Longitudinal Multicenter Follow-up of High-risk Infants: Why, Who, When, and What to Assess
Betty R. Vohr*, Michael O’Shea†, and Linda L. Wright‡

This article reviews experience with longitudinal follow-up of high-risk infants in the National Institute of Child Health and Human Development (NICHD) Neonatal Research Network. In 1993, the Network initiated a research protocol to provide longitudinal follow-up of infants with birth weight less than 1001 g and infants with higher birth weights who participated in certain Network randomized trials. Infants are assessed at 18 to 22 months corrected age (corrected for degree of prematurity) using measures of the infant’s health, growth, neuromotor, and early cognitive functioning, language, behavior, and family resources. Data from these assessments have been used to investigate potentially modifiable risk factors for cerebral palsy and delayed early cognitive functioning and to evaluate the risks and benefits of interventions assessed in randomized trials. The Network’s experience thus far suggests that longitudinal follow-up can provide valuable information about treatments given to fetuses and neonates.

The importance of longitudinal follow-up of high risk infants is evidenced by the increasing numbers of follow-up studies published in peer reviewed journals and the time dedicated to epidemiology and follow-up at national and international meetings. Historically, follow-up studies were initiated in response to the concern that efforts to salvage increasingly immature neonates were resulting in increased numbers of impaired survivors within a cohort of infants.1

Why Do We Need Longitudinal Follow-up?
As neonatal intensive care became increasingly common and outcomes of survivors improved, subsequent studies evaluated the relationships between maternal or infant characteristics and infant and childhood outcomes.4,6 The associations among respiratory distress syndrome, extremely low birth weight (ELBW), and outcome have been a major focus of follow-up studies. More recently, in recognition of the fundamental importance of neurodevelopmental outcome to families and society (rather than proximate surrogate outcomes), randomized controlled trials have been designed with neurodevelopmental outcome at follow-up as the primary endpoint.7,12 The contribution of follow-up data to clinical trials is particularly relevant because of accumulating evidence that some interventions that are associated with “short-term gain, may have long-term pain.”13 Perhaps the most notable example of the difference between short-term and long-term outcomes is postnatal steroids which may facilitate extubation of a very low birth weight infant, but may be associated with an increased risk of cerebral palsy.14

Acknowledgement of the importance of long-term outcomes has contributed to the development of multicenter research networks including the National Institute of Child Health and Human Development (NICHD) Neonatal Network, the Vermont Oxford Neonatal Network, the Canadian Network and European Networks. Networks provide the large sample sizes required for clinical trials, an infrastructure which includes data management support, funds for tracking and follow-up, and mechanisms to establish standardization of assessments. Another strength of multicenter networks is that they facilitate the collaboration and collegiality of...
highly skilled academicians from different institutions with a wide spectrum of expertise. Finally, the results of Networks may be more generalizable than the results obtained by a single clinical center.

The NICHD Neonatal Network was established in 1986 to conduct randomized controlled trials and studies to improve the management and care of neonates, especially low birth weight neonates; the follow-up component was added in 1993. In 2003, investigators from 16 tertiary care centers, a large Data Coordinating Center, and NICHD collaborate in multicenter clinical research. The infrastructure includes a Steering Committee composed of the principal investigator (PI) from each clinical center, the follow-up subcommittee chair, the PI from the Data Coordinating Center (DCC), and the NICHD project officer.

The Follow-up subcommittee is composed of a follow-up PI from each site. This committee developed the initial framework and design of the follow-up protocol and makes revisions to the protocol as required. Responsibilities of the follow-up subcommittee include: 1) To define follow-up study objectives; 2) To develop and revise the standard follow-up protocol to reflect these study objectives; 3) To identify appropriate assessments; 4) To develop methodology, manual of operations, and forms; and 5) To ensure that compliance rates are ≥80% for all infants ≤1000 g at birth and all infants participating in randomized controlled trials (RCTs). The individual study objectives of the NICHD Research Network follow-up for the 18- and 30-month corrected age visits are shown in Table 1.

**Table 1. Study Objectives of 18- and 30-month Visit**

<table>
<thead>
<tr>
<th>Objective</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. To determine the rate of developmental delay</td>
</tr>
<tr>
<td>2. To determine the rate of motor impairments</td>
</tr>
<tr>
<td>3. To determine the rate of language delay</td>
</tr>
<tr>
<td>4. To determine the rate of behavioral problems</td>
</tr>
<tr>
<td>5. To determine the rate of resource utilization, including special services</td>
</tr>
<tr>
<td>6. To determine the rate of failure to thrive</td>
</tr>
<tr>
<td>7. To determine the contribution of neonatal and environmental factors on outcomes</td>
</tr>
<tr>
<td>8. To determine the relationship between language delay and behavior problems</td>
</tr>
<tr>
<td>9. To determine the predictive value of the 18 month corrected age assessment on outcomes at 30-month corrected age</td>
</tr>
<tr>
<td>10. To provide the outcome assessments for randomized trials</td>
</tr>
</tbody>
</table>

**Who Will Be Followed?**

Although there are numerous categories of risk in both term and preterm infants, recent interest has focused on the vulnerability of “extremely preterm survivors.” Survival rates of extremely low birth weight (ELBW; ≤1000 g) infants have improved dramatically in the past 50 years.

In fact, the survival rates of infants less than 800 g have increased from 0% in 1943 to 1945 to 34% in 1987 to 1988 to 70% in 1994. Though the NICHD registry includes data on all infants with a birth weight less than 1500 g cared for at participating Network sites, the follow-up component was limited to the infants ≤1000 g because they are considered at highest risk for neurologic, developmental, and growth sequelae. In addition to all ELBW infants, all infants enrolled in RCTs since 1995 and infants enrolled in clinical trials whose primary outcome is neurodevelopment are enrolled in the follow-up study.

**When Is the Optimal Age of Assessment?**

Choice of the age of assessment reflects the purpose of the study, the corrected age of preterm infants, the presumed period of risk, the longest feasible duration of follow-up, and the burden of long-term tracking. Although 5-year evaluations provide valuable information about school readiness, assessment at 5 years requires costly long term tracking and interim assessments. In addition, 5 year assessments are strongly impacted by family socio economic status. The goal of the NICHD Network follow-up study is to obtain valid information on development, growth and motor impairments on each child with the support of the modest extramural funds available from NICHD. Site experience and a review of the literature suggested that at 18-22 months corrected age a valid developmental examination can be conducted and that a neurologic examination will identify the vast majority of children with cerebral palsy (CP), certainly all those with moderate or severe CP. In 2002, a pilot protocol was initiated to assess feasibility of tracking and evaluating ELBW infants at 28 to 32
months of age. The cohort of 1,400 infants in
the Glutamine Trial was chosen for this
project. The addition of this cohort provided the
opportunity to evaluate the developmental tra-
jectory of ELBW infants and to evaluate the
impact of nutrition (the glutamine trial col-
lected detailed nutritional intake data) on neu-
rodevelopmental outcome, as well as to deter-
mine the feasibility of conducting a routine 30
month assessment at participating Network sites.

What Are the Assessments?
The following categories of assessment were
identified for the comprehensive follow-up pro-
tocol: socioeconomic status, family resources,
growth, neurologic status, gross motor function,
developmental status, behavior, and language.
An attempt was made to identify assessment
tools that were user friendly and child friendly,
as well as reliable, standardized, valid assess-
ments of outcome.

The tests that have been administered at sites
since the onset of the follow-up study are shown
in Table 2. The assessments include the Hollings-
head Social and Economic (SES)24 status at dis-
charge, 18 and 30 months, a medical history
which includes the use of oxygen, monitors,
medications, and nutrition supplements, a hos-
pital readmission form, growth parameters, a
neurological examination,25 a gross motor func-
tional assessment,26 the Bayley Scales of Infant
Development,27 language28 and behavior assess-
ments,29 the Functional Status II question-
aire,30 the Family Resource Scale,31 and the
Impact on the Family Scale.32 All definitions,
protocols, and procedures are contained in a
manual of operations, which includes back-
ground information, instructions for protocol
implementation, and study forms. Details of the
procedures have been published.33 The majority
of the assessments in the follow-up battery have
remained unchanged since the study’s inception
in 1994. Revisions that were made to the proto-
col in 2002 include dropping the Impact on the
Family (used between 1994 and 2001) and the
Functional Status Assessment (used between Jan
1994 and Jan 1, 2002), and the addition of a
behavior assessment, The Brief Infant-Toddler
Social Emotional Assessment (The BITSEA),
and measures of receptive and expressive lan-
guage, based on Bayley subscores, in 2002. The
manual of operations contains protocols, defini-
tions, and instructions for administering assess-
ments and study forms and examiners are
trained to reliability on all assessments.

Individual components of the assessments are
shown in Table 2. Structured interviews are per-
formed to obtain social and economic status
information (SES)25 including maternal and pa-
ternal education and occupation, marital status,
and a detailed interim medical history with data
on hearing, vision, and rehospitalizations, and
resource utilization including early intervention.

Table 2. NICHD Research Network Follow-up Study Assessments

<table>
<thead>
<tr>
<th>Assessment</th>
<th>18-22 Months</th>
<th>28-32 Months</th>
</tr>
</thead>
<tbody>
<tr>
<td>Socioeconomic Status</td>
<td>Hollingshead24</td>
<td>Hollingshead24</td>
</tr>
<tr>
<td>Postdischarge Morbidity</td>
<td>Medical History</td>
<td>Medical History</td>
</tr>
<tr>
<td></td>
<td>Resource Utilization</td>
<td>Resource Utilization</td>
</tr>
<tr>
<td></td>
<td>Hospital Readmission</td>
<td>Hospital Readmission</td>
</tr>
<tr>
<td>Growth</td>
<td>Weight, Length, Head</td>
<td>Weight, Length, Head</td>
</tr>
<tr>
<td></td>
<td>Circumference- Standard</td>
<td>Circumference- Standard</td>
</tr>
<tr>
<td>Neurologic exam</td>
<td>Amiel-Tison25</td>
<td>Amiel-Tison25</td>
</tr>
<tr>
<td>Gross motor function</td>
<td>Palisano GMC26</td>
<td>Palisano GMC26</td>
</tr>
<tr>
<td>Development</td>
<td>Bayley Scales of Development II</td>
<td>Bayley Scale of Development II</td>
</tr>
<tr>
<td></td>
<td>Mental and Motor Scales27</td>
<td>Mental and Motor Scales27</td>
</tr>
<tr>
<td>Language</td>
<td>Bayley subscores27</td>
<td>Bayley subscores27</td>
</tr>
<tr>
<td></td>
<td>Peabody receptive scores28</td>
<td>Peabody receptive scores28</td>
</tr>
<tr>
<td>Behavior</td>
<td>Bayley Behavior Record27</td>
<td>Bayley Behavior Record27</td>
</tr>
<tr>
<td></td>
<td>Brief Infant Toddler</td>
<td>Brief Infant Toddler</td>
</tr>
<tr>
<td></td>
<td>Social-Emotional Assessment29</td>
<td>Social-Emotional Assessment29</td>
</tr>
<tr>
<td>Functional status</td>
<td>Functional Status Measure II30</td>
<td></td>
</tr>
<tr>
<td>Stress</td>
<td>Impact on the Family32</td>
<td></td>
</tr>
<tr>
<td>Family resources</td>
<td>Family Resource Scale31</td>
<td>Family Resource Scale31</td>
</tr>
</tbody>
</table>
Weight, length, and head circumference are measured using standard techniques and plotted at the corrected age adjusted for gender on National Center for Health Statistics growth curves. The child is weighed naked and weight is recorded to the nearest hundredth of a kilogram. Recumbent length is obtained using a Pediatric length board. The mean of three measurements is recorded to the nearest quarter of a centimeter. Occipital-frontal circumference is recorded as the largest of 3 measurements in centimeters to the nearest quarter of a centimeter.

The neurological examination is based on the Amiel Tison neurologic assessment. Neurologic examinations are performed by experienced certified examiners who have been trained to reliability in the exam procedure in an annual two-day practicum on neurologic assessment. New neurologic examiners at sites are an annual two-day practicum on neurologic and are neither comprehensive nor age-normative. Distinctions between levels of gross motor function are based on functional limitations, the need for assistive technology including mobility devices (such as walkers and wheeled mobility), and to a much lesser extent the quality of movements.

Mental and motor development are assessed using the Bayley Scales of Infant Development (ed 2) (BSID-II), the most widely used infant developmental assessment measure. The BSID-II correlates well with other infant assessment instruments and discriminates normal from delayed performance. The Mental and Motor Scales measure the current levels of functioning in the cognitive, language, personal-social, and fine and gross motor areas. The Mental Development Index (MDI) and Psychomotor Development Index (PDI) (mean = 100, SD [standard deviation] = 15) are used as general indicators of developmental status and are normed to 42 months. Reliability coefficients for the Mental and Motor Scales are 0.88 and 0.84, respectively. The BSID-II has been empirically shown to correlate with other infant assessment instruments and to adequately discriminate normal from delayed performance. Examiners are recertified annually.

In 2002, 2 approaches to language assessment were initiated. The first approach involves a more thorough analysis of individual Bayley items. Subscores have been published on the old Bayley but not on the new Bayley II. Currently, the Network sites are collecting individual item-by-item responses on the mental and motor scales. The data will be used to develop new language subscores on the Bayley II. The other approach is based on the Peabody Picture Vocabulary Test (ed 3), a measure of the child’s receptive (hearing) vocabulary. It consists of 175 items and takes approximately 12 minutes to administer. This is a nonverbal, multiple-choice test, which measures receptive knowledge of vocabulary words of individual ages 2 1/2 through...
adulthood. The child is required to identify a stimulus word from a selection of 4 pictures. Raw scores are converted to standard scores with a mean of 100 and a SD of 15. This test has been well normed. Psychometric properties are acceptable with reported reliability and validity estimates ranging from 0.50 to 0.91.

Currently, there are 2 behavior assessments used. One involves direct observation during the Bayley administration and the second is a parent interview. The Bayley Behavior Rating Scale (BRS)\textsuperscript{27} assesses the child’s behavior during a testing situation. The BRS has a 5-point scoring system with descriptors at each point; items are grouped according to appropriateness by age levels. Percentile rankings are obtained for the BRS Total Score and four Factor Scores, motor quality, attention/arousal, orientation/enjoyment, and emotional regulation.

The Brief Infant-Toddler Social Emotional Assessment (BITSEA) Version 1.0\textsuperscript{29} is designed as a first step in the early identification of 1 to 3 year olds with social-emotional problems and/or delays in social competence. It consists of 44 interview questions. A trained medical or non-medical professional administers this scale to the primary caretaker of the child. Estimated time to administer the BITSEA is 15 minutes.

Parent perception of the child’s functional status was assessed between 1994 and 2002 using the Functional Status II Measure developed by Stein, Reissman, and Jessop.\textsuperscript{30} This measure consists of 42 interview questions administered to the primary caretaker. The interview is designed for parents of children from 2 weeks to 16 years of age. Thirty-one questions are administered at the 18 month level. The questionnaire assesses the caretaker’s perception of the child’s overall functional status and responsibility.

The objective of the Impact on the Family\textsuperscript{32} questionnaire is to measure the effect of the child’s condition in producing change in the family. Impact was conceptualized as the effects of a child’s illness on the family system, in 4 dimensions: economic, social, familial, and strain. The developers of this assessment considered qualitative data from patient interviews, reviews of the literature and clinical experience of providers who treat chronically ill children. Psychometric data were used to further refine the scale and a factor analysis of the first 100 cases revealed 4 dimensions of impact in the measure: financial, social/familial, personal strain, and mastery. Three of the 4 hypothesized dimensions were empirically demonstrated by the factor analysis: 2 theorized dimensions combined into 1 (familial/social), and a new one was generated (coping or mastery). The twenty-four item scale elicits variability in response and is internally consistent. This measure was administered in the Network from 1994 until 2001. In 2002 at the annual follow-up meeting, a decision was made to introduce new assessments in two important areas, language and behavior. Because of time constraints when evaluating a young child, other components of the assessment (Impact on the Family and Functional Status) were dropped from the protocol.

The Family Resource Scale (FRS)\textsuperscript{31} measures the adequacy of different resources in households with young children. The scale includes 31 items rated on a five-point scale ranging from not at all adequate (1) to almost always adequate (5). The hierarchy is derived from a conceptual framework that predicts that inadequacy of resources necessary to meet individually identified needs will negatively affect both personal well-being and parental commitment to carrying out professionally prescribed regimes unrelated to identified needs. The reliability and validity of the scale were established in a study of 45 mothers of preschool retarded, handicapped, and developmentally at-risk children participating in an early intervention program. Coefficient alpha computed from the reliability was 0.95 corrected for length by using the Spearman-Brown formula. The short-term stability of the FRS was determined for all 45 patients administered the scale on 2 occasions 2 to 3 months apart. The stability coefficient for the total scale scores was \( r = .52 \ (P < .001) \).

**What Is the Value of Follow-up in Neonatal Network Studies?**

Collecting follow-up data is costly, so it is reasonable to consider the extent to which these data inform us about the risks and benefits of perinatal interventions, over and above the information that can be obtained during the acute neonatal hospitalization. From the perspective of individual patients, the Network follow-up protocol provides developmental screening so that developmental delays can be identified and re-
recommendations can be made for early intervention. From
a larger perspective, the Network’s experiences illustrate the concept that short-term outcomes may be misleading with regard to an intervention’s eventual effects on quality of life. In clinical trials, the addition of follow-up outcomes increases the likelihood that unanticipated risks or benefits will be detected. For example, the studies with postnatal steroids demonstrated that although short-term benefits are derived, the long-term outcomes may be an increased rate of neurodevelopmental handicap. The Network’s experience suggests that follow-up should be a standard aspect of all trials involving fetuses and newborns.

The first randomized trial reported from the Network was a trial of biweekly intravenous gamma-globulin (IVIG) to prevent nosocomial infections in very low birth weight (VLBW) infants. In contrast to the meta-analysis of trials that existed at that time, the Network trial did not show that IVIG was associated with a reduced risk of infection and this intervention has not been adopted by neonatologists. However, IVIG was also associated with a small increased risk of necrotizing enterocolitis (NEC), and NEC is associated with an increased risk of adverse outcome. Therefore, had the Network trial results agreed with earlier trials (in showing a decrease in the risk of sepsis) then long term follow-up would have been essential for deciding whether the risk (increased frequency of NEC) was outweighed by the benefit (decreased frequency of sepsis).

The second randomized trial reported by the Network was a comparison of 2 surfactant preparations. This trial indicated that infants treated with natural surfactant had lower oxygen and ventilator requirements in the 72 hours after randomization, but no difference was found in the primary study end point, which was death or bronchopulmonary dysplasia. Although long-term follow-up might have provided useful information about a possible difference in the long term effects on the lung of neonatal exposure to bovine versus synthetic surfactant, such data were not collected.

The first Network trials to show a benefit of the study intervention was a randomized trial of vitamin A to prevent neonatal chronic lung disease from 62% to 55%. Despite this benefit, vitamin A is not uniformly used by neonatologists (Tyson J, personal communication, Jan 2003). It is reasonable, however, to speculate that a decrease in the prevalence of chronic lung disease would result in improved neurodevelopmental outcome, and that had the Network trial demonstrated benefit in this long term outcome, its impact on clinical practice might have been greater.

The Network has performed several trials investigating the safety and efficacy of inhaled nitric oxide (INO) in the treatment of infants with severe respiratory failure. In the NINOS trial, INO-treated infants, as compared to control infants, had similar mortality risk but a lower risk of the need for extracorporeal membrane oxygenation (ECMO). Although there was no difference in neurodevelopmental outcome, the follow-up assessments demonstrated that INO was not associated with any specific increase in untoward long-term effects. One of the objectives of the Network’s second trial of nitric oxide, the early INO trial, was to determine if providing nitric oxide to infants with less severe respiratory failure would further reduce the need for ECMO and lead to improved neurodevelopmental outcome. However, that trial used an early versus late treatment design, almost all infants received INO, and the incidence of death or need for ECMO was about 20% in each group. Therefore, identifying a treatment effect of INO on long-term outcome may be difficult, especially given the prerrandomization difference in severity of illness between the NINOS and the Early INO trials. In the ongoing Network trial of nitric oxide in preterm infants, neurodevelopmental outcomes are planned for 18 to 22 months CA.

The first Network trial to include long term follow-up involved antenatal treatment with phenobarbital to prevent intracranial hemorrhage in preterm infants. Although the primary outcome was intraventricular hemorrhage, neurodevelopmental follow-up was performed on study participants, presumably in an effort to assure long-term safety. No treatment effect was found on either the primary outcome or the neurodevelopmental status at follow-up.

In the first of 2 trials involving dexamethasone, no difference was found in the prevalence of death or chronic lung disease for infants...
treated with a 14-day tapering course of dexamethasone, beginning at 2 weeks of age versus those given the same intervention beginning at 4 weeks of age.\textsuperscript{44} Because of a study design that compared early and late steroid treatment, it is possible that follow-up of study participants would not have identified a treatment effect. However, since it is also plausible that the adverse effects of dexamethasone on the brain depend, in part, on the timing of exposure (ie, the postmenstrual age at which the infant is exposed), such follow-up data would make an important contribution to the debate about the safety of postnatal steroids for treatment of chronic lung disease.

A second trial\textsuperscript{45} involving postnatal steroids and minimal ventilation, using a factorial design, was stopped early because of an increased risk of spontaneous gastrointestinal perforation in steroid-treated infants. The primary end point in this trial was the combined outcome of death or neonatal chronic lung disease. Implicit in the decision to stop this trial early was the assumption that the expected benefit of the interventions could not be large enough to offset the harm associated with isolated gastrointestinal perforation. On the other hand, the long-term effects of chronic lung disease are thought to be substantial, whereas those of isolated perforation are not well characterized. Thus, it is plausible that even if dexamethasone increased the risk of isolated gastrointestinal perforation, it might, by decreasing the frequency of chronic lung disease, improve outcome assessed at follow-up. Follow-up data from this trial suggests that neurodevelopmental outcome was not affected by minimal ventilation,\textsuperscript{45} but thus far analyses of the effects of postnatal steroids have not been reported.

In three recent trials, the primary focus has been to decrease the risk of death or adverse neurodevelopmental outcome at 18 months. One is the completed trial of indomethacin for prophylaxis of intracranial hemorrhage, in which no treatment effect on neurodevelopmental outcome was observed.\textsuperscript{46} The results of this trial have caused a re-examination of the use of prophylactic indomethacin, perhaps, in part, because the primary outcome is a compelling one.\textsuperscript{47} In ongoing trials of hypothermia for term infants with neonatal encephalopathy,\textsuperscript{48} and aggressive versus conservative phototherapy for extremely low birth weight infants (Tyson JE, personal communication, Jan 2003), death or adverse neurodevelopmental outcome is the primary end point.

In addition to the follow-up of babies participating in clinical trials, observational studies of neurodevelopmental outcomes have been completed. The first of these was reported in 2000, and suggested that a disturbingly high proportion of infants (49%) have one or more major neurodevelopmental or sensory impairment at 18 months CA (ie, cerebral palsy, significantly delayed mental development, blindness, or deafness).\textsuperscript{35} Among these, sensory impairments are considerably less prevalent. Further, the functional impact of deafness can be attenuated considerably by early intervention. Delayed mental development at 18 months CA is strongly associated with mental retardation, which has broad implications for a variety of outcomes such as school performance and social functioning. On the other hand, it is well known that a majority of cases of mild cerebral palsy diagnosed at one year of age resolve during early childhood,\textsuperscript{49} so the functional significance of cerebral palsy diagnosed at 18 months CA is not entirely clear. Other reasons for imperfect concordance between assessments at 18 months CA and eventual functional outcome include the relatively modest correlation between mental development assessed at 18 months CA and intelligence among children without delayed mental development, and the ongoing influences of environmental factors, such as family resources and early childhood intervention services. In contrast to the low predictive ability of assessments at 18 months CA for predicting later functioning among children who lack a major neurodevelopmental impairment, these early assessments do have a high positive and negative predictive value with respect to the presence, at school age, of major handicap.\textsuperscript{50}

The Network has used follow-up data for analysis of risk factors for adverse neurodevelopmental outcomes. These analyses indicate that short-term outcomes including intraventricular hemorrhage (IVH), bronchopulmonary dysplasia (BPD), NEC and late onset sepsis are associated with an increased risk of cerebral palsy or delayed early cognitive functioning (low scores on the Bayley Mental Developmental Index) outcome at 18-22 months CA. In addition to
analyses of risk factors, observational data on extremely low birth weight and extremely premature infants make possible the monitoring of trends in mortality and morbidity among infants at the limits of viability (ie, 23-24 weeks of gestation).

The assessments of babies at 18 months CA provide an infrastructure that can be used to provide assessments at older ages. More information about the eventual outcome of children without major handicap can be obtained from assessments prior to school entry. To this end, the Network has begun a pilot project in which a subset of Network children will be assessed at 28-32 months of age, when intelligence, language development, and early signs of attentional deficits can be evaluated. It is reasonable to anticipate that ongoing studies by the Network will indicate that outcomes at 28-32 months CA will be even more informative than outcomes at 18-22 months CA about eventual quality of life. In addition, the collection of follow-up data at 28-32 months will provide information about the usefulness of evaluations at 18 months for predicting persistent motor impairments, cognitive impairments, and, perhaps, behavioral challenges.

The efficacy of long-term follow-up is clearly shown by the findings of Ment et al51 in a multicenter study that serially evaluated the cognitive and receptive vocabulary outcomes of VLBW infants 600 to 1250 g birth weight at 3, 4 ½, 6, and 8 years of age. Peabody vocabulary scores and IQ scores improved with increasing age and only children with early grade I-II IVH and subsequent significant central nervous system injury (grade 3-4 IVH, PVL, or ventriculomegaly) had test scores that declined over time. Environmental factors associated with higher test scores over time included higher maternal education, residence in a two-parent household, and early intervention for children whose mothers had less than a high school education. Although the study provides longitudinal data on a large cohort of VLBW survivors, there are limiting factors in the study including the fact that there were no full-term controls and the investigators did not report data on visual motor skills or executive function. This longitudinal follow-up study, however, makes several important contributions to the literature: 1) It emphasizes the importance of evaluating the interaction of biologic and environmental factors on outcome over time; 2) It supports the concept of “plasticity and recovery” over time for cognitive and receptive vocabulary skills for VLBW infants without significant central nervous system injury; and 3) It alerts us to the importance of studies to decrease the incidence of severe central nervous system injury among VLBW infants. Multicenter longitudinal studies with serial assessments to adolescence and comparisons to full term controls are needed to fully assess and understand the impact of VLBW and environment on outcome.

A number of lessons can be taken from the Network’s experience with follow-up thus far. Perhaps foremost is the recognition that follow-up data are often needed when assessing the risks and benefits of interventions directed at fetuses and neonates. Short-term outcomes do not appear to be as compelling as long-term ones. For at least a few interventions, the direct treatment effects differ for short- versus long-term outcomes. Further, as the experiences with postnatal steroids suggest, the safety of some interventions cannot be assessed without follow-up data. It appears then, that at least a portion of perinatal research funding and effort should be directed at follow-up. Work is needed, however, toward the goal of increasing the predictive value, without increasing cost, of follow-up data collected in early childhood.

References


24. Hollingshead AB: Four Factor Index of Social Status, New Haven CT, Yale University Press, 1975


