Timing in Prenatal Nutrition: A Reprise of the Dutch Famine Study

Mervyn Susser, M.B., B.Ch., D.P.H., F.R.C.P. (E) and Zena Stein, M.A., M.B., B.Ch.

Interest in prenatal nutrition as a subject of serious scientific study has fluctuated during the past century as it seemed to produce results that did or did not have importance for the developing fetus. We recount some of that history, not for its own sake, but to make both a general and a particular epidemiological point.

The general point is that those hypotheses that stand the tests of time and replication are narrowly focused and, in turn, are matched by rigorous research design that is equally focused. One can add, in counterpoint, that the effects of human nutrition cannot be fully apprehended if research does not range outside the molecular and physiological and venture to explore the human realm of organic and systemic impairment, functional disability, and social handicap.

The particular point is that the effects of prenatal nutrition are obscured if the nutritional exposure is not closely defined in terms of developmental timing (as well as duration, severity, and type). We concentrate here on this particular point, and illustrate it in some detail by a single study, our own work on the Dutch Famine of 1944–1945.

Timing and the History of Prenatal Nutrition Research

Interest in prenatal nutrition waxed in the late 19th century with the movement led by Budin in France, Ballantyne in Britain, and others toward promoting healthy maternity and infancy, especially among the poor majority. When animal experimentalists failed to detect important developmental effects, this interest waned. It was restored by the efforts of leading nutritionists such as Boyd-Orr in Britain as the Great Depression lingered on through the 1930s. Hunger beset the poor of the western world almost as in preindustrial times, despite its developed industrial capitalist economies. Several studies in Canada, the United States, and Britain reported favorable effects of food or vitamin supplementation on the outcome of pregnancy.

These results found vindication in the consequences both good and ill of World War II. In a besieged Britain, where rationing distributed food evenly to all classes, maternal and child health were not only sustained, but improved. The adverse condition of acute deprivation was even more striking. The effects of starvation on the size of infants at birth were undeniable in the 18-month Leningrad siege, in the 6-month Dutch famine, and in the acute food shortage in Wupperthal, Germany during the postwar chaos of 1945.

However, the findings of the late 1950s once more assailed confidence about what had seemed self-evident. Despite the favorable but inconclusive result of a protein and vitamin supplementation trial, the iconoclasts were two major observational studies that undermined the previous understanding. Neither study demonstrated any persuasive effects of nutrition during pregnancy on the size of the newborn infant. The theory gained favor that the fetus was a “perfect parasite,” battenning on maternal provender while it was afforded protection from nutritional damage that might be inflicted on the mother.

Interest in prenatal nutrition was reawakened by a revival of the theory of critical periods in development. This theory postulates that particular stages are critical for aspects of fetal development in that, if they are not successfully negotiated, the damage may be irremediable and the course of development irreversibly changed.

The theory was given new force by the animal experiments and interpretations of McCance and Widdowson. They postulated a critical period early in development when “the regulating centers
of the hypothalamus are being coordinated . . . with rate of growth." They were struck by the fact that runts in pig litters, born small because of deficient placental nutrient supply, were rehabilitated by adequate postnatal nutrition in terms of growth rate but never attained the size of their littermates.

In the late 1960s, the theory of critical periods was further strengthened by several researchers concerned with development of the brain.26-28 Winick and his colleagues studied rats severely deprived of protein at different stages of early brain development and subsequently rehabilitated. In the early phase when the cells were rapidly multiplying (hyperplasia), it could be inferred from biochemical measures of DNA and RNA that the number of cells was irreversibly depleted. Later, in the phase when the cells were no longer hyperplastic but enlarging (hypertrophy), the slowed growth in cell size that ensued could be reversed.

These illuminations of the importance of timing were graphically illustrated by Davison and Dobbing.29 Their emphasis was on the confusion caused by extrapolations from animal experiments that ignored differences between species. Experimental nutrition interventions took little account of the fact that the timing of the same developmental phases of the brain in relation to birth was not the same across species (Figure 1). Consequently, results were inconsistent and causal inferences weak. The argument was that if comparisons and extrapolations are to be made across species, one must respect developmental milestones of the brain. Thus, intervention must be accurately timed to fit phases of cell and organ development (and, implicitly, dissociated from the social, psychological, and physiological species-specific event of birth).

Eventually, we elaborated these prenatal phases for epidemiological purposes in human studies of prenatal nutrition. The framework we arrived at was divided into periconception, embryonic, fetal (early and late), and early postnatal phases.21 Periconception is defined by the process of ovulation and the establishment of pregnancy, a 3-week period which begins after menstruation (two weeks before ovulation) and ends with implantation, 1 week after fertilization. The embryonic phase is conterminous with the embryo and is defined by the period beginning about 8 weeks after implantation and ending as specific organs begin to form, at a developmental age of about 9 weeks (and a gestational age of 11–12 weeks). The fetal phase covers the remainder of gestation. For some purposes it can be usefully divided into three: an early fetal phase continues to viability at 20–22 weeks of gestation, a middle fetal phase to about 34–36 weeks, and a late fetal phase to term. The definition of the early postnatal phase varies with the topic of study (from the first week for mor-


tality, to more than a year for the completion of brain development, to 2 years before measured intelligence can be described with reliability).

The Development of a Hypothesis:
Prenatal Nutrition and Mental Performance

Working from the foundation of the work on developmental timing described above, we designed our studies of the effects on development of prenatal nutrition. We had the advantage of knowing that the Dutch famine of 1944–1945 had not only lowered birth weight as noted above, but that this effect was, to all appearances, a result of exposure during the third trimester of gestation.8 Although our studies ultimately addressed several aspects of development and its disorders, the initial focus was on mental performance.

In this regard, the proceedings of a conference on early brain development and cognition held in Boston in 196722 were something of a landmark. Biochemical, physiological, pathological, psychological, and epidemiological studies were assem-
bled. The results convinced many that early nutritional deprivation caused irreparable damage to the brain and consequently to cognition. Reports of the proceedings were featured prominently in leading newspapers. Hundreds of millions of people, especially but not only in the less developed world, were thought to be at risk of, if not affected by, retarded mental development.

Our prime interest at the time had been the epidemiology of mental retardation. We could not neglect such a momentous issue. We began with a close examination of the literature. To a critical epidemiological eye, it was not conclusive. The biochemical results of animal experiments were persuasive, but left room for alternative explanations. While several animal studies of severe food deprivation were suggestive, the results were not consistent. None were sufficiently comprehensive to demonstrate the whole sequence leading from irreversible effects on brain development to persistent disorders of brain function.

We decided to commit ourselves to testing the hypothesis at issue. Our model of developmental timing was less refined than that just outlined but observed the same principle.

The Design of a Study

The necessity for definitive studies, on the model of T. H. Huxley’s “crucial experiment,” seemed clear to us. Experience had not yet tempered our understanding with the knowledge that even crucial studies are not the end of a story. A causal sequence cannot be taken as established until the result repeatedly and consistently survives rigorous tests that might disprove the hypothesis.

That our own studies in this vein should be epidemiological seemed self-evident. This was not solely because we were epidemiologists. Human studies alone could avoid extrapolating from animal experiments, and such extrapolations are especially tenuous in the matter of intelligence. Only human population-based epidemiological studies could establish the causal role of environmental factors. They were population-based in the sense of having denominators as well as numerators and, hence, the capacity to compare rates of those exposed and unexposed or affected and unaffected.

Studies of the effects of nutritional deprivation in free-living human populations face immense difficulties in adequately measuring exposure in the form of food intake and in disentangling that exposure from its intimate involvement with poverty. It is also difficult to assess the many other associated factors capable in themselves of producing the effects we seek to attribute to undernutrition.

The most rigorous design available to epidemiologists is the randomized controlled trial, and under many circumstances the gain in specificity and precision outweighs the loss in generalizability. The design is applicable in humans only under severe experimental restrictions, however. To be ethical, an intervention must be benign. One of our two studies used this design. This first study, begun in 1968, was one of two contemporaneous randomized trials in which the intervention was nutritional supplementation, and which alone had sufficient numbers and statistical power to test the hypotheses at issue.

Our purpose was to prevent fetal growth retardation and subsequent retarded development. We undertook this study in New York. We used our understanding of C. A. Smith’s work on the Dutch famine to select the third trimester as the targeted phase of gestation. The selected intervention was in line with the wisdom of the time, in which the limiting nutrient was thought to be protein, and we devised a supplement that was essentially a highly concentrated protein. No favorable effects on fetal growth nor on infant development were achieved.

The second study was observational, not experimental. Here we aimed to turn to human advantage the Dutch famine of 1944–1945, a deliberate creation of the Nazi occupation. The famine met the criteria for a natural experiment. Thus, it was an observational study with three characteristic elements of such studies: the exposure was a sharp measurable change from the norm, and of short duration; it occurred in definable environmental circumstances; and definable populations were exposed. A necessary condition for establishing causal sequences is the observation of changes in outcome in undeniable time sequence following measurable changes in circumstances (or “exposure”), and natural experiments provide that condition.

The Dutch famine was unique among famines in the precision with which many elements vital to strong research design could be described. Although the study was undertaken more than 20 years after the event, it was possible to account for the entire population affected, to measure the severity of exposure, and to define many outcomes. Equally well-measured comparison groups and controls for potential confounders could also be identified. Long-term follow-up could be managed with facility.

The famine was the direct result of a transport embargo the Nazi occupation imposed on occupied western Holland. Dutch transport workers had mounted a strike in anticipation of a British advance northward across the Rhine that had failed. In the coldest winter in a half-century, ice blocked canal traffic to exacerbate the embargo on road and rail. The famine was remarkably circumscribed in both time and place. In time, it lasted 6 months from November 1944, a few weeks after the embargo be-
gan, to May 7, 1945 when Holland was liberated from the Occupation. In place, it was marked out by the delimitation of the embargo to the west Holland region; the regions of north and south Holland were exempt. The cities were worst affected; they could not draw as easily on the surrounding farms as rural people could. We therefore confined the study to the cities with populations greater than 40,000 people in the three regions of the west, north, and south (Figure 2).

This demarcation of exposure was enhanced by records of the weekly food rations allocated to each individual throughout the war, including the Nazi Occupation. The rations available to west Holland for each 3-month period from 1941 to 1945 are shown in Figure 3. As a result of the well-defined boundaries of the famine, individual exposure could be related to gestational age, provided that date and place of birth were known.

The exposed and comparison populations (national, regional, and local) could be reconstructed both from national vital statistics and from registers maintained by all local authorities. These kept track of the composition and the migration of all households and their members. The two sources provided data that were in some respects duplicated and allowed validity checks one against the other, and that were in other respects complementary, each filling gaps in the alternative source.

Outcome measures of cognition were first sought. We looked for national data on mental performance that would also give the date and place of birth for the person tested. We found this information in the records of military induction, which was compulsory at the time for all 19-year-old men. These computerized records alone were sufficient to test the primary hypotheses relating to mental performance, since they allowed us to construct birth cohorts by date and place of birth, and hence by famine exposure and outcome at 19 years of age.26

The data for 3 years of births, 1944–1946, were assembled in the form of cohorts born before, during, and after the famine. Further, and of central importance to tests of hypotheses based on phases of prenatal development, these cohorts could be arranged according to the stage of gestation at which they experienced famine exposure (Figure 4).

We did not let the matter rest there. The two decades in the life course between exposure and outcome were an interval empty of data in the military induction records. This raised concern because without such data, we could not retest the finding of intrauterine growth retardation in the famine-exposed cohorts postulated from Sindram and Smith.7 Nor could we be sure that selective attrition by mortality and migration might not confound the results. For these reasons, we decided also to collect the available hospital records for deliveries, as well as records for all births (i.e., fertility), deaths, and mi-

Figure 2. The Netherlands, showing cities with population ≥40,000, by size and famine exposure. Adapted with permission from: Stein ZA, Susser MW, Saenger G, Marolla F. Famine and human development: The Dutch Hunger Winter of 1944/45. New York: Oxford University Press, 1975.


grations. These data made it possible not only to fill many gaps of inference but also to describe a range of famine effects otherwise inaccessible to study.

Effects of Prenatal Famine Exposure on Development

We can now proceed to assess the effects of famine on those developmental outcomes measurable in the data collected (Figure 5). We shall order them in the sequence in which they appear in the life course: fertility; maternal weight; birth in terms of prematurity; infant weight and growth rate; mortality from birth through two decades up to 19 years; in survivors at age 19 years, physique, especially obesity, physical and mental performance, and mental disorders; in the fourth decade, mental disorder, especially schizophrenia; and, finally, in the second generation, the relation of birth outcome to various aspects of maternity. In accord with our starting postulates about the centrality of developmental stages, each outcome will be considered in relation to the timing of prenatal famine exposure. To clarify the presentation, the figures that follow illustrate the results for the famine cities only, omitting the two control areas; the whole of the data are reported in Stein et al.27

Fertility

During the famine, fertility declined dramatically (Figure 6). The decline was unique to the time and place of the famine. It was most severe among those with the least resources, to the extent that in the affected cohorts the proportions in social classes were substantially altered, with a deficiency among the poorest (Figure 7). This was a periconceptional effect, since recovery was immediate once the famine ended.

Figure 5. The Dutch Hunger Winter, 1944/1945. Scheme of famine effects on successive developmental stages culminating in young adults aged 19 years. Taken from: Susser MW. The logic in ecologic part II. The logic of design. Am J Pub Health 1994 (in press). Reprinted with permission from the Am J Pub Health, Washington DC.
Maternal Weight

After the frugal war rationing of the years before the famine, women carried no surplus weight at the onset of the famine. A measure of postpartum weight was available to the study from hospital records in Rotterdam. A sharp loss of weight followed by a quick recovery was very obvious in women exposed to the height of the famine in the third trimester (Figure 8). A subsequent analysis from Amsterdam confirms these results. Among those women whose pregnancy coincided with the height of the famine, maternal weight at term was less than at early or mid-term registration for prenatal care. The correlation of birth weight with maternal weight held below a dietary threshold only, which suggests that birth weight was mediated through maternal weight but only below that threshold.

Birth Outcomes

Length of Gestation

Births exposed in the third trimester at the height of the famine had a gestation somewhat shorter than average. Those exposed in the first trimester were sharply affected (a result retested in more detailed data). However, with this outcome, inference must be guarded. The northern region experienced a similar shortening in the face of food deprivation of lesser order. Hence a cofactor like epidemic infection of the mothers must be entertained.
Birth Weight

The effect of third trimester exposure on birth weight is unequivocal (Figure 9). Rehabilitation was rapid, just as for maternal weight. Third trimester famine exposure had a much more modest effect on preterm delivery, as mentioned above, which points to growth retardation as the underlying process in lowered birth weight. Placental weight marched in parallel with birth weight, if not in perfect step. Length and also head circumference (Figure 10) were diminished but to a lesser degree. In all these respects the famine produced effects on size consonant with acute nutritional deprivation in the rat experiments of Winick et al. and thus validated the expectation of similar effects on the brain in the famine study.

Mortality

Mortality among older adults living in the larger cities during the famine was considerable. The major impact in children was seen in the summer epidemic immediately after the famine ended. This epidemic, of entirely obscure origin at the time, decimated the infant population. The answer to its enigmatic pattern lay in large part in the preceding prenatal exposure to famine. Analysis by birth cohort, enjoined by the developmental hypothesis and the historical cohort design of the famine study, uniquely demonstrated the overwhelming effect of the famine on the epidemic.

Stillbirths increased sharply and solely with first trimester exposure to the famine (Figure 11). A part of the increase may have been due to interaction with an accompanying factor (most likely infection) since at the same time small increases occurred also in the two control regions.

First-week death rates also rose sharply with first trimester exposure (again with the appearance of an interaction with an accompanying factor that was most likely infection). In the live newborn, however, an equally powerful effect on first-week deaths was seen with third-trimester exposure. The first week marks a transition in that no further mortality effects of first-trimester exposure appear at later ages.

Deaths up to 90 days constitute by far the greatest part of the unequivocal contribution of the famine to the summer epidemic of infant deaths in 1945. They are related solely to third-trimester exposure to famine, with maximum effect having occurred at 30–90 days (Figure 12).

The cohorts exposed to famine in the third trimester showed an increased mortality to some degree from 3 months to 18 years. Because there was a comparable rise in the north and south, this could not be attributed solely to prenatal famine exposure.

Physique

Height

The military induction records on 19-year-old men yielded data on height and weight. Height was clearly responsive to postnatal environment, in that a distinct trend toward increase was present over 4
years of observation. Within this trend, however, we were unable to demonstrate a famine effect; it did not differ between exposed and unexposed cohorts.

**Weight**

Weight showed a trend similar to height. Obesity, in contrast, did exhibit famine effects. Exposure early in gestation (the first 4 months) produced an excess of obesity, perhaps related to deranged hypothalamic function, or else to the rapid rehabilitation of maternal weight in late pregnancy or, more likely, to both together. Exposure in the latter half of gestation produced a reduction in obesity. A subsequent experimental study, designed to test these findings, provided congruent results. Pregnant rats, first starved and then nutritionally rehabilitated (analogous with first-trimester famine exposure) produced an excess of obese offspring. Although the military induction data refer only to men, the results have since been replicated in a cohort of women exposed to the famine in Amsterdam. Those exposed early in gestation had an excess of obesity and those exposed late, a deficiency.

**Mental Performance**

The hypothesis which initiated this study was tested in a total sample of some 120,000 military inductees drawn from six cities in the famine region and 10 from the two comparison regions. The hypothesis rested on a premise of brain cell depletion without organ damage.

A second premise was that any large and prevalent socially determined environmental factor in mental performance, such as deprivation of macronutrients in nonfamine times, must reflect the marked social distribution of mild mental retardation and mental performance generally. Severe mental retardation and frank organ damage to the brain exhibit only minor social gradients, whereas measured intelligence and mild retardation exhibit marked gradients. In accord with both assumptions, the appropriate outcome measures were therefore tests of mental performance and frequencies of mild mental retardation.

Five test scores and one global mental assessment, as well as a clinical diagnostic assessment, were available for each young man examined in the induction process. The tests were sensitive enough to demonstrate and differentiate subtle effects of family size, birth order, and height on performance. Because of the changed social composition of the cohorts conceived at the height of the famine, with a marked deficiency of the members of the lowest classes (see Figure 7), it was necessary to control closely for social class to avoid the spurious appearance of a beneficial effect on mental performance in those cohorts. Ultimately, all tests yielded the same essentially null result as the Raven test scores (Figure 13).

**Physical Impairment**

Birth defects were the main manifestation of physical impairment. The data on the anomalies ob-
Figure 13. Raven scores and prenatal famine exposure: mean scores (high scores are worse) manual and non-manual occupational class of father in successive urban west Holland birth cohorts (January 1944 to December 1946) at military induction at 19 years of age. Adapted with permission from: Stein ZA, Susser MW, Saenger G, Marolla F. Famine and human development: the Dutch Hunger Winter of 1944/45. New York: Oxford University Press, 1975.

served were drawn in part from mortality statistics based on death certificates, but chiefly from examination of male survivors at military induction. Among a large array, the sole finding with a clear relation to famine was an excess of central nervous system defects with first trimester famine exposure. This could not be attributed to any specific deficiency from the available data. It was of course a harbinger of an important future discovery, namely, that periconceptional folate deficiency is a cause of neural tube defects.

Mental Disorder

At the age of 19, no mental disorders, in either the induction record or the national register then being compiled, were at first recognized as having an obvious relation to famine exposure. For uncommon, major disorders such as schizophrenia and depression, we assumed that an adequate test had to wait on the accumulation of numbers with the aging of the population.

Very recently, in the same birth cohorts now nearing the age of 50, an association has been demonstrated for schizophrenia. Famine exposure in early gestation carried a twofold risk for women. In a larger set of data subsequently acquired, the risk in men reached much the same level. The more narrowly both famine exposure and schizophrenia are defined, the stronger the association proved to be. It is apparent also in the induction data for 19-year-old men that with exposure in early gestation, the prevalence of conditions antecedent to schizophrenia—in particular, schizoid personality—was raised in the same birth cohorts. These effects could be the result of embryonic exposure, since the raised frequencies exactly coincide in time with the excess of neural tube defects in two monthly birth cohorts conceived at the height of the famine.

The Second Generation

Intergenerational effects in reproduction have in recent years attracted scientific interest. At least two studies have found a modest association between the birth weights of mothers and their offspring. Whether this association resides in heredity, maternal uterine environment, or elsewhere is as yet obscure. Second-generation data on reproduction are now being collected in Amsterdam on the offspring of the women who comprised the births of the cohorts of the original Dutch Famine Study. Preliminary results from the Amsterdam data show similar associations between generations. What is of special interest in this new result is that the correlation of birth weight in mother and child persists at the same level in all cohorts regardless of the sharp changes in relation to the famine experience of the mothers. One might therefore deduce that whatever the determinants of maternal birth weight may be, any environmentally induced change in maternal birth weight, such as the result of starvation during famine, is transmitted to the offspring with the same force as any other determinant, including genetic determinants.

<table>
<thead>
<tr>
<th>Table 1. Results Assembled by Stage of Gestation</th>
</tr>
</thead>
<tbody>
<tr>
<td>Periconception</td>
</tr>
<tr>
<td>Fertility</td>
</tr>
<tr>
<td>Organic brain defects (NTD mainly)</td>
</tr>
<tr>
<td>Schizophrenia</td>
</tr>
<tr>
<td>Schizoid and antisocial personalities</td>
</tr>
</tbody>
</table>

92 Nutrition Reviews, Vol. 52, No. 3
Summary

The array of results discussed in this paper is assembled by stage of gestation in the following table (Table 1). We must allow some uncertainty for some outcomes as to the precise stages of gestation in which they originate. For instance, with obesity, the effects of famine exposure may extend beyond the first trimester into the second. Yet, the table is perhaps sufficient in itself to support the point with which the paper opened, namely, that stage of development is crucial to the appreciation of the effects of prenatal nutrition.

7. Sindram IS. De invloed van ondervoeding op de groei van de vrucht. Nederlandsch Tijdschr Verlosk 1945;45:30–48
34. Jones AP, Friedman MI. Obesity and adipocyte ab-

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.
41. Susser E, et al. a. submitted
42. Susser E, et al. b. submitted