 8 IDA (<109 g/L) = 40 Group A, IDA iron treated after first test = 20 Group B, IDA iron treated after 2nd test = 20 No other iron status cut-off	3–25 mo	None given	RCT, IDA tested 3 times 10 d apart. Randomly assigned to treatment after 1st or 2nd test. Non-A treated after 1st test and tested twice Treatment = IM FE 50 mg, no placebo, Duration = 10 days	Bayley Scales	Initial: No significant difference, MDI; A = 90.2 ± 14.5, B = 95.2 ± 17.4, C = 106.5 ± 26.3 Treatment: No significant treatment effect on PDI Group A significantly improves between 1st and 2nd (+7.0 points) but not between 2nd and 3rd test (+6.6 points) but not between 1st and 2nd (+2.8) Non-A no significant increase (+2.2 points)	0	N placebo ?blind testers Some IDA may not have been iron deficient Analysis of treatment effect not reported
Aukett et al. 1986, UK	Invited to attend clinics, 470 attended and were screened. All with Hb between 80–110 g/L (no other Fe cut off) enrolled, 90% were Asian Treated = 54 Placebo = 56 No other iron status cut off. In treated and placebo, respectively: mean ferritin 8.7 and 7.7 μg/L, transferrin saturation 6.9 and 13.5%, serum iron 8.3 and 13.0 μmol/L	17–19 mo	Hb < 8, lead poisoning, chronic health problems,	DBRCT treatment = 24 mg Fe + 10 mg vit. C per day Placebo = 10 mg vit. C/day Duration = 2 mo.	Denver screening test. Anthropometry.	*Treatment*: No significant treatment effect on increase in number of psychomotor skills No difference between those with Hb increase > 20 g/L and those with less Expected rate of development was achieved 31% of iron treated and 12% of the placebo group ($P < 0.05$) Greater weight gain in iron treated group ($P < 0.001$) Hb increased a mean of 22 g/L in Rx and 0.3 g/L in placebo group	13	Denver not sensitive Rate of development a post-hoc analysis
Lozoff et al. 1987, Costa Rica	Total = 191 from house to house survey Four groups: a) IDA Hb < 105 g/L Fe deficient (+2 other measures) = 52 b) Intermediate, Hb = 10.6–119 g/L Fe deficient = 45 c) Non-A Hb > 120 g/L Fe deficient = 21 d) Non-A Fe depleted (+ low ferritin) = 38 e) Non-A Fe replete = 35	12–23 mo	LBW, multiple pregnancy, perinatal complications, congenital anomalies, iron therapy after 6 mo, IM Fe at any age, acute or chronic ill health, retardation, abnormal Hb, or missing iron data	1st study DBRCT Groups a and b randomized to IM Fe/oral Fe or placebo Group c, d, e randomized to oral Fe or placebo Duration of Rx 1 week 2nd study after 1st week, IM treated infants and iron replete infants were for given placebo and the rest treated with oral Fe. No randomization Duration of Rx 12 weeks	Bayley Scales	Initial: Hb < 100 g/L significantly lower MDI than rest combined 96.6 ± 1.9 vs. 104.6 ± 0.9 $P = 0.0002$. Hb < 105 g/L significantly lower PDI than rest combined 103.0 ± 2.2 vs. 113.0 ± 1.3 $P = 0.0001$ 1st study All groups increased in MDI, IDA treated (+6) and placebo (+2) had small increase in PDI, no significant treatment effect Only Fe treated IDA increased Hb by 10 g/L 2nd study No significant difference in change of scores between non-A & IDA groups But IDA (<100 g/L) whose anemia and iron lack were corrected caught up to infants with initial Hb > 10 who became Fe replete, but those remaining Fe deficient continued to have lower MDI IDA (<100 g/L) who became Fe replete increased PDI significantly 10 points to catch up with non-A iron sufficient those remaining Fe deficient still had significantly lower PDIs Increase in Hb of at least 10 g/L in 93% of Fe treated infants, 64% no longer anemic but still Fe deficient	6.7% left after agreeing to participate	1st study: short duration 2nd study: differences remained after controlling for covariates not randomized

TABLE 2 (*continued*)

Therapeutic treatment trials in children < 2 y old[1]

Study	Sample	Age	Exclusions	Study design and treatment	Outcome measures	Findings	Drop-outs	Remarks
Lozoff et al. 1998 Costa Rica	Same as above	Same	Same	Same as study 2 above	Videotape of 15 minutes free play, IBR; behaviour ratings & quality of maternal participation HOME	Initial: IDA infants were more wary, and hesitant, maintained closer contact with caregiver, showed less pleasure and delight, tired easily, made fewer attempts test items, less playful, and less attentive. Their caregivers showed less pleasure in them, less affection. Testers made fewer attempts to demonstrate or encourage them. Treatment: no significant treatment effect, formerly IDA continued to spend more time nearer caregivers, made fewer attempts at tasks, more likely to be crying, irritable, asleep, less likely to be playing interactively Post Rx Hb see above	?none	
Idjradinata 1993, Indonesia	Total 126 IDA Fe treated = 25 IDA placebo = 25 Non-A Fe deficient Fe treated = 14 Non-A Fe deficient placebo = 15 Non-A Fe replete Fe treated = 24 Non-A Fe deficient placebo = 24 IDA = Hb ≤ 105 g/L, TS ≤ 10%, Ferritin ≤ 12. Fe deplete = Hb ≥ 120 g/L, TS ≤ 10% Ferritin ≤ 12. Fe replete Hb ≥ 120 g/L, TS > 10%, Ferritin > 12	12–18 mo	BW < 2500 g, multiple pregnancy, congenital anomalies, perinatal complications, hemoglobinopathy, wt and ht < 2 SD of reference standards, micronutrient suppl in past 6 mo, acute or chronic illness, Hb between 10.5 and 120 g/L	DBRCT Stratified by iron status group then randomized to treatment or placebo Rx = Fe sulphate 3 mg/(kg · d) Duration = 4 mo	Bayley Scales	Initial: IDA scored significantly less in MDI (treated = 88.8, placebo = 92.4) than non-A Fe deficient (treated 102.4, placebo = 101.8) and Fe replete (treated = 105.4, placebo = 104.7). IDA scored significantly less in PDI (treated = 88.5, placebo = 92.4) than iron deficient (treated = 102.9, placebo = 103.5) and Fe replete group (treated = 105.3, placebo = 105.9), latter two groups not significantly different Treatment: IDA groups: significant treatment effect in MDI and PDI Treated IDA, MDI = +19.3 points, PDI + 23.5 points; placebo IDA, MDI = +0.5 points, PDI + 5.1 No longer any difference between treated IDA and non-A Fe deficient and Fe replete groups Fe replete and deficient groups had no significant treatment effect Significant treatment effects on Hb in IDA and Fe deficient group	? 7 lost	Previous delay reversed. Non-A Fe deficient group small
Lozoff et al. 1996 Costa Rica	Survey of low-middle class neighbourhood IDA = 32 Non-A = 54 Non-A = Hb > 125 g/L IDA = Hb ≤ 100 g/L + 2 of 3 measures ferritin ≤ 12 μg/L, transferrin ≤ 10% free erythrocyte protopophyrin > 1 mg/L of pcv	12–24 mo.	BW < 2500 g, birth complications, multiple pregnancy, acute or chronic health problems	IDA all treated Non-A randomized to treatment or placebo Rx = 6 mg/kg/day Duration = 6 mo	Bayley Scales and IBR at baseline, 3 mo & 6 mo later	Initial: IDA group's MDI was 6 points lower than non-A group (P < 0.05) No group difference on PDI IBR: IDA were more fearful (P < 0.03), unhappy (P < 0.01) Treatment: No significant treatment effect IDA group disadvantaged in maternal education and home stimulation and less breast feeding When controlled for all covariates IDA not significantly different from non-A group in MDI (limited power) No difference in IBR post-Rx but treatment effect not reported		No randomization
Harahap et al. 2000 Indonesia	Subsample from a randomized trial of infants in day care centres. 18 IDA from group 1 and 2, 18 non-A from group 3 IDA = Hb < 110 g/L, TS > 16% or change in Hb > 10 g/L non-A = > 110 g/L, TS > 16%	10 IDA 12 mo, 8 IDA 18 mo, 9 non-A 12 mo, 9 non-A 18 mo		Day care centres were randomised to high energy + micronutrients skimmed milk + micronutrients skimmed milk. IDA children in groups 1 and 2 compared with non-A children matched for age and sex in Group 3 Duration = 6 mo	Bayley Response to novelty Object concept Motor milestones Activity at home Behaviour at home	Initial: Bayley PDI and motor activity significantly different. Milestones, MDI, object concept, novelty recognition all not significant. Treatment: IDA improves significantly more than non-A in motor development (slopes P < 0.05). MDI, object concept, novelty recognition and milestones all not significant.		Small groups, children on ceiling of test for milestones. No randomized placebo anemic group.

[1] Abbreviations: IDA, iron-deficiency anemia; Hb, hemoglobin; DBRCT, double-blind randomized controlled trials; IBR, Infant Behavior Record; MDI, mental development index; PDI, psychomotor development index; MCV, mean corpuscular volume; IM, intramuscular; RCT, randomized controlled trials; LBW, low birth weight; TS, transferrin.

significantly more treated anemic children gained the normal number of items than did children in the placebo group. The other study (Idjradinata and Pollitt 1993) showed a large significant treatment effect in both MDI and PDI.

Discussion. There is no good evidence from RCT that short-term iron treatment benefits the development of anemic young children. However the anemic groups were very small in all five of the studies and one contained only 12 subjects. No investigators reported the study's statistical power to show differences, but it must have been extremely low. In the four studies reporting the scores of placebo and iron-treated anemic groups, the iron-treated group improved more than the placebo group (Driva et al. 1985, Lozoff et al. 1982b and 1987, Oski and Honig 1978). Thus the hypothesis cannot be considered to have been tested rigorously. However, it may be that it takes longer than 2 wk for the type of skills measured by the Bayley Scales to develop. It is possible that other behaviors such as attention and motivation could change.

TABLE 3

Treatment trials in iron deficient children > 2 y old and adults[1]

Author, country	Sample	Age	Exclusions	Study design Hb response to treatment	Outcome measures	Findings	Dropouts	Remarks
Pollitt et al. 1982 USA	Fe depleted = 15 Non-A = 15 Fe depleted, Change in transferrin saturation from −1 SD to +1.5 SD, mean Hb = 112 g/L, Non-A, TS > 20%, Hb > 110 g/L	3–6 y	Physical handicap	Both groups treated Treatment = 4–5 mg/(kg · d) Duration = 4–6 mo Mean Hb increase in iron depleted = 13 g/L	Discrimination-learning tasks Oddity learning Short term memory Stanford Binet IQ	Initially: IDA group took more trials to reach criterion in 1st part but not in reversal learning, IQ, short-term memory and oddity learning not significantly different Treatment: At post-test, no significant difference between the groups in discrimination		No randomization Small groups
Soemantri et al. 1985 Central Java	588 children from 3 schools screened. 78 IDA (43 treated and 35 placebo) 41 iron replete (16 treated and 25 placebo). IDA = Hb ≤ 110 g/L and transferrin ≤ 15% Iron replete = Hb ≥ 120 g/L + transferrin ≥ 20%. Mean Hb IDA = 97 g/L, Fe replete = 132 g/L	IDA = 10.6 y Nonanemic = 11.07 y	<80th percentile of weight and height and <85th percentile for MUAC, parasite egg after deworming, malaria, hematological diseases, severe illness, physical handicap, IQ < 75	DBRCT, both groups randomized to: Treatment = ferrous sulphate 10 mg/(kg · d) Placebo = tapioca and saccharin. >90% compliance Duration = 3 mo. Change in Hb in Fe treated IDA = 26.7 g/L Placebo IDA = −11.7 g/L Fe treated non-A = 7.6 g/L Placebo Rx nonanemic = 6.7.	On enrollment: Raven Progressive Matrices (IQ), Pre- and post-treatment: abbreviated standard achievement test (mathematics, biology, social science, and language) used in the public schools, Bourden-Wisconsin test for concentration.	Initial: IDA and non-A not significantly different in IQ and concentration Non-A had significantly higher school achievement (42.3 ± 10.8) than IDA (31.8 ± 10.3) Treatment: significant treatment effect in school achievement and concentration Fe treated IDA improved significantly more in concentration (+20.6) than placebo IDA (−38.6) Iron treated IDA improved significantly more in school achievement (3.64) than placebo IDA (−0.7) No significant different between Fe treated and Placebo non-A Post-treatment score of non-A children were still significantly better than iron treated IDA group	?	Clear treatment effect Non-A made no improvement in school achievement in 3 mo ?ceiling effect ?control for school and grade level extremely low post test Hb in IDA placebo group
Pollitt et al. 1985, Egypt	Total of 203 out of which 68 were chosen: 28 IDA (18 treated and 10 placebo) and 40 iron replete [19 treated and 21 placebo]) selection based on discriminant function analysis IDA = ≤115 g/L + TS ≤ 25%, or ferritin ≤ 20 μg/L Fe replete = Hb > 130 g/L + TS = >25% or ferritin > 12 μg/L	Mean 9.5 y	?	DBRCT Treatment: 50 mg of ferrous sulphate daily Placebo: not stated Duration: 3–4 mo	Matching familiar figure test	Initial: Non-A children had significantly greater efficiency (fast + accurate) than IDA children. Treatment: Efficiency of Fe treated anemic children was significantly greater (P < 0.05) than placebo treated anemic children. No significant effect of treatment on iron replete children. Efficiency score Fe treated IDA became similar to that of non-A	0 after final selection	Limited details (letter) Clear treatment effect not reported Small groups
Pollitt et al. 1986, Guatemala	Total 50 out of 153 in a previous study. 25 IDA, 25 non-A. IDA: Hb ≤ 100 g/L and FEP > 1 mg/L. At T2 Hb > 110 g/L and FEP ≤ 175 or FEP < 100 + HB > 10 or ΔHb > 20 g/L. Non-A = Hb > 110 g/L FEP ≤ 1.5 mg/L. At T2 Hb > 110 g/L and FEP ≤ 1.5 mg/L.	30–72 mo	GA < 38 weeks, birth weight < 2500, chronic illness, severe malnutrition, primary hematological disorder	All infants treated with iron Not randomized Treatment: Fe sulphate 3 mg/kg/day. Duration 11–12 weeks Mean Hb increase in IDA = 29 g/L	Discriminant learning Short term memory Oddity learning tasks. These measure attention, memory and conceptual learning.	Initial: Discrimination task; IDA group had more trials to criterion compared with non-A group (33.2 ± 2.69 Vs 26.6 ± 1.46 P < 0.05), Memory and Oddity task groups not significantly different Treatment: Discrimination test; IDA improved significantly, groups no longer significantly different Memory test; no differences Oddity test: non-A improved more than IDA (not significantly) and had significantly better scores at post test T2.		No randomization
Deinard et al. 1986 USA	IDA = 25 Fe deficient non-A = 45 (22 given Fe and 23 given placebo) Non-A Fe replete matched to IDA group for sex, age mother's education, and race. 2nd control matched for initial DQ/IQ IDA = Hb ≤ 110 g/L, PCV ≤ 33%, EP ≥ 35, MCV < 74 fl, Ferritin < 20 μg/L. Non-A Fe deficient = Hb ≥ 11, PCV ≥ 34, EP = ≥ 35, MCV ≥ 75, Ferritin < 20. Non-A iron replete = Hb ≥ 11, PCV ≥ 34, EP < 35, MCV ≥ 75, Ferritin > 20.	18–60 mo	gestational age ≥ 38 wk, birth weight ≥ 2500 g, head circumference, height and weight within ± 2 SD of NCHS, chronic illness, developmental retardation.	All IDA treated Non-A Fe deficient: alternately assigned to Fe Rx or placebo Non-A Fe replete given placebo Double blind. Treatment = 6 mg/kg/ day of elemental Fe Duration = 6 mo IDA iron treated and both Fe deficient groups showed complete hematological correction.	Bayley MDI for infants 18–24 mo. Stanford Binet IQ for > 2 y. Behavior rating. Tested at baseline 3 and 6 mo Nutritional assessment, life stress and SES.	Initial: (standard scores of pooled MDI and IQ tests used) No significant difference between IDA and Fe replete group, Iron treated but not placebo Fe deficient group significantly lower than Fe sufficient group IDA significantly less responsive to the environment and more unhappy, responsive to examiner Treatment: Fe replete group's score significantly improved over baseline at 3 and 6 mo, and were significantly higher than IDA at 3 mo but not 6 mo. IDA and iron deficient groups showed no significant improvement, no treatment effect for IDA or Fe deficient Non-A group rated significantly more responsive to examiner than IDA (3 and 6 mo), and more responsive to environment (T1, 3 and 6 mo) and emotional tone (T1 and 3 mo). Treatment effect not reported	7	Wide age range Placebo treated Fe deficient non-A group had full hematological correction ?a valid trial No analysis of group differences in change of scores

TABLE 3 (continued)

Treatment trials in iron deficient children > 2 y old and adults[1]

Author, country	Sample	Age	Exclusions	Study design Hb response to treatment	Outcome measures	Findings	Dropouts	Remarks
Groner et al. 1986, USA	38 women < 16 wk pregnant all enrolled regardless of Hb level	14–24 y	Hematocrit ≤ 30%	DBRCT Treatment = 60 mgs elemental iron + vitamins Placebo = vitamins Duration = 1 mo	Digit span, digit symbol, arithmetic, vocabulary, Consonant Trigrams, Rey Auditory Verbal Learning Test,	Treatment: Treatment effect in arithmetic ($P < 0.02$), one subsection of Consonant Trigrams improved in treated group significantly more than placebo.	10 controls, 3 treated, 3 poor compliers	Small groups biased loss
Soewondo et al. 1989 Indonesia	Total of 139 IDA = 49 (26 Fe Rxed and 23 placebo). Fe depleted = 57 (24 Fe Rxed and 33 placebo). Fe replete = 70 (27 Fe Rxed and 43 placebo). IDA = Hb < 110 g/L plus two of Ferritin < 12 mcg, TS < 16% FEP > 1.77 μmol/L RBC Fe depleted = Hb ≥ 110 plus two of above Fe replete = Hb ≥ 110, plus two of Ferritin ≥ 12 μg/L, TS ≥ 16%, FEP ≤ 1.77	Anemic = 54 mo Fe replete = 58 mo	?	DBRCT, all groups randomized to: Treatment = 50 mg Fe/day in syrup or placebo Duration = 8 wk Hb change in iron treated IDA = 9 g/L Placebo IDA = 1 g/L	Peabody Picture Vocabulary test (PPVT). Two Discrimination learning tasks Four Oddity tasks	Fe deficient and non-A not different in any test Initial differences between IDA and Fe replete: no differences in PPVT or Discrimination learning tasks (49% on floor of test, subgroup on test; iron replete better than IDA in one task of discrimination learning) Fe replete faster than IDA on Oddity task 3 and 4 ($P < 0.05$) Treatment: no treatment effect PPVT no treatment effect, Discrimination learning tasks remained too difficult, Discrimination learning: (subgroup on test) treated IDA learned significantly faster than treated Fe replete group $P < 0.05$ at post test Oddity task; no effect in trials 1 and 2 In trials 3 and 4 interaction between the iron treatment and group, iron treated IDA had significantly higher scores than the iron treated Fe replete group and Fe replete placebo group had higher scores than the IDA placebo group	Finished with 205. Unclear what loss was	Too many on floor of discrimination learning tasks, using subgroup who were on the test of doubtful validity
Pollitt et al. 1989, Thailand	Children in grades 3–5 in 16 schools screened, Total of 1358 101 IDA, 47 Fe depleted, 1210 iron repleted groups IDA = Hb < 120 g/L + two of: ferritin < 10 μg/L, TS < 16%, FEP > 700 μg/L Fe depleted = Hb ≥ 12.0 g/L + same as above Fe repleted = Hb ≥ 120 g/L, + two of: ferritin > 9 μg/L, TS > 15%, FEP < 701 μg/L	9–11 y	Thalasemia, cyanotic heart disease	All dewormed on enrolment and after 3 mo DBRCT Randomization to iron treatment or placebo before iron status was known. Then divided into groups by iron status Treatment = 50 mg/d of ferrous sulphate for 2 wk, then 100 mg/d for 14 wk Duration = 16 wk Hb change: IDA placebo and iron treated groups increased in Hb + 14 g/L, +20 g/L respectively, Fe depleted placebo and iron treated groups increased + 1 g/L and + 5 g/L respectively, both Fe replete placebo and iron treated decreased—2 g/L	Raven Progressive Matrices Thai language and maths test, given in groups Controlled for school and grade	Pre and post treatment scores averaged—no treatment effect IQ score of IDA group 90.8 ± 1.0 significant lower than Fe replete group (94.2 ± 0.3, $P = 0.001$) Thai language score of Fe replete group (59.0 ± 0.5) significantly higher than IDA group (55.9 ± 1.5, $P < 0.05$) and Fe depleted (51.8 ± 2.2, $P < 0.010$). Math scores not significantly different Fe status groups still significantly different in language and IQ after controlling for anthropometry and SES		Hb of IDA placebo group improved, therefore validity threatened ? due to deworming
Seshadri and Gopaldas 1989 (study 1)	94 children attending school of low socio-economic status Anemia defined as Hb < 110 g/L, no other measure	5–8 y	Severe malnutrition (weight-for-age < 60% of NCHS standards).	Before Hb level known, children stratified by age, then every third child randomly assigned to control group and the next two to iron treatment group Treatment = 20 mg elemental iron and 0.1 mg folic acid/d. Placebo not mentioned Duration = 60 d Significant increase in Hb in treated group	Nutritional assessment. Indian adaptation of WISC. Tested at baseline and 60 days.	Initial: Anemic children had significantly lower WISC scores than non-A children only in 7 to 8-y-olds Treatment: Iron-treated group improved significantly in verbal, performance, and global IQ in all ages, whereas controls did not In iron treated group both anemic and non-anemic children improved in IQ The mean improvement in the global IQ of anemic children was significantly higher than for non-anemic children for the 7–8-y-olds only (11 vs. 4 points) not in 5–6 or 6–7 y olds		Analysis of treatment effect not reported by randomized groups Age groups probably too small to interpret separately (6–10) No placebo Folic acid may have independent benefits

TABLE 3 (continued)

Treatment trials in iron deficient children > 2 y old and adults[1]

Author, country	Sample	Age	Exclusions	Study design Hb response to treatment	Outcome measures	Findings	Dropouts	Remarks
Seshadri and Gopaldas 1989 (study 2)	60 at one school screened, 14 pairs of boys matched for anemia, height, weight, Hb, IQ, per capita income, and mother's educational level, selected, total = 28 anemia = Hb < 105 g/L. Hypochromic microcytic red cell morphology.	5–6 y	Weight-for-age < 61% of NCHS standard Draw-a-man (DAM) IQ of <70 and >110.	DBRCT Matched then each pair randomized to iron or placebo Both groups dewormed with 2 tablets of mebendazole for 3 days. Treatment = 40 mg elemental iron + 0.2 mg folic acid/day. Compliance confirmed Placebo = sugar. Duration = 60 days Mean change in Hb = +24 g/L in treated group and −8 g/L in controls.	Anthropometry Draw a man IQ WISC	Initial: No significant differences at baseline. Treatment: Both groups improved significantly in WISC but more (? Significant) in the iron group, 10 points vs. 5 points for verbal scale, and 17 points vs. 7 points for performance scale At end the iron treated group significantly better than controls in verbal and performance		Significance of group differences in change of scores not reported Folic acid may have independent benefits May be over matched Small groups
Seshadri and Gopaldas 1989 (study 3)	Total of 48 selected from 210 boys at one school 16 groups of three, each matched for age, Hb level, and baseline scores in cognitive function tests, Anemic = <105 g/d Non-anemic = >115 g/L	8–15 y		DBRCT Matched, then each triplet randomized into 3 groups: Treatment (a) 30 mg elemental iron/d Treatment (b) 40 mg elemental iron/d Placebo = brown sugar Duration = 60 d	Visual-recall, Digit-span, Maze (visual motor coordination) Clerical task *Initial:* No baseline differences because of matching *Treatment:* The 40 and 30 mg iron groups both significantly improved in all cognitive tests except for the maze test in the 30 mg group, no change in the placebo group (?significant difference) At end: Compared with placebo, the 30-mg group had significantly higher scores in the clerical-task and visual-recall tests, and the 40-mg group had significantly higher scores in digit-span, mazes, clerical-task and visual-recall Anemic placebo boys showed no significant improvements whereas the 40-mg and 30-mg iron treated anemic group significantly improved in several tests			No analysis reported of differences between groups in change of scores May be over matched small groups	
Seshadri and Gopaldas 1989 (study 4)	207 girls from 4 schools enrolled but after loss, 83 pairs matched for age, Hb, and cognitive scores began study (Total 166) but final data on only 65 pairs (130).	8–15 y	Family income > Rs500,	DBRCT Matched pairs randomized to: Treatment = 60 mg elemental iron/d. Placebo = sugar tablets. Duration = 60 d starting at the beginning of each 4 mo term for one school year (8 mo)	Visual-recall, Digit-span, Maze, Clerical task.	Initial: No baseline differences in Hb or cognitive test scores in treated or placebo group, nor between anemic and non-A children Treatment: Fe treated anemic children significantly better than placebo anemic children in overall scores and in clerical tasks, and mazes Fe treated non-A group improved significantly only in mazes	95	Analysis not presented by randomized group Adding test scores of doubtful validity
Bruner et al. 1996 USA.	Four high schools, grades 9–12 81 non-A Fe deficient girls; Treated = 40 Placebo = 41 Anemia cut off = 120 g/L for white and 115 g/L for black girls, Fe deficient = normal Hb + ferritin < 12.0 dL	13–18 y	Boys,	DBRCT Treatment = equivalent to 260 mg elemental iron/d Duration = 8 wk Hb before and after. Treated = 13.1 ± 0.7 vs. 13.5 ± 0.8 Control = 13.0 ± 0.7 vs. 12.7 ± 0.7.	Brief Test of Attention (BTA, auditory divided attention) Symbol Digits Modalities Test (SDMT) Visual Search and Attention (VSAT) Hopkins Verbal Learning Test (HVLT).	Treatment: Significant treatment effect Iron treatment had no significant effect on BTA, SDMT, or VSAT On HVLT iron treated group improved significantly more in total score of 3 free recall items than the placebo group (P < 0.02). No significant difference in delayed recall or recognition parts of test	8: 3 controls and 2 treated dropped because they became anemic	Benefits limited to free recall Ferritin only measure of iron status
Lynn and Harland, 1998, England	12–16 y attending 7 schools. All children regardless of Hb level. Treated = 208, placebo = 205, 2.9% hemoglobin < 12 gms, 16.9% ferritin > 12 μg/L.	12–16 y		Divided into 2 groups matched for age, sex and IQ, method of assignment not given. Treatment = 17 mg elemental iron + 17 mg ascorbic acid Duration = 16 wk	Ravens Progressive Matrixes (IQ)	Initial: Correlation between Hb and IQ R = 0.17, P < 0.01, ferritin and IQ not significant Treatment: No significant difference between the groups Sub group with ferritin < 12 μg/L improved significantly more with treatment than placebo (treated + 3.1 IQ points, placebo − 2.7 IQ points, P = 0.02), subgroups with ferritin 0.12–0.2 mg/L not different, subgroups with ferritin > 20 μg/L treated improved more than placebo (P < 0.05)	Full data on 413 original sample ~ 600	Moderate and high ferritin groups combined showed no treatmenteffects

[1] Abbreviations: IDA, iron-deficiency anemia; Hb, hemoglobin; IQ, intelligence quotient; TS, transferrin; MUAC, mid-upper arm circumference; DBRCT, double-blind randomized controlled trials; NCHS, National Center for Health Statistics; EP, erythrocyte protophyrin; PCV, packed cell volume; MCV, mean corpuscular volume; SES, socioeconomic status; MDI, mental development index; FEP, free erythrocyte protophyrin

The Indonesian study (Idjradinata and Pollitt 1993) is unique in being a DBRCT with long treatment and using the Bayley test. It is also unique in showing a clear significant treatment effect in both MDI and PDI. The treated children showed an extremely large improvement in both indices. The size of the increase is surprisingly large for a 4-mo period. However, in one other study (Oski et al. 1983), children showed a similar increase in less time. In the only other DBRCT with this age group (Aukett et al. 1986), the Denver screening test was used as the outcome measure. This test is not sensitive to small differences and was intended to screen for children with abnormal development. Most studies indicate that the development of anemic children is within the normal distribution.

In four long-term trials (Harahap et al. 2000, Lozoff et al. 1987 and 1996, Walter et al. 1989), using only nonanemic iron-replete groups as controls, the anemic group did not catch up to the nonanemic group in three but did in one (Harahap et al. 2000). There are at least two reasons why it is not possible with this design to infer the presence or absence of a causal relationship. First, most children improve slightly with test practice and we cannot be sure that untreated anemic and nonanemic children respond in the same way. It is possible that anemic children have poor test behavior and do not learn as much as the nonanemic children from the test experience and thus improve less. In this case, improving the same amount as nonanemic children could actually represent an iron response that would not be detected. Another possibility is that anemic children's fearful and unresponsive behavior causes them to perform badly at the first test and subsequently improve more that the nonanemic group.

When we examined the four studies with placebo anemic and nonanemic iron-replete groups there were no consistent differences between the groups in their changes in Bayley scores between tests. The change in scores in nonanemic and anemic placebo groups were +6 and +2 (Lozoff et al. 1987), +5.1 and +5.5 (Lozoff et al. 1982b), +8.3 and +6.7 (Walter et al. 1989), +2.1 and +0.5 (Idjradinata and Pollitt 1993), respectively.

Nonanemic controls are helpful in assessing whether iron-deficient anemic children catch up to nonanemic children. However, anemic children usually come from poorer environments that should be controlled for before catch-up is examined. Most investigators restricted the range of social background and biomedical conditions found in the study children, but few controlled further for environmental factors when examining catch-up. Of the five longer-term studies that had nonanemic groups, the treated anemic group failed to catch up to the nonanemic group in scores on the Bayley test in three studies (Lozoff et al. 1987 and 1996, Walter et al. 1989). Anemic children in the two other studies caught up to the nonanemic group (Harahap et al. 2000, Idjradinata and Pollitt 1993). It is difficult to explain why children in two studies caught up and those in the others did not. Duration of treatment did not explain this because the Costa Rican study (Lozoff et al. 1996) had the longest treatment period. Severity of anemia did not explain this either because two of the studies had less severe criteria for anemia than did the Indonesian study (Idjradinata and Pollitt 1993). Disparities in social background between the anemic and nonanemic groups may explain some of the failure to catch up.

Outcome measures. The Bayley test is a global measure and gives little indication as to any specific cognitive deficit. Its predictive ability in y 1 of life is extremely limited, but increases in y 2 (Colombo 1993). Other measures of infant cognitive development such as novelty preference and fixation time (Colombo 1993, Goswami 1998) might be more sensitive and predictive in y 1. However, the Bayley test was sensitive to initial differences between anemic and nonanemic groups in nearly all studies. Furthermore, marked changes were found with iron treatment in the Indonesian study (Idjradinata and Pollitt 1993) in both MDI and PDI. The PDI of the Bayley test has also been sensitive to changes from nutritional supplementation in studies of protein-energy malnutrition (Husaini et al. 1991, Joos et al. 1983).

Summary. In general, these studies are difficult to interpret mainly because so few were RCT and the samples were often extremely small. There has been a hesitancy to use placebo groups in the field of iron deficiency on ethical grounds. This is the main reason we do not have clear answers to the important question of whether iron treatment can benefit the development of anemic children. There is no clear evidence that short-term iron treatment has such a benefit; however, the question has not been tested rigorously. Long-term treatment has clearly been shown to benefit the development of anemic children in only one relatively small study. We could find no other study that had rigorously evaluated (with an RCT) the effect with sensitive outcome measures. In several studies, but not all, anemic children have failed to catch up to nonanemic children with iron treatment. We located no study that looked at the effect of anemia in high-risk children (e.g., low birth weight)

Therapeutic treatment trials in children >2 y old

We identified 13 iron treatment trials with anemic children >2 y of age that included iron-deficient children and one trial in pregnant women. One reported no statistical analysis of the treatment and will not be discussed in detail (Soemantri 1989). The more important details of the others are given in Table 3. (Bruner et al. 1996, Deinard et al. 1986, Groner et al. 1986, Lynn and Harland 1998, Pollitt et al. 1983, 1985, 1986 and 1989, Seshadri and Gopaldes 1989, Soemantri et al. 1985, Soewondo et al. 1989).

Comparison with nonanemic iron-replete children. Eight studies (Deinard et al. 1986, Pollitt et al. 1983, 1985, 1986 and 1989, Soemantri et al. 1985, Soemantri 1989, Soewondo et al. 1989) had nonanemic comparison groups and in one (Deinard et al. 1986), no initial difference was found on enrollment between the groups in developmental quotient and IQ. This study had a very wide age range (18–60 mo) and scores on the Bayley test were combined with scores on the Stanford Binet test. In the other seven studies, the anemic children had significantly poorer cognitive scores or school achievement than did the nonanemic iron-replete groups in at least one test. A variety of tests have been used. In the two studies that used a test of discrimination learning, the groups were different (Pollitt et al. 1983 and 1986); this test was reported to depend largely on attention. In two of three studies using an oddity learning test, which measures concept learning, the groups were different (Pollitt et al. 1986, Soewondo et al. 1989). IQ tests did not show differences in three studies [the Stanford Binet test used by Pollitt et al. (1983) and the Ravens Progressive Matrices used by Soemantri (1989) and Soemantri et al. (1985)] but did in one study [the Ravens Progressive Matrices used by Pollitt et al (1989)]. Differences were not detected in short-term memory in two studies (Pollitt et al. 1983 and 1986). The Matching Familiar Figures test, which measures the efficiency of solving a visual perceptual problem, was different in another study (Pollitt et al. 1985), whereas scores on the Peabody Picture Vocabulary Test were not different in one study (Soewondo et al. 1989). Tests of school

achievement were given in two studies and anemic children had lower scores in both (Pollitt et al. 1989, Soemantri et al. 1985).

Catch-up to nonanemic. In the six studies showing initial differences, the anemic children in three studies caught up to the level of nonanemic children in scores on cognition; the anemic children in one caught up in discrimination learning but not in oddity learning (Pollitt et al. 1986). In contrast, children in neither of the two studies measuring school achievement caught up (Pollitt et al. 1989, Soemantri et al. 1985) although they improved on one of them (Soemantri et al. 1985). Catching up in school achievement would presumably take a long time because not only do children have to function well but they also need the opportunity to learn missed material. This may not occur in schools without facilities to give children personal attention.

Study design. Three early studies (Deinard et al. 1986, Pollitt et al. 1983 and 1986) had only nonanemic children as controls and suffer from the problems discussed previously; we will therefore not discuss these three further. One study had controls matched for age, gender and IQ but the method of assignment is not clear (Lynn and Harland 1998). Nine of the remaining studies were randomized controlled trials. One did not give a placebo (Seshadri and Gopaldes 1989, study 1), but the others were all double blind. Several had sample sizes of <30 in each treatment group (Table 3) but four had reasonably adequate or good numbers (Bruner et al. 1996, Pollitt et al. 1989, Seshadri and Gopaldes 1989, study 4; Soemantri et al. 1985). The study conducted by Lynn and Harland (1998) had large groups but most children were not anemic.

School achievement. Two studies had robust designs and adequate sample sizes and examined school achievement, which has obvious policy implications. In Indonesia, anemic children showed a clear improvement with treatment (Soemantri et al. 1985), but it is puzzling that neither the placebo anemic group nor the two nonanemic groups made any improvement in 3 mo of schooling. This lack of progress suggests that the school was not teaching effectively; if this was the case, it may be that the treated anemic children's ability to take tests in general improved rather than they had learned more. This would impute such factors as attention and motivation. An alternative explanation is that the test was insensitive over a certain level, causing a ceiling effect. In the other study (Pollitt et al. 1989), children in Thailand showed no improvement with treatment. The criterion for anemia in that study was high, hemoglobin <120 g/L, and the placebo group actually improved by 14 g/L, probably secondary to the deworming. It is possible that the placebo group responding to improved iron status may have threatened the integrity of the trial. In contrast, in the Indonesian study, the criterion for anemia was lower (<110 g/L) and the placebo group's hemoglobin levels declined by 17 g/L during the study. A further consideration is that the children were tested in groups in the Thailand study, which is unlikely to be as accurate as individual testing; it is not clear how testing was conducted in the Indonesian study.

Cognition. In another Indonesian study with only 24 treated anemic children (Soewondo et al. 1989), no effect was found on vocabulary scores (Peabody Picture Vocabulary Test), and many children were on the floor of the discrimination learning test, making its validity doubtful. In the oddity test, after initial test scores were controlled for, there was no significant difference between the iron-treated and placebo anemic groups, but an interaction occurred between treatment and group in two of four tasks. The treated anemic group improved more than the treated nonanemic group, indicating

some catch-up. Only limited details were available from an Egyptian study (Pollitt et al. 1985). The Matching Familiar Figures test was used; the efficiency of the treated anemic children improved significantly, and they were significantly better than the placebo group at post-test. Group differences in change of scores were not reported. There were four studies from India (Seshadri and Gopaldes 1989) and in each the children were randomly assigned to treatment or control, regardless of hemoglobin level. The first study had no placebo; in the second study the children were given folic acid as well as iron, which may have had an independent benefit. However, the iron-treated group was significantly better than the control group at the end in the second study. In the third study, the treated group improved significantly in most of the cognitive tests, whereas the placebo group did not. At the end, the treated groups had higher scores than the nontreated groups. In the fourth study, iron-treated anemic children improved more than the placebo-treated anemic children in two of four tasks. There was a suggestion of a treatment effect in all four studies, but none reported the significance level of differences in change of scores by the randomized groups. Only the fourth study reported difference between the groups in change of scores, but they restricted the analysis to anemic children only, thus breaking the paired randomized design. It is probable that there was a treatment effect in the third and fourth studies.

In an English study (Lynn and Harland 1998), there was no overall treatment effect. When the children were divided into subgroups by iron status, the subgroup with low iron status (ferritin <12 μg/L) showed a significant treatment effect on a test of visual reasoning. However, the findings were inconsistent and the group with moderate iron status showed no benefit from treatment, whereas the group with high iron status showed significant benefits. Also, ascorbic acid was given in the iron treatment and may have had an independent effect.

In a trial with pregnant women, the treatment group improved significantly more than the placebo group in a test of short-term memory (digit span) and a test of attention (Consonant Trigrams). There was however, a large loss ($n = 10$) from the control group, leaving only 9 children.

The final study was with nonanemic iron-deficient older girls (Bruner et al. 1996). It was a well-conducted trial except that they did not have three measures of iron deficiency. Several cognitive functions were assessed, including auditory divided attention, speed of coding, visual search and attention, and verbal learning. There was no treatment effect on the first three tests. The learning test comprised three parts, and a significant effect was found in free call but not in delayed recall or recognition. Three other randomized trials included nonanemic iron-deficient or iron-depleted children in this age range (Pollitt et al. 1985 and 1989, Soewondo et al. 1989). They all failed to find any treatment effect, but the samples were smaller.

Summary. As with the younger children, anemic children usually had poorer cognition and school achievement than nonanemic children. They tended to catch up with repeated testing and treatment in cognition but not in school achievement.

There were eight DBRCT with anemic subjects or a mixture of anemic and nonanemic subjects and another in which the method of assignment is not given and ascorbic acid was given with the treatment (Lynn and Harland 1998). In two trials (Groner et al. 1986, Soemantri et al. 1985), significant treatment effects were reported by randomized group. In another two, significant treatment effects were found in the

subgroups of anemic or most iron-deficient children (Lynn and Harland 1998, Seshadri and Gopaldes 1989, study 4). Data suggestive of treatment benefits was reported from three other studies (Pollitt et al. 1985, Seshadri and Gopaldes 1989, studies 2 and 3). All three reported significant differences at the end but did not analyze differences in change in scores. These studies were reported as conference proceedings or a letter, and details were not available. In two studies, no significant treatment effect was reported (Pollitt et al. 1989, Soewondo 1995).

Some of the studies were small and must have had limited power to show differences. None reported long-term follow-up of children to determine whether benefits arose later or whether benefits were sustained.

Nonanemic iron-deficient or depleted children. We identified three randomized trials with nonanemic iron-deficient children >2 y old. Although one trial found a treatment effect in one of several tests (Bruner et al. 1996), three other studies with smaller samples did not (Pollitt et al. 1985 and 1989, Soewondo et al. 1989). In addition, one study in the <2 y age range included iron-deficient children and found no benefit from treatment (Idjradinata and Pollitt 1993). Therefore the evidence for an effect of treatment is weak.

Preventive treatment trials

We identified six preventive trials (Heywood et al. 1989, Lozoff 1997, Moffatt et al. 1994, Morley et al. 1999, Walter et al. 1989, Williams et al. 1999) (**Table 4**). Two trials are difficult to interpret. A large number of the children became infected with malaria and this confused the results in one study (Heywood et al. 1989). In the second study (Walter et al. 1989), the analysis was not reported for the original randomized groups; instead the groups were pooled and all anemic children were found to have significantly lower scores than the nonanemic groups in MDI and PDI. These two studies will not be discussed further.

The remaining four studies were all randomized trials. In two, the untreated groups were given true placebos (Lozoff 1997, Moffatt et al. 1994), whereas in the other two, the groups were assigned to receive cow's milk or fortified formula in one or cow's milk or formula with less iron in the other; thus the subjects were not blinded to the treatment. The age of the subjects ranged from 2 to 9 mo and treatment lasted from 6 to 13 mo.

In three of the studies, some beneficial response to iron treatment was found. In Canada (Moffatt et al. 1994), children were supplemented from 2 mo of age and were tested at 6, 9, 12 and 15 mo. The iron-fortified group had significantly higher PDI scores than did the placebo group at 9 mo (4 points) and 12 mo (6.3 points), but the benefit was only 2.9 PDI points at 15 mo and no longer significant. There was no significant effect on MDI. The difference between the groups in percentage anemic was greatest at 6 mo (19.9%) and had become small by 15 mo (8%). The loss from the study was approximately one third but the investigators controlled for any difference between lost and tested children. The findings suggest that the effect of iron deficiency is transient.

In England (Williams et al. 1999), children were supplemented from 7 to 18 mo of age and benefits were not found until 24 mo. The global developmental quotient fell 5.4 points more in the nonfortified group than in the fortified group (*P* < 0.05). The scores of all subscales fell less in the iron-fortified than nonfortified group but the personal social subscale was the only one to show a significant treatment effect. This was the only study to follow children for as long as 17 mo, including 6 mo after the cessation of treatment; thus, it is conceiv-

able that other studies may have had undetected benefits. The main problem with the study is that fortified formula was given to one group and money to buy cow's milk to the other group. It is possible that other constituents in the formula were responsible for the benefits or that cow's milk reduced the absorption of other nutrients. Also, the study was not double blind. The loss was not large (15%), but unfortunately this included two children excluded from the nonfortified group because of anemia and two from the fortified group excluded because of failed protocol. Some children were already anemic on enrollment at 7 mo (fortified 13%, nonfortified 16%). The difference in hemoglobin levels between the groups was considerable at 12 mo (31%) and at 18 mo was 24%.

A Chilean study (Lozoff 1997) probably had the greatest statistical power. The children showed no benefit on the Bayley test, but at 12 mo the fortified group had shorter fixation times on the Fagan test (Lozoff, personal communication), which is thought to indicate better attention and ability to encode stimuli. However, within the groups, the anemic children did not have longer fixation times than non-anemic children. This study was the shortest and, furthermore, all anemic children were excluded at enrollment; therefore, any case of anemia would have been of short duration and not have been present for at least the first 6 mo of life. Both of these factors may have played a role in the lack of effect on the Bayley test scores. The Fagan test predicts later mental development and is probably more sensitive to small cognitive differences. However, the lack of consistency within the groups in the relation between fixation times and anemia makes the finding difficult to interpret. Details of the study are not yet fully reported; thus, in-depth evaluation is not possible.

In contrast to the above, another English study (Morley et al. 1999) found no benefit from 9 mo of iron treatment begun at age 9 mo. Fewer than one third of the children had hemoglobin levels available at the beginning and end and these were considered not valid because of technical problems (Lucas, personal communication). The other measures of iron status indicate very little iron deficiency; thus, the study would have had limited power to show benefits from treatment. The results are therefore not possible to evaluate except that there was no apparent harmful effect of giving iron to nonanemic children.

Conclusions. Only three of the prophylactic studies can be assessed. All found at least a hint of some improvements, albeit in one study, treatment was confounded by formula (Williams et al. 1999). The other two studies were double-blind control trials. In the Canadian study (Moffatt et al. 1994), the benefits were small, transient and limited to motor development. In the other study (Lozoff 1997), benefits were not found on the Bayley test and inconsistent benefits were found on the Fagan test. These findings provide extremely limited evidence that preventing iron-deficiency anemia produces benefits to development. When benefits were found, they were transient or small. Longer-term follow-up may help to interpret the Chilean data.

OVERALL COMMENTS

We have discussed the problems of individual studies under the specific sections, but there are some comments common to several different types of studies. Child development is essentially longitudinal, changing over time. Events happening in early childhood may show immediate benefits or detrimental effects that disappear quickly; on the other hand, benefits or detrimental effects may appear at a later stage of development.

TABLE 4

Prophylactic trials[1]

Study	Sample	Age on enrollment	Exclusions	Study design and treatment	Outcome measures	Findings	Dropouts	Remarks
Heywood et al. 1989, Papua New Guinea	Subgroup of a larger cohort of hospital births, those born in a 4-mo period and present at 2 mo = 96 Treated = 49 Controls = 52.	2–12 mo	BW < 1.4 kg, neonatal illness, congenital anomalies	DBRCT. Infants matched by sex, birth weight and residential area then randomized to: Fe treated = IM dextran, 150 mg of elemental iron or saline IM Duration = 10 mo	Visual attention: total fixation time (TFT), mean length of fixation, rate of habituation and dishabituation.	No treatment effect in subjects with malaria parasitemia Aparasitemic children: TFT: Fe treated significantly longer than placebo group But there was an interaction between iron and malaria, the placebo parasitemic group had shorter TFT than all other groups. Mean length of fixation, rate of habituation, dishabituation, no significant treatment effect Iron treatment had a significant benefit on Hb and parasitemia was detrimental Hb before vs. after: Fe treated = 87 g/L vs. 84 g/L (+ malaria) or 101 g/L (− malaria). Placebo = 89 g/L vs. 80 g/L (+ malaria) or 93 g/L (− malaria)	23 excluded at 12 mo loss before 12 mo not clear	High level (93%) prevalence of α-thalassemia results confused by malaria therefore ambiguous
Walter et al. 1989, Chile	Low-middle class infants Total 196 Breast-fed, iron fortified cereal Breast-fed unfortified food Weaned, iron-fortified formula Weaned, unfortified formula	3–12 mo	BW < 2500 g, neonatal complications, chronic or congenital disorders, inadequate growth or development	DBRCT Stratified by feeding then randomized to four groups: Exclusively breast-fed; 1) heme iron-fortified cereal 2) health clinic food (unfortified). Weaned; 3) iron-fortified formula 4) unfortified formula. Duration = 9 mo Sample continued in a 2nd then 3rd treatment trial	Bayley, IBR.	Treatment effect not reported by group assignment See Table for relation between iron status and MDI. Infants with anemia > 3 mo scored sig. lower on MDI and PDI than those whose anemia lasted < 3 mo	0	Analysis by treatment group not reported
Moffatt et al. 1994 Canada	Antenatal clinic attenders, motly poor Amerindian, bottle fed, 283 = total Fortified Unfortified	<2–15 mo	Perinatal complications, congenital anomalies, BW < 2,500 kg, prematurity	DBRCT Randomized to: fortified formula (12.8 mg/L Fe) low formula (1.1 mg/L Fe) Duration = 13+ mo	Bayley Scales at 6, 9, 12 and 15 mo and IBR	Significant treatment effect at 9 and 12 mo in PDI but no longer at 15 mo No effect on MDI Hb significantly higher in fortified group at each test. % < 110 g/L in fortified and unfortified was 8.1 and 28.0 at 6 mo, 8.1 and 18.6 at 9 mo, 2.3 and 12.4 at 12 mo and 2.6 and 10.4 at 15 mo, respectively.	225, 204, 186 and 186 infants tested at 6, 9, 12 and 15 mo respectively	Controlling for HOME Rx effect not significant but sample smaller Large loss
Lozoff et al. 1995 Chile (Abstract)	944 healthy non-A infants, almost all had been breast-fed 625 iron treatment 319 placebo	6–12 mo	Premature, BW < 3.0 kg, low Hb, acute or chronic illness	DBRCT Randomized to iron or placebo Duration = 6 mo Dose ?	Bayley Fagan Test	No significant treatment effect on the Bayley treated vs. untreated: MDI 104 ± 12.6 vs. 105.1 ± 11.2, PDI 96.6 ± 15.0 vs. 97.0 ± 14.8 Fagan test: significant treatment effect on fixation time Anemia 4% vs. 24% (P < 0.01) iron def 15% vs. 49% (P < 0.01)	?	Almost all breast-fed and low risk Short duration All non-A in first 6 mo
Williams et al. 1999, UK.	Total = 100 infants who had started on unmodified cows milk at 6 mo from a poor neighbourhood Fe fortified formula = 50 Unmodified cow's milk = 50 Fortified = 13%, Hb < 10 g/L Cow's milk 16% Hb < 110 g/L	7–18 mo (5.7–8.6)	preterm, Hb < 90 g/L, hemoglobinopathy, chronic ill health	RCT Randomized to: Fe fortified formula (1.2 mg Fe/100 mL) Usual cow's milk = (0.05 mg Fe/100 mL) + money to buy cow's milk Duration = 12 mo then all returned to unmodified cow's milk. Observation for 18 mo	Griffiths Scale Anthropometry	Significant treatment effect on development mean scores fell in both groups but fortified group developmental quotient fell significantly less by 24 mo 14.7 Vs. 9.3 P < 0.05). The difference was not significant at 18 mo. Drop in all subscale scores was less in fortified group but only significant in Personal social subscale P < 0.05 %Hb < 110 g/L at 12 mo was 2% in fortified and 33% in cows milk, and at 18 mo fortified 0% and cow milk 24% Hb (g/L) 18 mo: fortified = 123 ± 7, cow milk = 115 ± 13	15	Subjects not blind to treatment Other constituents of formula may have caused the effect
Morley et al. 1999 UK	Volunteers from birth register in 3 areas who were feeding cows milk Total = 493 cow's milk = 166 (Hb 99 g/L) Low Fe formula = 165 (Hb 96 g/L) High Fe formula − 162 (Hb 99 g/L) Only one center had Hb estimations	9 to 18 mo	Prematurity, BW < 2500 g, Multiple pregnancy, Previous iron supplement or blood transfusion, Delayed Development, Non-English speakers.	DBRCT, stratified by 3 areas and Asian/other then randomized to: Cow's milk (0.05 mg Fe/L) Low Fe formula (0.9 mg Fe/L). High Fe formula (1.2 mg Fe/L) Duration = 9 mo	Sherard's screen on enrolment Bayley II at 18 mo	No significant treatment effect on MDI or PDI or antorpometry. Similar ferritin levels in μg/L in cow's milk and low Fe formula groups (14.3 ± 0.8 and 13.1 ± 0.24), both significantly lower than high Fe formula group (21.7 ± 0.22, P < 0.01) Insufficient Hb data therefore % anemic unknown	Hb missing from most at beginning and end, 13% of subjects lost	Doubtful statistical power. Large loss of Hb data.

[1] Abbreviations: DBRCT, double-blind randomized controlled trials; IM, intramuscular; PDI, psychomotor development index; MDI, mental development index; IBR, Infant Behavior Record; Hb, hemoglobin.

Very few trials have followed the children after the treatment stopped, and this should be remedied.

Another problem is that many studies suffered from lack of statistical power; future studies not only should be randomized trials but also should have adequate sample sizes. A further problem is that when studies use a battery of tests and analyze several different scores from each test, there is a danger of spurious significant effects.

Last, there is an extreme lack of data on high risk children. In countries with large populations of high risk children, such as children with low birth weight, this may be important.

OVERALL CONCLUSION

It is clear that iron deficiency identifies children at concurrent and future risk of poor development. It is also clear that iron deficiency is usually associated with many pyschosocial, economic and biomedical disadvantages.

Anemic children <2 y old

There is no good evidence from RCT that short-term iron treatment benefits the development of anemic young children, but this has not been examined with rigor. The evidence of benefits from long-term trials is insufficient for drawing conclusions or extrapolating to other populations. There are insufficient RCT on the topic and larger studies are required in which both short- and long-term effects are assessed.

In several but not all studies, anemic children have failed to catch up to nonanemic children with iron treatment. This indicates that either their poor development is not caused by anemia or that the effect is irremediable by iron treatment alone, at least in the short term. This is surprising, considering the plasticity of child development; however, a vulnerable period exists for iodine deficiency and there is limited evidence that the first 2 y of life may be critical for protein-energy malnutrition (Grantham-McGregor and Ani 2001). It may be that improvements in the environment may be necessary for anemic children to catch up.

Anemic children >2 y old

Anemic older children also usually had poorer cognition and school achievement than did nonanemic children. They usually catch up in cognition with repeated testing and treatment but not in school achievement.

There are more RCT with this age group, and it was clearly shown that children benefited from iron treatment in four studies and a treatment benefit was highly likely in three others. However, two studies showed no effect. At present, the evidence for a beneficial effect of iron treatment on cognition in anemic older children is reasonably convincing, but it would be helpful to run one or two more rigorous RCT with detailed reporting of the results.

Preventive trials

Only limited data from preventive studies support a causal relationship. Preventing iron deficiency anemia can produce benefits to development but they are small and may be transient. The Chilean study should be followed up for longer to determine the implications of the benefits on the Fagan test.

It is difficult to come to unequivocal overall conclusions concerning the effects of iron deficiency in the first 2 y. There is some evidence of a causal relationship but this tends to be inconsistent. There are too few randomized trials of adequate size and appropriate analyses to make firm conclusions. More large randomized trials with anemic children are required before we can inform policy with confidence.

However, considering the stronger evidence of a causal relationship in school children it would be surprising if younger children were not also affected.

LITERATURE CITED

Agarwal, D. K., Upadhyay, S. K., Tripathi, A. M. & Agarwal, K. N. (1987) Nutritional Status, Physical Work Capacity and Mental Function in School Children. Nutrition Foundation of India, New Delhi, India.

Aukett, M., Parks, Y., Scott, P. & Wharton, B. (1986) Treatment with iron increases weight gain and psychomotor development. Arch. Dis. Child 61: 849–857.

Bruner, A. B., Joffe, A., Duggan, A. K., Casella, J. F. & Brandt, J. (1996) Randomised study of cognitive effects of iron supplementation in non-anaemic iron-deficient adolescent girls. Lancet 348: 992–996.

Cantwell, R. J. (1974) The long term neurological sequelae of anemia in infancy. Pediatr. Res. 342: 68.

Clarke, N., Grantham-McGregor, S. M. & Powell, C. (1991) Nutrition and health predictors of school failure in Jamaican children. Ecol. Food Nutr. 26: 1–11.

Colombo, J. (1993) Infant Cognition: Predicting Later Intellectual Functioning. Sage Publications, London, UK.

Czajka-Narins, D. M., Haddy, T. B. & Kallen, D. J. (1978) Nutrition and social correlates in iron deficiency anemia. Am. J. Clin. Nutr. 31: 955–960.

de Andraca, I., Walter, T., Castillo, M., Pino, P., Rivera, P. & Cobo, C. (1990) Iron deficiency anemia and its effects upon psychological development at pre-school age: a longitudinal study, pp. 53–62. Nestle Foundation, Lausanne, Switzerland.

Deinard, A., Gilbert, A., Dodds, M. & Egeland, B. (1981) Iron deficiency and behavioral deficits. Pediatrics 68: 828–833.

Deinard, A. S., List, A., Lindgren, B., Hunt, J. V. & Chang, P. N. (1986) Cognitive deficits in iron-deficient and iron-deficient anemic children. J. Pediatr. 108: 681–689.

Dommergues, M. P., Archambeaud B., Ducot, Y., Gerval, Y., Hiard, C., Rossignol, C. & Tchernia, G. (1989) Carence en fer et tests de developpement psychomoteur: etude longitudinale entre l'age de 10 mois et l'age de 4 ans. Arch. Fr. Pediatr. 46: 487–490.

Driva, A., Kafatos, A. & Salman, M. (1985) Iron deficiency and the cognitive and psychomotor development of children: a pilot study with institutionalised children. Early Child. Dev. Care 22: 73–82.

Fairchild, M. W., Haas, J. D. & Habicht, J. P. (1989) Iron deficiency and behavior: criteria for testing causality. Am. J. Clin. Nutr. 50: 566–574.

Florencio, C. (1988) Nutrition, Health and Other Determinants of lcademic Achievement and School-Related Behavior of Grades One to Six Pupils. University of the Philippines, Quezon City, Philippines.

Goswami, U. (1998) Cognition in Children. Psychology Press Ltd., East Sussex, UK.

Grantham-McGregor, S. M. & Ani, C. (2001) Undernutrition and mental development. In: Nestle Nutrition Workshop Series, Nutrition and the Brain. Raven Press, New York, NY (in press).

Grantham-McGregor, S. M., Fernald, L. C. & Sethuraman, K. (1999) Effects of health and nutrition on cognitive and behavioural development in children in the first three years of life. Part 1: Low birth weight, breastfeeding, and protein-energy malnutrition. Food Nutr. Bull. 20: 53–75.

Grantham-McGregor, S. M., Lira, P. I., Ashworth, A., Morris, S. S. & Assuncao, A. M. (1998) The development of low birth weight term infants and the effects of the environment in north-east Brazil. J. Pediatr. 132: 661–666.

Grantham-Mcgregor, S., Schofield, W. & Haggard, D. (1989) Maternal-child interaction in survivors of severe malnutrition who received psychosocial stimulation. Eur. J. Clin. Nutr. 43: 45–52.

Grindulis, H., Scott, P. H., Belton, N. R. & Wharton, B. A. (1986) Combined deficiency of iron and vitamin D in Asian toddlers. Arch. Dis. Child. 61: 843–848.

Groner, J. A., Holtzman, N. A., Charney, E. & Mellits, D. E. (1986) A randomized tial of oral iron on tests of short-term memory and attention span in young pregnant women. J. Adolesc. Health Care 7: 44–48.

Harahap, H., Jahari, A. B., Husaini, M. A., Saco-Pollitt, C., & Pollitt, E. (2000) Effect of an energy and micronutrient supplement on iron deficiency anemia, physical activity and motor and mental development in undernourished children in Indonesia. Eur. J. Clin. Nutr. 54: S114–S119.

Heywood, A., Oppenheimer, S., Heywood, P. & Jolley, D. (1989) Behavioral effects of iron supplementation in infants in Madang, Papua New Guinea. Am. J. Clin. Nutr. 100: 630–637.

Honig, A. S. & Oski, F. A. (1984) Solemnity: a clinical risk index for iron deficient infants. Early Child Dev. Care 16: 69–83.

Huda, S. N., Grantham-McGregor, S. M., Rahman, K. M. & Tomkins, A. (1999) Biochemical hypothyroidism secondary to iodine deficiency is associated with poor school achievement and cognition in Bangladeshi children. J. Nutr. 129: 980–987.

Hurtado, E. K., Claussen, A. H. & Scott, K. G. (1999) Early childhood anemia and mild or moderate mental retardation. Am. J. Clin. Nutr. 69: 115–119.

Husaini, M. L., Karyadi, L., Husaini, Y. K., Sandjaja-Karyadi, D. & Pollitt, E. (1991)

Developmental effects of short-term supplementary feeding in nutritionally-at-risk Indonesian infants. Am. J. Clin Nutr. 54: 799–804.

Idjradinata, P. & Pollitt, E. (1993) Reversal of developmental delays in iron-deficient anemic infants treated with iron. Lancet 341: 1–4.

Johnson, D. L. & McGowan, T. J. (1983) Anemia and infant behavior. Nutr. Behav. 1: 185–192.

Joos, S. K., Pollitt, E., Mueller, W. H. & Albright, D. L. (1983) The Bacon Chow study. A maternal nutritional supplementation and infant behavioral development. Child. Dev. 54: 669–676.

Lansdown, R. & Wharton, B. A. (1995) Iron and mental and motor behaviour in children. In: Iron: Nutritional and Physiological Significance. The Report of the British Nutrition Foundation's Task Force, pp. 65–78. Chapman and Hall, London, UK.

Levitsky, D. A. & Strupp, B. J. (1995) Malnutrition and the brain: changing concepts, changing concerns. J. Nutr. 125: 2245S–2254S.

Logan, S. (1999) Commentary: iron deficiency and developmental deficit—the jury is still out. Br. Med. J. 318: 697–698.

Lozoff, B. (1997) Does preventing iron-deficiency anemia (IDA) improve developmental test scores? Pediatr. Res. A 39: 136 (abs.).

Lozoff, B. (1998) Explanatory mechanisms for poorer development in iron-deficient anemic infants. In: Recent Advances in Research on the Effects of Health and Nutrition on Children's Development and School Achievement in the Third World: Policy Implications (Grantham-McGregor, S. M., ed.). Pan American Health Organization, Washington, DC.

Lozoff, B., Brittenham, G., Viteri, F. E. & Urrutia, J. J. (1982a) Behavioural abnormalities in infants with iron deficiency anemia. In: Iron Deficiency: Brain Biochemistry and Behavior (Pollitt, E. & Leibel, R. L., eds.). pp. 183–193. Raven Press, New York, NY.

Lozoff, B., Brittenham, G. M., Viteri, F. E., Wolf, A. W. & Urrutia, J. J. (1982b) The effects of short-term oral iron therapy on developmental deficits in iron deficient anemic infants. J. Pediatr. 100: 351–357.

Lozoff, B., Brittenham, G. M. & Wolf, A. W. (1987) Iron deficiency anemia and iron therapy: effects on infant developmental test performance. Pediatrics 79: 981–995.

Lozoff, B., Jimenez, E., Hagen, J., Mollen, E. & Wolf, A. W. (2000) Poorer behavioral and developmental outcome more than 10 years after treatment for iron deficiency in infancy. Pediatrics 105: E51.

Lozoff, B., Jimenez, E. & Wolf, A. W. (1991) Long-term developmental outcome of infants with iron deficiency. N. Engl. J. Med. 325: 687–694.

Lozoff, B., Klein, N. K., Nelson, E. C., McClish, D. K., Manuel, M. & Chacon, M. E. (1998) Behavior of infants with iron deficiency anemia. Child. Dev. 69: 24–36.

Lozoff, B., Klein, N. K. & Prabucki, K. M. (1986) Iron-deficient anemic infants at play. J. Dev. Behav. Pediatr. 71: 52–158.

Lozoff, B., Wolf, A. W. & Jimenez, E. (1996) Iron-deficiency anemia and infant development: effects of extended oral iron therapy. J. Pediatr. 129: 382–389.

Lozoff, B., Wolf, A. W., Urrutia, J. J. & Viteri, F. E. (1985) Abnormal behavior and low developmental test scores in iron-deficient anemic infants. J. Dev. Behav. Pediatr. 6: 69–75.

Lynn, R. & and Harland, P. (1998) A positive effect of iron supplementation on the IQs of iron deficient children. Pers. Individ. Differ. 24: 883–885.

Moffatt, M.E.K., Longstaffe, S., Besant, J. & Dureski, C. (1994) Prevention of iron deficiency and psychomotor decline in high-risk infants through use of iron-fortified infant formula: a randomized clinical trial. J. Pediatr. 125: 527–523.

Moock, P. R. & Leslie, J. (1986) Childhood malnutrition and schooling in the Teri region of Nepal. J. Dev. Econ. 20: 33–52.

Morley, R., Abbott, R., Fairweather Tait, S., MacFadyen, U., Stephenson, T. & Lucas, A. (1999) Iron fortified follow on formula from 9 to 18 months improves iron status but not development or growth: a randomised trial. Arch. Dis. Child. 81: 247–252.

Oski, F. A. & Honig, A. S. (1978) The effects of therapy on the developmental scores of iron-deficient infants. J. Pediatr. 92: 21–25.

Oski, F. A., Honig, A. S., Helu, B. & Howanitz, P. (1983) Effect of iron therapy on behaviour performance in non-anaemic iron-deficient infants. Pediatrics 71: 877–880.

Owen, G. M., Lubin, A. H. & Garry, P. J. (1971) Pre-school children in the United States: who has iron deficiency? J. Pediatr 79: 563–568.

Palti, H., Pevsner, B. & Adler, B. (1983) Does anemia in infancy affect achievement on developmental and intelligence tests? Hum. Biol. 55: 183–194.

Palti, H., Meijer, A. & Adler, B. (1985) Learning achievement and behavior at school of anemic and non-anemic infants. Early Hum. Dev. 10: 217–223.

Pollitt, E., Gorman, K. S., Engle, P. L., Martorell, R. & Rivera, J. (1993) Early supplementary feeding and cognition. Monogr. Soc. Res. Child. Dev. 58: 1–99.

Pollitt, E., Hathirat, P., Kotchabhakdi, N. J., Missell, L. & Valyasevi, A. (1989) Iron deficiency and educational achievement in Thailand. Am. J. Clin. Nutr. 50: 687–697.

Pollitt, E., Leibel, R. L. & Greenfield, D. B. (1983) Iron deficiency and cognitive test performance in preschool children. Nutr. Behav. 1 137–146.

Pollitt, E., Saco Pollitt, C., Leibel, R. L. & Viteri, F. E. (1986) Iron deficiency and behavioral development in infants and pre-school children. Am. J. Clin. Nutr. 43:555–565.

Pollitt, E., Soemantri, A. G., Yunis, F. & Scrimshaw, N. S. (1985) Cognitive effects of iron-deficiency anaemia. Lancet 19: 158.

Popkin, B. & Lim-Ybanez, M. (1982) Nutrition and school achievement. Soc. Sci. Med. 16: 53–61.

Ramdath, D. D., Simeon, D. T., Wong, M. S. & Grantham-McGregor, S. M. (1995) Iron status of schoolchildren with varying intensities of *Trichuris trichiura* infection. Parasitology 110: 347–351.

Roncagliolo, M., Garrido, M., Walter, T., Peirano, P. & Lozoff, B. (1998) Evidence of altered central nervous system development in infants with iron deficiency anemia at 6 mo: delayed maturation of auditory brainstem responses. Am. J. Clin. Nutr. 68: 683–690.

Seshadri, S. & Gopaldes, T. (1989) Impact of iron supplementation on cognitive functions in pre-school and school-aged children: the Indian experience. Am. J. Clin. Nutr. 50: 675–686.

Soemantri, A. G. (1989) Preliminary findings on iron supplementation and learning achievement of rural Indonesian children. Am. J. Clin. Nutr. 50: 698–702.

Soemantri, A. G., Pollitt, E. & Kim, I. (1985) Iron deficiency anemia and educational achievement. Am. J. Clin. Nutr. 42: 1221–1228.

Soewondo, S. (1995) The effect of iron deficiency and mental stimulation on Indonesian children's cognitive performance and development. Kobe. J. Med. Sci. 41: 1–17.

Soewondo, S., Husaini, M. & Pollitt, E. (1989) Effects of iron deficiency on attention and learning processes in pre-school children: Bandung, Indonesia. Am. J. Clin. Nutr. 50: 667–674.

Waite, J. H. & Neilson, I. L. (1919) Study of the effects of hookworm infection upon the mental development of North Queensland school children. Med. J. Austr. 1: 1–8.

Walker, S. P., Grantham-McGregor, S. M., Himes, J. H., Williams, S. & Duff, E. M. (1998) School performance in adolescent Jamaican girls: associations with health, social and behavioural characteristics, and risk factors for dropout. J. Adolesc. 21: 109–122.

Walter, T., Kovalskys, J. & Stekel, A. (1983) Effect of mild iron deficiency on infant mental development scores. J. Pediatr. 102: 129–522.

Walter, T., de Andraca, I., Chadud, P. & Perales, C. G. (1989) Iron deficiency anemia: adverse effects on infant psychomotor development. Pediatrics 84: 7–17.

Wasserman, G., Graziano, J. H., Factor-Litvak, P., Popovac, D., Morina, N., Musabegovic, A., Vrenezi, N., Capuni-Paracka, S., Lekic, V., Preteni-Redjepi, E., Hadzialjevic, S., Slavkovich, V., Kline, J., Shrout, P. & Stein, Z. (1992) Independent effects of lead exposure and iron deficiency anemia on developmental outcome at age 2 years. J. Pediatr. 121: 695–703.

Wasserman, G. A., Graziano, J. H., Factor Litvak, P., Popovac, D., Morina, N., Musabegovic, A., Vrenezi, N., Capuni Paracka, S., Lekic, V., Preteni Redjepi, E., Hadzialjevic, S., Slavkovich, V., Kline, J., Shrout, P. & Stein, Z. (1994) Consequences of lead exposure and iron supplementation on childhood development at age 4 years. Neurotoxicol. Teratol. 16: 233–240.

Watkins, W. E. & Pollitt, E. (1998) Iron deficiency and cognition among school-age children. In: Nutrition, Health, and Child Development. Research Advances and Policy Recommendations (Grantham-McGregor, S. M., ed.), pp. 179–197. Pan American Health Organization, Tropical Metabolism Research Unit of the University of the West Indies, and The World Bank, Washington, DC.

Webb, T. E. & Oski, F. A. (1973) Iron deficiency anemia and scholastic achievement in young adolescents. J. Pediatr. 82: 827–830.

Williams, J., Wolff, A., Daly, A., MacDonald, A., Aukett, A. & Booth, I. W. (1999) Iron supplemented formula milk related to reduction in psychomotor decline in infants from inner city areas: randomised study. Br. Med. J. 318: 693–698.

Yu, G. S., Steinkirchner, T. M., Rao, G. A. & Larkin, E. C. (1986) Effect of prenatal iron deficiency on myelination in rat pups. Am. J. Pathol. 125: 620–624.

DISCUSSION

Participants: Lozoff, Pollitt, Grantham-McGregor, Beard, Haas, Habicht, Schultink, Sazawal, Allen, Lynch.

Dr. Lozoff: First, it is very sobering. I have been working on iron deficiency for 25 years now, and really trying to do good studies. After all this effort, we still cannot give definite answers.

Second, I was pleased to see that we did not buy in totally to the reasoning that if we could show that we could correct it, that would be the proof that iron was the cause. There was a period in the literature where, if you did not show that you corrected it with the iron treatment, people would conclude iron deficiency could not be the cause, because you correct the anemia, you correct the development. I was trying to think of an analogy for this. Take tuberculosis. A child comes into the hospital with tuberculosis meningitis. You give antituberculosis drugs. The child still has neurological sequelae. Well, the child might have had a neglectful or a misguided family, but the child nonetheless had tuberculosis meningitis. Tuberculo-

sis has been cleared from the body. There was unequivocally a treatment response, but you still have a long-lasting effect. So, it is just fallacious thinking to hold as the only criterion that you reverse it with iron treatment and I was pleased that Dr. McGregor made that so clear. It would be great if we could do it, because that would settle the whole thing, but the reverse does not hold true.

Third, I want to highlight the unresolved issues as I see them. We have not really been looking at specific central nervous system functions. As we do that, we are going to have to revisit this iron deficiency without anemia. We have got to revisit these more sensitive measures. Similarly, we have got to revisit everything about what treatment affects or does not affect as you get more measures that make sense in terms of what iron is doing.

Finally, one comment about the magnitude of effect. The Bayley for infants and toddlers has been the only measure that we really have had. So, it has been used not just in iron deficiency but in a whole host of early biological risks. The magnitude of the differences that we have observed between anemic and nonanemic kids on the Bayley is of the same order of magnitude that is widely accepted as being clinically relevant and important, whether it is low birth weight or cocaine or alcohol or any of those things. I shared Dr. Pollitt's concern about the Bayley Scales, but I wanted to be sure that we at least gave that information to everybody.

Dr. Pollitt: The data have shown differences in the Bayley Scales that were based on children who were 18 mo or older. They probably do have significant value. I am arguing that it is at the younger ages that the data are not good. Second, we also have to understand that a 3–5 point difference in the Bayley Mental Development Index may actually represent only about 1–2 wk of difference in development. In other words, a 3 or 5 point difference, as has been reported, does not correspond to a difference of 3 mo of development. It is equivalent to 15 d.

Dr. Lozoff: Six to 19 points difference in your study?

Dr. Pollitt: Nineteen points actually represents about a month and a half.

Dr. Grantham-McGregor: That is quite a lot in a young infant's life. It is not like a month and a half in my life.

Dr. Beard: I would like all three of the presenters to perhaps argue the point, if you would, about the reversibility of these issues.

Dr. Lozoff: One of the beauties of having had the advances in neuroscience that we have had is that you start to talk about different functions that iron might play. Then you start being able to say that you have some basis to think some should respond to treatment and some maybe will not respond to treatment. It is really going to depend on what function we are going to talk about.

If the formation of myelin is thrown off, maybe that is not going to reverse so quickly. There is a whole cascade of events that are rapidly happening. Then you talk about a neurotransmitter—which might reverse right away. In development, though, if it is a neurotransmitter where they are laying down the paths, even though you correct the neurotransmitter levels, you could still have some effects because you altered the developmental course. I think that we could start being able to make sense rather than simply saying no response or response, get back to some mechanistic hypotheses.

Dr. Beard: Dr. McGregor, do you want to bring up the protein-energy malnutrition literature and hypomyelination?

Dr. Grantham-McGregor: Yes, I get a bit nervous when we talk about irreversible changes. When you think about severely malnourished children in a developed country, if they

have a good environment, they show quite a lot of improvement, if not total catch up. These are children who are really bad. The adoption studies suggest that an enormous amount of catch up can take place. So, it may be irreversible in the environment they are in, but it may not be irreversible elsewhere.

Dr. Lozoff: The plasticity of the brain is so important. In fact, you do not have to prove that they are long-lasting effects to have a causal relationship. You can have a causal relationship and it can be ameliorated by something else.

The other thing about reversibility or irreversibility, until you systematically try treating it earlier, treating it more, treating it longer, or treating it in different conditions, you do not know for sure. There is really uncertainty.

Dr. Pollitt: I would like to make two comments. One is, you made the point about the case of tuberculosis and having permanent effects in some cases. A contrast is the case of low birth weight. Low birth weight represents one of the most important developmental risk factors that there is. That is true if you are talking about low-socioeconomic groups and the range of 2000–2500 g. That is not the case in a high-socioeconomic group. In other words, the probabilities of actually finding a developmental delay, a learning disability, or an IQ deficit in a low-birth-weight baby born to an upper-middle-class family in the United Kingdom or the United States are very low. It is the same risk factor in a high- or a low-socioeconomic condition. It is just that the probability of full rehabilitation in one group is much higher than in the other group.

Second, the information on hypomyelination and on dopamine, γ-aminobutyric acid, and serotonin is tremendously valuable for understanding the neurobiological effects of iron deficiency. Those kinds of effects may actually produce some functional consequences. That is very different from saying that higher cortical or higher cognitive functions are going to be affected by any of these types of changes. In other words, although it is valuable information, that information by itself is not going to tell you anything about the potential that a child has to do well in school.

Dr. Lozoff: Some of the cognitive measures now available for infants—in recognition memory and reaction times—do predict those more integrative cognitive functions later on better than anything we had before.

Dr. Haas: I would like to address one of the issues that Dr. McGregor raised at the very end—that it is somewhat of a dilemma that you are finding these significant effects in schoolchildren and you are not seeing the effects in preschool children, which defies every logic of development. My question has to do with something that Dr. Pollitt ended with–the lack of sensitivity of the Bayley at 12 mo. Could it just be that the tests that you have for preschool children in general, across all these different domains, are still relatively insensitive compared with the ones that you are using in schoolchildren?

Dr. Grantham-McGregor: I think undoubtedly the data are not a great measure. We are all saying that we cannot show the effect. We had one randomized controlled trial and we showed the effect. We should do a couple more with bigger samples.

Now, I think this is an enormous issue. Is it not ethical? I would love to do a randomized controlled trial in the first 2 y of life to see whether there is an effect. I would not do it in the 1st y, because it is really difficult to show an effect on the Bayley in the 1st y. To use these other measures is a very welcome idea. Can we do a randomized controlled trial and settle the question? Is it unethical? That is what we need to address.

Dr. Habicht: You are proposing a preventive trial; is that right?

Dr. Grantham-McGregor: No, those are difficult to do because you need 1000 kids, you need to show that you are going to get anemia, and the effects in the 1st y are not as good. It is much better to do it in the 2nd y and to do it with anemic kids.

Dr. Habicht: One lesson from this general work in the area of development of child learning is that there are real effect modifiers. In other words, in one situation you will see an effect and in another situation you will not. One of the ways of doing this is starting with children who are low birth weight or small-for-gestational age or something of this sort.

Dr. Grantham-McGregor: Your study children were not like that, were they, Dr. Lozoff?

Dr. Lozoff: The mean hemoglobin in all these studies that we have been talking about is 96 g/L. In community study after community study in these populations in Latin America, with kids who are otherwise healthy, the lowest hemoglobin that I have ever seen in a community study is 78 g/L.

Dr. Sazawal: We are proposing a design to basically identify anemic children and then randomly assign them to get iron or not get iron.

Dr. Grantham-McGregor: You would obviously have to have a cut-off point. If you found children below a certain level, you would eliminate them from the study on ethical grounds.

Dr. Habicht: You are using anemia as a proxy for iron. I mean, you could be looking at other parameters of iron.

Dr. Grantham-McGregor: I would screen for iron-deficiency anemia.

Dr. Habicht: So, in other words, it is moderately severe.

Dr. Grantham-McGregor: Not severe, obviously not. I think everybody is agreed that children should not be severely anemic.

Dr. Schultink: How long would you need to do the trial? How long would you keep the placebo group untreated?

Dr. Grantham-McGregor: Dr. Pollitt's trial ran for 4 mo.

Dr. Habicht: The larger issue seems to be a catch-22. We do not have very firm evidence of detrimental effects in mild anemia. If ultimately we believe that mild anemia really is not all that bad and that we should be focusing much more on severe anemia where we have clear effects, there are a number of implications relative to programs and policy. The major implication relative to progress in this field is that we will then be able to look at these milder levels to decide whether or not we are really right.

Dr. Allen: I have a quick question. What do you do with anthropometric data in all of these studies? Has anybody looked at the association between current anthropometric status and the effect of this supplementation?

Dr. Lozoff: I look at it. I do it either as a main effect outcome or as a covariate. Pretty much nothing.

Dr. Grantham-McGregor: Did you have malnourished populations.

Dr. Lozoff: No, we had well-nourished kids and it is a different question.

Dr. Beard: An issue that we really have not addressed is the age dependency of any of these effects. Dr. Pollitt and Dr. Lozoff in earlier comments noted that there are certainly developmental issues at different times where plasticity may be a concern. Everybody needs to recognize that lots of the neurochemistry work that we do and other people have done is in animals that are essentially young adults. These are not animals that were made iron deficient during early brain development and it is a very reversible phenomenon in terms of only manipulating iron. So neurochemistry, independent of development, is sensitive to iron status.

Dr. Allen: Is demyelination reversible?

Dr. Beard: It does not appear to be—do you mean demyelination or did you mean inhibition of myelination?

Dr. Allen: No, demyelination. I mean, the vitamin B-12 literature leads you to believe that it is not reversible.

Dr. Lynch: Maybe partially, but not completely, no.

Dr. Lozoff: I want to make a comment about the myelin. You could ultimately have, for example, in a hypomyelinated condition, that they catch up, but developmentally, things could not have been on track. The myelinated axons have to be there when things are coming on line. So, you could have something that ultimately caught up at the brain level and still have had long-lasting effect on development.

Dr. Beard: This is the point that Dr. Pollitt was making before, in terms of environmental deprivation. To get some of these pathways to actually be put into place, you have to have the appropriate interactions with the environment.