THE DESCRIPTIVE EPIDEMIOLOGY OF COGNITIVE AND ACADEMIC OUTCOMES AT AGE 9 OF CHILDREN BORN WEIGHING LESS THAN 2000 GRAMS

By

Daniel E. Harrison

A THESIS

Submitted to
Michigan State University
in partial fulfillment of the requirements
for the degree of

MASTER OF SCIENCE

Department of Epidemiology

1999

ABSTRACT

THE DESCRIPTIVE EPIDEMIOLOGY OF COGNITIVE AND ACADEMIC OUTCOMES AT AGE 9 OF CHILDREN BORN WEIGHING LESS THAN 2000 GRAMS

By

Daniel E. Harrison

This study presents the descriptive epidemiology of four cognitive and academic outcomes in a geographically defined group of 471 low birthweight (LBW) children from New Jersey followed prospectively to age nine. The outcomes assessed were learning disability (LD), and low scores (< 85) on IQ and Woodcock-Johnson (WJ) measures of reading and math achievement. The group was found to face more than twice the risk of LD usual for children in New Jersey. Birthweight had a significant gradient effect on all the outcomes even after adjusting for gender, race, and home environment. The HOME score measure of home environment was found to greatly attenuate the effect of race and to render its predictive value non-significant for all outcomes.

ACKNOWLEDGMENTS

The author would like to thank all the members of his thesis committee, Dr. Nigel Paneth, Dr. Alka Indurkhya, Dr. Jennifer Pinto-Martin, Dr. Agnes Whitaker, and Dr. Michael Lambert for all their kind help and insightful guidance. Thanks are also due to James Jetton for his help with data management and to the department secretaries, particularly Lora McAdams and Edyth Holland without whom I could never have navigated the graduate paperchase. I must also thank my friend Becky Oakes for helping me endure the vicissitudes of graduate life. A smaller, but not insignificant debt of gratitude is due to the kind barristas at Espresso Royale Caffe and Beaner's Cafe who could be counted on to ask how the thesis was going and occasionally slip me an extra shot of espresso. There is, of course, a redundant thanks to be offered to my mother and father for obvious reasons. Finally, the author could not have made it through the program without the social and occasionally financial support of his mother and stepfather. Thanks for keeping the homefires burning. To these people and the others who are too numerous to mention, thank you all.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

LIST OF ABBREVIATIONS

BWT - birthweight

CI - confidence interval

CP - cerebral palsy

DQ - developmental quotient

DCP - disabling cerebral palsy

ELBW - extremely low birthweight (BWT < 1000g)

GA - (gestational age)

HOME - Home Observation for Measurement of the Environment

LBW - low birthweight (BWT < 2000g)

LD - learning disability

LP - learning problem

MR - mental retardation / mentally retarded

MP - math problem

NBHS - Neonatal Brain Hemorrhage Study

NBW - normal birthweight

NICU - neonatal intensive care unit

OR - odds ratio

PI - Perceptually Impaired. New Jersey's Department of Education's equivalent to learning disability.

RP - reading problem

RR - relative risk or risk ratio

SD - standard deviation

SES - socioeconomic status

VLBW - very low birthweight (BWT < 1500g)

WJ - Woodcock-Johnson Psychoeducational Battery

INTRODUCTION

Advances in perinatal care and neonatal intensive care have increased the survival rate of low birthweight (LBW, <2500g) babies in ways that would seem miraculous to doctors trained as recently as the 1970's[1]. Survival for extremely low birthweight (ELBW, <1000) babies in first world countries, for instance, has increased from almost nil to around fifty percent since the early 1970's[2]. This kind of increased survival is now drawing attention to questions concerning the new survivors. It is well known that low birthweight babies are at increased risk for cerebral palsy (CP) and various other neurological abnormalities such as blindness or hearing loss, and we are beginning to appreciate that LBW may adversely affect intelligence, cognitive functions, and academic success in later years. Thus, as improved care for LBW newborns aids doctors in their primary task of saving these babies, it also raises concerns about the costs and benefits of doing so. These new survivors, after all, may face increased risk for difficult and expensive developmental problems. This concern raises questions about the effects of different perinatal, pediatric, and psychological interventions and their roles in mitigating or exacerbating the consequences of LBW. It is important to assess the effects of LBW and associated treatments as completely as possible so as to apply our resources to reap the greatest benefits in survival at the least cost in suffering, impairment, and expense.

One of the first steps in this cost-benefit analysis is careful descriptive epidemiology. The particular concern of this thesis will be school-age deficits in cognition and academic aptitude in a geographically defined LBW population followed from birth to age 9. In concert with previously collected data on perinatal care and health outcomes at preschool and early school age, this descriptive epidemiology will comprise a necessary step in the process of understanding the public health impact of LBW, and eventually, of comparing different treatment strategies in a cost-benefit analysis.

REVIEW OF THE FIELD

School age outcomes are important in their own right. In addition to quantifying the more subtle effects of LBW, the study of school-age outcomes should help to identify causes, predictors, and effective preventions for cognitive and academic deficits. Because LBW studies are best done longitudinally[3] and follow a high-risk group, they present a rich opportunity for identifying early markers of, for instance, learning disability (LD). Lyon[4] argues that the earlier we identify a child's LD, the better the effect of the intervention.

Early Deficits:

Most research on the LBW population has focused on early outcomes[2]. This research has made it clear that babies born LBW face abnormal risk of CP, and vision and hearing loss [1, 2, 5, 6, 7, 8]. Ornstein et al.[2], after reviewing 25 studies of ELBW and/or very low birthweight (VLBW, <1500g) survivors reported that the prevalence of CP in this birthweight group varied between 2.4% and 9%. Visual impairment affected between 2% and 38% of the children in the populations reviewed, and the prevalence of hearing loss ranged between 2% and 44%. The wide range in the proportions of these disorders is probably attributable to the use of different definitions of the disorders, different birthweight groups, differences in perinatal care, and to imprecision arising from small sample sizes. Escobar et al.[9], in a meta-analysis, estimated the rate of CP in the VLBW population at 7.7%. They estimated the incidence of overall disability at 25%. The range of these estimates greatly exceeds the 2 to 3 per thousand prevalence estimate for CP in the general school age population.

Outcome Measurement:

Recent research has focused on the more subtle deficits that frequently do not present until school age. The literature on school-age outcomes focuses, broadly speaking, on three areas. These are psychological/behavioral problems, cognitive deficits, and academic or performance outcomes. This paper will deal solely with cognitive and academic outcomes. Before reviewing the literature in this area though, a few words should be said about the particular difficulties and pitfalls of studying these outcomes. The difficulties arise from uncertain definitions, confounding variables, and biases.

Three academic and cognitive outcomes will be used here. These are IQ, learning problems, and learning disability. IQ is ostensibly a measure of intelligence or mental potential. While it is a hotly debated construct[10], it is attractive to researchers because a few common standardized tests of IQ have been normed on large populations and are often used in the literature. Whereas it may not please all theorists, a result can at least be compared to extant literature. A learning problem (LP), simply denotes poor performance in school (or sometimes on an achievement test). It is not only theoretically straightforward, but easy to measure. To facilitate comparisons, this paper will use a score below 1 SD below the mean on Woodcock-Johnson achievement tests to indicate a LP. However, to evaluate validity, this will be compared to a teacher's designation of the child as at or below grade level using a subset of the sample. Finally, LD represents a more subtle construct and has more problems associated with its operationalization. The concept refers to specific, presumably neurological, deficits which may impede the achievement of students, even those who are of average or above average intelligence, and who are otherwise unaffected by medical complications such as sensory deficits or CP. Many of the difficulties associated with operationalizing LD are peculiar to the field and will be discussed later.

All three outcomes are affected by environmental stimuli. Socioeconomic status (SES), home environment, infant-caregiver interaction, race, gender, and a litany of other measures each represents some aspect of a child's intersection with his or her environment. A variable may represent a factor that has direct effect on the outcome of interest (proximal), or it may be a proxy for other unidentified variables (distal). For instance, while Blacks in the US generally do not perform as well on IQ tests as Whites, it appears that this difference is a function of test-bias and differential access to education due to financial and social disadvantages. Indeed Van Rossem et al.[11] showed that the effects of race on IQ (in the LBW population at least), can be reduced to non-significance by controlling for the amount of stimulation available to the child in the home environment. In this example race is a more distal variable grossly subsuming the more proximal effects of home environment. While we would prefer to understand the proximal causes of an outcome, we must often use more distal constructs to minimize the effects of our ignorance.

The effect of a child's environment certainly cannot be left out of the equation relating birthweight to IQ or learning, for it is both strongly predictive and possibly confounding[2, 7, 12, 13, 14, 15, 16, 17, 18]. Breslau et al.[13] found that IQ was as related to urban vs. suburban residence as to birthweight. Cohen et al.[19] reported that infant-caregiver interaction is as predictive of poor developmental outcome as are perinatal and pediatric history. Lloyd et al. found a similar effect on IQ when using father's occupation as a marker for social class [15]. In fact, in the review by Ornstein et al., SES, however measured, was often the most important variable for predicting a child's IQ and academic performance[2].

One should be careful in interpreting these results, however, as they may be confounded. The possibility of confounding comes from the likelihood that children who are born into low SES families are at once more likely to be born LBW and to have poorer cognitive and academic outcomes. In some cases it is hard to say how much effect comes from being born LBW and how much from the various factors subsumed in low SES.

In a multisite randomized trial executed by the Infant Health and Development Program, the investigators intervened to improve the learning environments of the experimental group[20]. They found clinically and statistically significant effects of intervention on IQ, which paralleled previous findings in disadvantaged youth. They also found that there was an interaction effect with birthweight, with heavier born children responding better to the intervention. This study suggests that the stimulation provided by a child's environment is one of the important factors subsumed under SES, and that models of the relationship between BWT and IQ may be incomplete if they are missing some measure of the child's access to intellectual stimuli. It is a suggestion further strengthened by Van Rossem et al.'s work showing that a measure of home environment accounts for the usually significant predictive value of race on IQ[11].

Another factor that may affect a child's interaction with his or her environment and thus influence diagnoses of at least one of our outcomes, is gender. Girls are at about a quarter of the risk of being diagnosed with LD [4, 18] as boys, and one half the risk of LP [21]. Why this should be is not entirely apparent. Researchers in the field believe that to some extent girls are being underdiagnosed. Nevertheless, this is not to say that all the difference proceeds from a diagnosis bias. Ornstein et al. found that in a number of studies girls did better in school than did their male counterparts[2]. However, they also found that only 28% of the literature they reviewed dealt with gender at all. This is unfortunate, as in the field of LD, gender may have complex and profound effects, both real and apparent.

The main difficulty that impedes research in the field of LD, however, is the lack of a standard operational definition. As in many areas, the presence of a rough ideological consensus where a construct is concerned does not imply similar convergence regarding a working definition. The theoretical definition in the US and in the field of LD in English-language journals, at least, is defined largely thanks to US Public Law 94-142 (1977) which calls for additional instruction for children with LD. This document defined LD as the existence of a severe discrepancy between a child's ability and achievement in one of seven areas. The areas are 1) oral expression, 2) listening comprehension, 3) written expression, 4) basic reading, 5) reading comprehension, 6) mathematic calculation, and 7) mathematic reasoning. It further stipulated that the disparity should not be due to 1) visual, hearing, or motor handicap (MI for major impairment), 2) mental retardation (MR), 3) emotional disturbance, or 4) environmental, cultural, or economic disadvantage.

This definition is clear in its identification of exclusionary criteria and domains wherein disability may occur, but it is vague in other ways. Much leeway remains, for instance, in choosing measures of ability and achievement, in defining "severe discrepancy," and in statistically limiting the problems inherent to comparing scores from two normally distributed, correlated, and imperfectly precise measures such as achievement and ability. Also, some researchers ask whether there might not be a better indicator of the problem than a score derived from a literal translation of the definition[22, 23, 24, 25, 26]. A construct measuring achievement subtracted from a construct measuring aptitude, might not be as good an indicator of LD as a test of specific deficits in areas necessary for academic functioning. For instance, dyslexia is thought to stem, in some cases, from inappropriate viewing of the page of text. The fact that the standard definition does not focus on specific processing deficits indicates that it is not etiologic. As we do not have a good understanding of what constitutes the process of learning, we cannot test how well a child performs the requisite steps. Also, a definition may be confused by political and financial concerns. Algozzine & Ysseldyke [27] go so far as to say that, because defining LD affects who is given certain educational services, it should not be treated as a technical matter but as a social-policy consideration.

Myklebust has said of LD definitions, "tell me how many students you want to find and I'll write you a definition that will find that many" [quoted in 28]. This is obviously discomfiting to researchers, who as a matter of course, prefer definitions that are valid, stable, etiologic, and at least standard enough to facilitate comparisons to other research. While few fields work with ideal definitions, fewer still are burdened by such disarray as that which is seen in the study of LD.

A brief perusal of the debate in this field is all it takes to see that proposing a satisfactory definition of LD is well beyond the scope of this paper [4, 12, 23, 24, 25, 26, 27 28, 29, 30, 31, 32, 33, 34, 35]. Nevertheless, proceeding with no definition whatsoever is equally untenable. In deference to both the necessities of research and the legitimately thorny issues of the LD field, this paper will opt for a definition that is pragmatic, if not ideal. The definition in this paper combines ease of comparison with uncomplicated practicality. A number of studies have used school placement as a definition of LD. New Jersey, where our subjects attend school, may classify a child as "perceptually impaired" (PI) if the student is properly referred and shows symptoms of LD to a panel of professionals employed by the school district. Though the state of NJ does not yet have a standardized definition of LD by which to classify children, the designation of perceptually impaired conforms, in spirit, to the definition of LD outlined by the National Joint Committees on Learning Disability. Using this definition, one identifies at least those children whose disability is placing a direct financial burden on the state. However, one runs the risk of missing children who are not referred[4] (largely girls), and we lack information necessary to understand the biases or imprecision at work in the diagnoses (no study of reliability, for instance, has been done in NJ to date; nor are the criteria for "perceptually impaired" standardized). It is a definition that reflects the impact of LBW and allows for comparison with many studies, but it is of questionable theoretical validity and may or may not be reliable. Because these problems plague all definitions in the field (indeed there is no strong definition), and because the objective financial cost of LD is reflected in the school district's definition, this operationalization of LD will be used in this thesis. It should be recognized that the moniker "perceptually impaired" is misleading and does not refer to a sensory deficit, but to the same discrepancy between ability and achievement embedded in the majority of LD definitions.

Another popular definition in the research is a regression-discrepancy model. This definition subtracts an achievement score such as the Woodcock-Johnson (WJ) test of achievement from a measure of ability, such as the Wechsler Intelligence Scale for Children - Revised and controls for the bias introduced by regression to the mean[33]. While it may be the best of the discrepancy models, such models have faced a lot of well-reasoned criticism in recent years [22, 23, 24, 25, 27, 28]. Their validity has been questioned because they are not etiologic and because certain of their theoretical underpinnings remain unproved. Moreover, even though the problems of regression to the mean may be attenuated, hosts of other statistical and psychometric difficulties attend the use of these discrepancy definitions. In order to use a regression discrepancy formula it is necessary to know the means and standard deviations of the achievement and ability scores in the normal population, as well as the correlation between the two instruments. Wolke et al.[36] have shown that these parameters change depending upon the composition of your population and have demonstrated the dangers of using outdated or otherwise inappropriate norms to determine the fraction of a population that is of abnormal IQ. Despite having the IQ and achievement scores needed in a discrepancy model of LD, this paper will not operationalize LD in this way. While the use of the discrepancy model might seem to facilitate comparison to a large portion of previous research, the shaky underpinnings of the model as well as the difficulties of obtaining appropriate norms lead this author to believe that the benefits would be undermined by the uncertain validity and reliability of the findings. It is mentioned here only to explain why such a common definition will not be used.

REVIEW OF THE RELEVANT LITERATURE

Having settled upon which definitions to use, it remains to summarize the literature. Previous studies have focused on the Very and Extremely Low Birthweight ranges (VLBW <1500g and ELBW <1000g), and less information is available on children born at heavier weights. The proportion of LD varies widely based on birthweight (BWT) and on the definition used. Saigal et al.[37], who used a discrepancy definition and studied a large regionally defined cohort of ELBW children, found that as many as 49% of the ELBW children were LD. However, this was only about 1.5 times the proportion found in normal birthweight (NBW) children (33.4%). In the same population, 37% of the ELBW group was utilizing special education vs. 16% of the controls (RR=2.3). Vohr et al.[38] found 54% of ELBW children using special education. Ross et al.[16] found that 48% of VLBW children were identified as using special education, but only 19%, or a little over a third of the special education recipients, were diagnosed by the school as LD. This caution about using special education as a marker for LD in the LBW population is echoed in a study by Lefebvre et al.[6], who found that in the ELBW group, 24% received special education, but only 8%, or one third were LD.

Special education enrollment is probably better used to quantify a portion of the burden of LP in a community. In the ELBW studies reviewed for LP[6, 15, 16, 17, 37, 38, 39] the percentage of children receiving special education ranged from 24% to 54%. Hack[8] and Saigal[39], who used $\leq$-2 SD on a standard achievement test, found numbers between 20% and 28% depending upon which subtests and combinations thereof were used. Teacher reports of children performing below grade level identified 36.6% of the ELBW population in Saigal et al.'s study and 53% of the VLBW population in a study by Lloyd et al.[15]. The relative risks in comparison to control children were 2.3 and 2.4 respectively. The numbers from these studies seem to imply that fewer children are in special education than are doing poorly by teachers' estimations, and fewer still are identified by simple diagnostic statistics.

As for the association between IQ and birthweight, the available data suggest that LBW raises a child's likelihood of having an IQ below 85 or even below 70. Numerous studies have found that IQ is significantly associated with birthweight[2, 7, 8, 13, 15, 16, 17, 37, 39, 40, 41, 42, 43]. Breslau et al.[13], Robertson et al.[43], Damman et al.[7], Lloyd et al.[15], and Saigal et al.[39] all found that the differences in mean IQ between the NBW and LBW populations they studied were significant. Breslau et al.[13] found that LBW children scored, on average, 6.8 points lower than NBW children. This difference was attenuated to a 4.9 point difference when population site, maternal IQ, maternal education, and race were controlled. While this is a small difference in IQ (one standard deviation is 15 points), it translates into a large difference in the percentage of children falling more than one standard deviation below normal intelligence (85 points). In the same study the authors investigated the question of whether or not there is a gradient in effect. Their findings were positive. The odds ratios they computed showed that children born weighing less than 2,000g were approximately 2.7 times more likely to have an IQ below 85 than NBW children, whereas children weighing between 2,500g and 2,001g at birth were not significantly more likely to have IQ's lower than 85. The findings of this study indicate that LBW does correlate with reduced IQ.

Hack et al.[17] performed a regional study which compared NBW children to children with birth weights ranging from 750g-1,499g or from 500g-749g. The VLBW comparison group for the lowest group was matched by race, sex, site at which children were born, and by roughly when they were born. The NBW children who acted as comparisons were chosen from the same classrooms as the study group, and matched by sex, race and birth date within three months. In addition to birth weight, neonatal complications and various maternal risk factors were measured for each child. The VLBW group had 28% in the subnormal intelligence range (IQ < 85), 4 times as large a proportion as in the control group. The proportion in the group weighing less than 750g was 50%. In the studies by Rantakallio and vonWendt[40], Ross et al.[16], Hall et al.[41], and McCormick et al.[42], the proportions of ELBW children with IQ's below 85 were 20%, 22%, 23%, and 24% respectively. Rantakallio and vonWendt found that 12.6% of children between 1500g and 1999g fell into this category, as did 7.5% of children between 2000 and 2499g. Of children born NBW, however, only slightly more than 2% fell into this category. The literature shows that children born LBW are more likely to fall below the cutoffs for subnormal intelligence or mental retardation, and that they generally have slightly lower IQ's than their NBW counterparts.

This thesis will present the descriptive epidemiology of IQ, LP, and LD in a large, geographically defined cohort of children born weighing between 501 and 2000 grams. A goal will be to identify which personal factors predispose a child to IQ below 85, to LP, and to LD. The factors of interest include gender, SES, race, gestational age (GA) and birthweight.

METHODS

Study Population:

The population studied in this thesis has been described in previous publications[44, 45]. The Neonatal Brain Hemorrhage Study (NBHS) was originally assembled to study the causes and outcomes of germinal matrix/intraventricular hemorrhage in preterm infants. In an effort to develop a large geographically defined LBW cohort, all children born in, or transferred to any of the three major neonatal intensive care units (NICU's) in three New Jersey counties during a defined period were considered for entry. Three NICU's, St Peter's Medical Center, Monmouth Medical Center, and Jersey Shore Medical Center, share clinical responsibility for virtually all LBW babies born in Middlesex, Monmouth, and Ocean counties. In the period between September 1984 and June 1987 all the newborns weighing 501-2000g seen in these three NICU's were enrolled in the study. The sum of all births in the tri-county area during the period of enrollment amounted to 55,107. Of these, 1,318 (2.4%) were born weighing between 501 and 2000g. 1,105 of these 1,318 children were born in, or transferred to the centers in our study. Thus, 83% of all the LBW children born in the three counties were enrolled. Ninety percent of the children born weighing less than 1500g were reportedly enrolled.

Attrition and Exclusion:

The enrolled children were followed up at 2, 6, and 9 years of age. The measures of interest in this paper were collected at the 9 year follow-up (mean age 9.5 years) for the most part. The measure of home environment was collected at 6 years. The Bayley Scale of Infant Development, which was used to compare the followed children to those who were not tested at 9 years, was administered at 2 years. Due to the small number of children from any ethnic group other than Blacks and Whites (Latinos were the largest other group and numbered only 17), only Black and White children were used in these analyses.

Of the original 1105 children enrolled, 1041 (94%) were either Black or White. Of this group, 195 (19%) died before the age nine follow up, leaving 846 (81%). Due to loss to follow-up, only 629 Black or White children were followed up in any capacity at age nine. This represents 74% of the surviving Black and White group. A decision was made to include only children for whom IQ and the two achievement scores (Math and Reading) were available. These children number 471 (56% of the surviving Black and White children, 75% of 629). The pattern of the attrition is best illustrated by the flow chart in fig. 1.

In some cases a child was not functioning at a testable level and was therefore assigned the lowest score for that variable. Because means are not used in these analyses, the value assigned in such a case is not important except insofar as it is above or below a cut-off. The means that are reported exclude the assigned scores and so may overestimate the group's scores. All assigned values less than 2 SD below the mean. If an untested child could not accurately be assumed to fall into this group, he or she was excluded from the analyses.

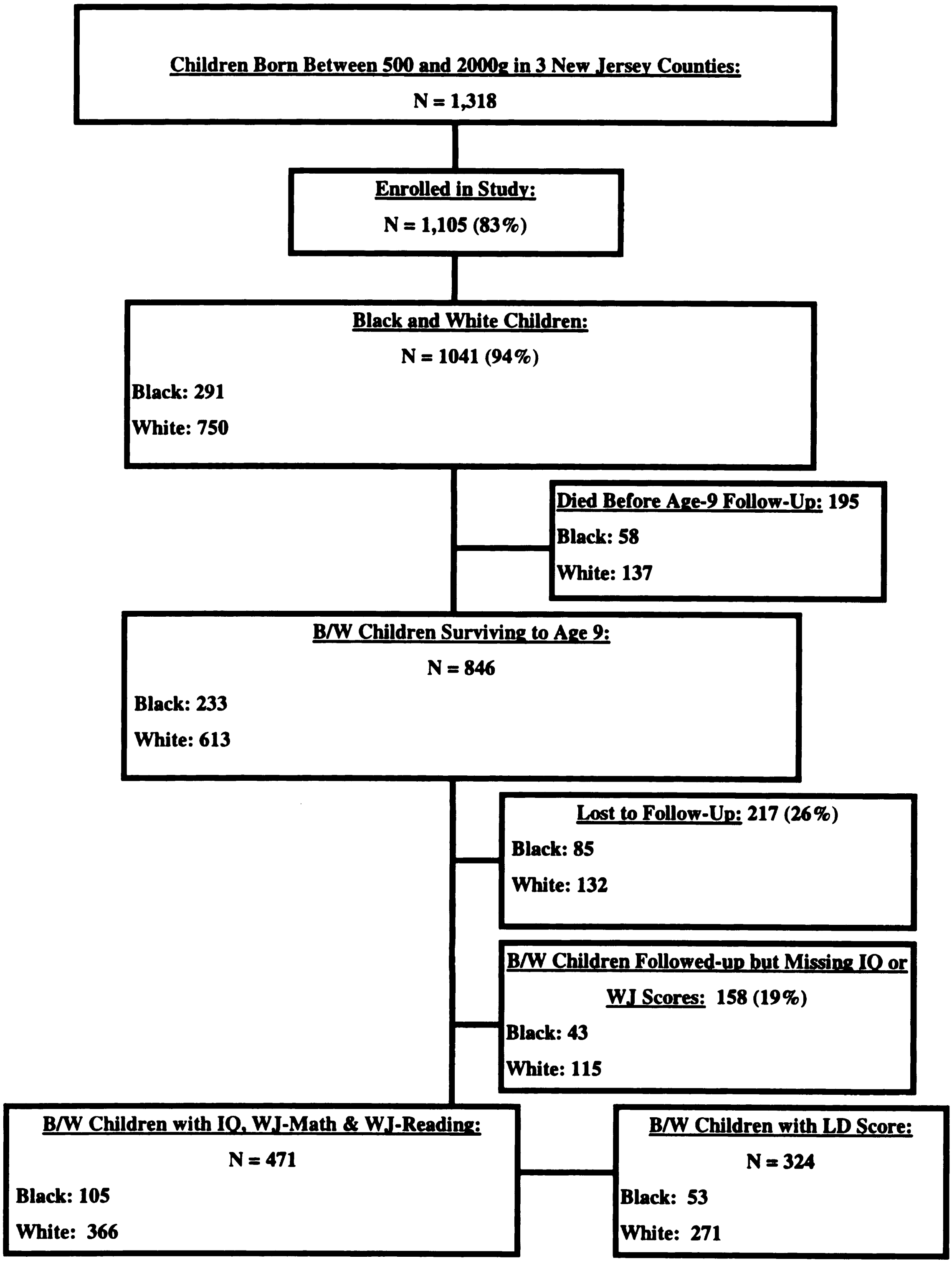

Figure 1 - Subject Enrollment, Attrition, and Exclusion

Instruments:

The variables used in these analyses are birthweight, gender, GA, race, 2 year developmental quotient (DQ), IQ at nine year follow-up, age by month at 9 year IQ testing, achievement in math and reading, LD, home environment, teacher ratings of children's performance in math and reading, and a diagnosis of disabling CP. Birthweight and GA were obtained at birth. These two variables were made into categorical variables. BWT was divided by 500g increments into 3 categories, and GA was divided into term children ($\geq$ 37 weeks), 33 -36 weeks, 29 - 32 weeks, and fewer than 28 weeks. GA was determined from prenatal ultrasound, prenatal medical records, postnatal maternal interview, labor and delivery records using a hierarchical algorithm described in Holzman et al.[46]. Race was reported by the mother.

IQ was assessed at the 9 year follow-up (1994 - 1997) using the Wechsler Intelligence Scale for Children III (WISC-III) which was standardized in 1991[47]. It provides a scale of intelligence scores, with a mean of 100 and a standard deviation of 15. The WISC-III has been used extensively in the field of child IQ testing. The reliability coefficients for internal reliability and interscorer reliability (when scoring is subjective) range between .90 and .98[47]. The validity of the test is argued in the WISC-III manual [47]. It correlates at levels between .59 and .92 with other non-Wechsler measures of ability, and at a .47 level with student grades. In addition to this, it is a valid instrument for this study because IQ is frequently determined with this measure and is in turn used to classify and categorize students at school. DQ was measured at age 2 using the Bayley

Scale of Infant Development[48]. This score, and the CP diagnoses were used to compare the children who were lost to follow-up to those included in the study. The CP diagnosis was made at the age 2 follow-up by trained nurse-practitioners and neurologists[see 49].

A child's achievement in a number of areas was measured using subscales of the Woodcock-Johnson Psychoeducational Battery (WJ) which was normed between 1986 and 1988[50]. It was normed on a large representative sample of US children. It shows internal consistency reliability coefficients in the mid .90's for the cluster scores. Concurrent validity (the tests' correlations with other tests of similar constructs) were mostly in the .80's[See 50]. As with the WISC-III, this test is a commonly used instrument in diagnosis.

The scales from the WJ used in this analysis are the reading and mathematics cluster scores. The reading score is a linear combination of the Letter-Word ID and Passage Comprehension subscales. The math score is likewise a combination of the Calculation and Applied Problem Solving subscales. All scores on the WJ can be normalized by age to a scale with a mean of 100 and a standard deviation of 15.

When dealing with children born prematurely, it is not unusual to norm a child's score to an age that takes into account the prematurity. For instance, a child born two months early is expected to perform like a newborn only after two months have elapsed. In this case, however, the raw scores were normed to the child's uncorrected age. Because the tests are designed to measure a cognate of academic achievement, and children enter school according to their birthdate rather than an age adjusted for their GA, uncorrected age was felt to be more accurate than corrected age.

Learning disability, as discussed previously, is measured here using the school-districts' designations. This information was gathered via a questionnaire given to the subjects' teachers. The state of New Jersey's designation "Perceptually Impaired", despite its name, corresponds to the general theoretical concept of LD. This is to say that a PI/LD child has been noted to perform below his or her potential in school for reasons other than sensory and physical deficits.

When the subjects were six years of age the Home Observation for Measurement of the Environment (HOME) scale was used to rate how well the child's home environment could be expected to foster learning and intellectual development. Using this instrument, trained observers measured such facets of the home environment as the parent's attitude toward the child's learning, the child's access to educational toys, and the degree to which the child was encouraged to express opinions.

Teacher ratings of children's performance were available for 324 children from the questionnaires filled out by the teachers when the children were 9 years of age. The scale on the teachers' questionnaire that provided this information allowed the teacher to rate a child as performing above grade level, at grade level, or below grade level in math or reading. For these analyses, the teacher ratings were dichotomized into below grade level vs. at or above grade level. These categories capture the information relevant to a study of negative scholastic outcomes, and parallel the distinction embedded in the WJ cutoff score. In short, the children who are performing at an admirable or even mediocre level are differentiated from those whose performance inspire worry. A child who was performing below grade level in reading will hereafter be designated as having a reading problem (RP) and a child performing below grade level in math will likewise be designated as having a math problem (MP).

ANALYTIC METHODS

An alpha of .05 was used in all these analyses except the analyses of difference among the followed and non-followed groups. For these analyses, an alpha of .10 was used instead in order to highlight possible sources of bias.

Assessing Possible Bias:

Using two sample t-tests and chi-squares, the 471 subjects were compared to the 375 Black and White children who were enrolled in the study at birth and survived to age 9 but were not included in further analyses because one or more outcome variables were missing. The variables on which they were compared are race, gender, BWT, and GA. They were compared to children seen at the 2 year, but not at the 9 year, follow-up on the prevalence of mental retardation and disabling CP

The same procedure was carried out for the subset of the subject children for whom the state-defined LD variable was available. They were compared to the children who could have entered this category had more information been available. As a child cannot be defined as learning disabled if he or she is in home-school (1 child), or is handicapped or mentally retarded (29 children), any such child whose LD status in the eyes of the school district had not been determined was not compared to the children for whom this variable was available. These children, even had the information been available, would not have been included in any analyses concerning LD.

Validating Learning Measures:

In order to gain some insight into the validity of the WJ measures as signifiers of low achievement, low WJ scores (< 85) were compared to teacher ratings of children's performance. Kappa coefficients were used to measure how well the scores agreed beyond chance. The Kappa measures chance corrected agreement. It is the difference in the number of cases the two methods agree upon and the number one would expect simply by chance, divided by the quantity 1 minus the number of cases one would expect the methods to agree upon solely based on chance [see 51].

Measuring Associations:

The relationships of birthweight category, race, gestational age category and gender to the four outcome measures (low IQ, low WJ-Math, low WJ-Reading, and LD) were measured using simple crosstabulations and chi-squares. These produced uncontrolled measures of effect. The same relationships were also estimated as odds ratios using multiple logistic regression analyses. Hierarchical logistic modeling was used. In accordance with suggestions made by Kleinbaum[52] regarding hierarchical modeling, the first step was to test for interactions and the next was to control for confounders while preserving accuracy and precision. As discussed previously, the variables which were considered necessary in all the initial models regardless of outcome include gender, HOME score, race, GA and of course birthweight. Age was also included in the model with IQ as the outcome because they were found to be negatively correlated. ($r = -.154$ $p < .001$). The achievement test scores are normed by age to the month and so it was not deemed necessary to include age in the models with achievement as an outcome.

RESULTS

Possible Bias:

The 471 subject children differed from the untested 9 year survivors on prevalence of both disabling CP and MR at age 2, and by race. They do not differ, however, by BWT, GA, or gender. The prevalence of disabling CP was 5% in the follow up-group, but 10% in the non-followed group ($p = .0005$). The follow-up group similarly had 4.5% of children with IQ < 70 at age 2, versus 11.9% in the non-followed group. Also, 34% of the untested survivors but only 22 % of the followed children were Black.

The comparison of the 335 children for whom the LD variable was available to the 449 survivors for whom no such information was available showed differences in racial composition and Bayley score at age 2. The children who were followed-up were less likely to be Black and had a mean DQ almost 10 points higher than their counterparts' (see Table 1).

Table 1: Comparisons of Children Followed-up vs. Lost

	Study Group	**Compariso n Group**	**p value**	**LD Group**	**Compariso n Group**	**p value**
N:	471	375	---	335	449	---
% Black:	22.3	34.0	.0002	16.4	36.3	≤ .0001
% With DCP:	4.9	10.4	.0049	Excluded from Analyses		---
% With MR:	4.5	11.9	.0008	Excluded from Analyses		---
Mean DQ at 2 Years:	106.8	93.9	< .0001	111.4	101.6	≤ .0001
Birthweight:	1496g	1461	.148	1501	1494	.44
Gestational Age:	31.7	31.58	.726	31.66	31.73	.67
% Female:	50.9	47.8	.389	51.3	48.3	.93

Validity of Achievement Measure:

The kappa coefficients relating the teacher ratings of a child's math or reading performance to the corresponding WJ cutoff score were both significant ($p < .001$). They show that the teacher ratings and WJ cutoffs agree at a rate better than chance. In the domain of math the two scores conform in 45% of those instances which would not be expected to agree simply by chance. For

Reading the proportion is 36%. In both math and reading domains, the teachers were generally more critical than the WJ cutoff criteria (see Tables 2 and 3).

This is appropriate if one considers that the teachers were asked to classify children into the broad categories of below, at, or above grade level, whereas the WJ cutoff of 85 points represents a significantly low level of performance. A child may be performing quite poorly but still not at the extreme level of 1 SD below the mean. Because of the definitional imprecision which attend psychological outcomes in general and academic outcomes in particular, extreme outcome measures such as the WJ cutoff scores are appropriate. The limited test of accuracy which is accomplished by measuring how reliably the WJ cutoff scores compare to the corresponding teacher ratings reflects favorably on the WJ scores as tests of academic achievement.

Table 2: WJ-Math <85 vs Teacher Reported Math Problem

		WJ Math (<85)		
		Norm	Low	Total
MP (from teacher forms)	-At or above Grade Level	242	6	248
	-Below Grade Level	52	34	86
Total		294	40	334

Kappa = .45 $p < .001$

Table 3: WJ-Reading < 85 vs Teacher Reported Reading Problems

		WJ Read (<85)		
		Norm	Low	Total
RP (from teacher forms)	-At or above Grade Level	256	.00	256
	-Below Grade Level	59	22	81
Total		315	22	337

Kappa = .36 $p < .001$

Table 4: Demographic Frequencies and Means Within Study Groups

	Whole Study Group	LD study group
N	471	335
% White:	77.7	83.6
% Female:	51.0	51.3
Mean Birthweight in grams:	1496.07 (348.96)	1501.30 (346.49)
Mean Gestational Age in weeks:	31.7 (3.05)	31.66 (2.83)
Mean HOME score:	47.5 (6.84)	48.2 (6.10)

numbers in brackets indicate standard deviations

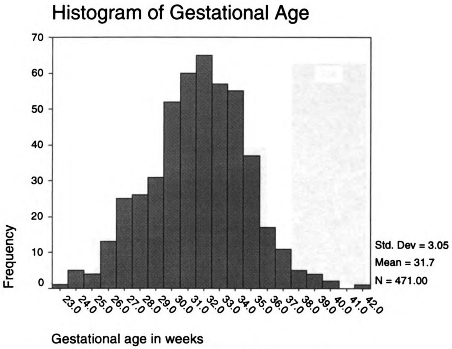

Figure 2 - Histogram of Gestational Age in Whole Group (N = 471)

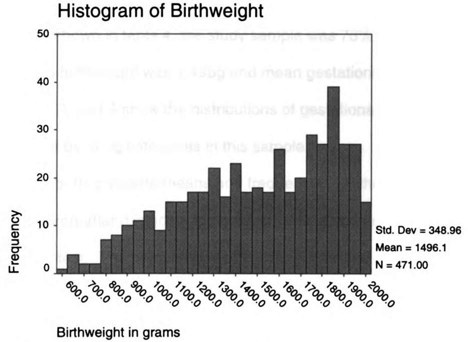

Figure 3 - Histogram of Birthweight in Whole Group (N = 471)

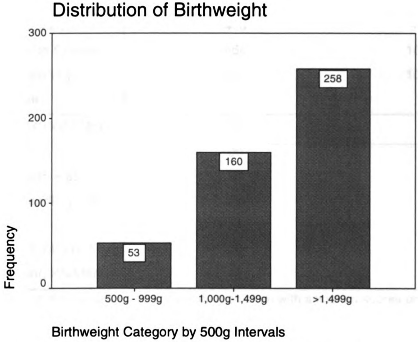

Figure 4 - Distribution of Birthweight in 500g Categories in Whole Group (N = 471)

Group Characteristics:

As shown in table 4, the study sample was 78% White and 51% female. The mean birthweight was 1,496g and mean gestational age was 31.7 weeks. Figures 2, 3, and 4 show the distributions of gestational age, birthweight, and birthweight by 500g categories in this sample.

Table 5 reports outcome means and frequencies for the study group of 471 LBW children after excluding the children whose scores were assigned due to inability to test. The percentages of children classified as below normal are 7.9% for the WJ-Reading Cluster, 13% for the Math Cluster, and 19.1% for full-scale IQ.

Table 5: Outcome measures - Means and Frequencies

Means†	**N**	**Mean (SD)**
Full Scale IQ:	468	97.9 (14.81)
WJ Math Cluster:	464	105.1 (17.38)
WJ Reading Cluster:	465	103.2 (13.40)
Frequencies	**#**	**%**
IQ < 85:	90	19.1
WJ-Math < 85:	61	13.0
WJ-Reading < 85:	37	7.9
LD:*	51	15.2
Math Problem:‡	86	25.7
Reading Problem:‡	81	24.0

† Means are calculated after excluding children with assigned scores on the relevant test.
* The denominator for this prevalence is 335, or all the children eligible to be considered for the designation of LD by their school district.
‡ Math and Reading Problem scores are based on samples of 334 and 337 respectively. They reflect a teacher's report that the child is performing below grade level in that domain.

Associations:

Learning disability (called Perceptual Impairment by the New Jersey Department of Education) is diagnosed in 15.2% of the children who were eligible to be considered for classification. This number may be contrasted to the 6.6% prevalence of the same Perceptual Impairment (PI) diagnosis found in 8-13 year old children in New Jersey as a whole[53, 54]. While this background statistic was derived from census information about the whole state, and therefore includes children born LBW, the LBW children would comprise less than 2.4% of the whole group. Thus, based on the outcomes found in this study and the information from the state of New Jersey, children born below 2000g seem to face 2.3 times the risk of other children of being diagnosed as LD by the school district by age 9 ($p < .0001$).

Tables 6 and 7 present unadjusted relative risks (RR) for different groups within the LBW sample. African-American children are at increased risk of showing a score of below 85 on IQ or either WJ measure when compared to Whites. They are not, however, statistically significantly more likely to be diagnosed as LD. This is theoretically acceptable, since the diagnosis of LD is usually based to some extent on the discrepancy between IQ and achievement scores. If both of these scores are lowered, it will not affect the discrepancy nor, by extension, the diagnosis of LD.

Boys face 1.94 times the risk of being diagnosed as LD that girls do. This conforms to the findings in the literature that boys are more likely than girls to be diagnosed as LD[4,17]. It is a less profound difference, however, than the RR of 4 reported by Lyon[4]. Why this should be is not clear. Boys are also 1.92

times as likely as girls to fall below 85 on the WJ-Reading scale. With respect to Math and IQ, boys and girls do not differ significantly.

Table 6: Proportions and Relative Risks of Poor Outcome by Gender and Race

	IQ < 85		**WJ - Math < 85**		**WJ - Reading < 85**		**LD**	
	%	**RR 95% CI**	**%**	**RR 95% CI**	**%**	**RR 95% CI**	**%**	**RR 95% CI**
Gender:								
Female	18.8	**Ref.**	10.4	**Ref.**	5.4	**Ref.**	10.5	**Ref.**
Male	19.5	**1.04** .72 - 1.51	15.6	**1.5** .93 - 2.41	10.4	**1.92*** 1.001 - 3.68	20.2	**1.94*** 1.14 - 3.3
Race:								
White	13.7	**Ref.**	10.9	**Ref.**	4.9	**Ref.**	14.3	**Ref.**
Black	38.1	**2.79*** 1.96 - 3.98	20.0	**1.83*** 1.13 - 2.96	18.1	**3.68*** 2.01 - 6.75	20.0	**1.4** .76 - 2.56

* $p < .05$

Birthweight had a significant relationship to the likelihood of falling below 85 on any of the three tests and to being diagnosed LD (see Table 7). For each outcome, the children born weighing the least were significantly more likely (between 2.14 and 3.65 times as likely) to have a negative outcome than the heaviest born children. While the children born weighing 1000-1,499g were not at significantly more risk of negative outcome than the heaviest group except on

Table 7: Proportions and Relative Risks of Poor Outcome by Birthweight and Gestational Age

	IQ < 85		**WJ - Math < 85**		**WJ - Reading < 85**		**LD**	
	%	**RR 95% CI**	**%**	**RR 95% CI**	**%**	**RR 95% CI**	**%**	**RR 95% CI**
Birthweight Categories:								
1500g- 2000g	15.9	**Ref.**	9.7	**Ref.**	4.7	**Ref.**	11.4	**Ref.**
1000g - 1499g	19.4	**1.2** .8 - 1.86	15	**1.5** .92 - 2.62	10	**2.15*** 1.04 - 4.43	18.3	**1.61** .92 - 2.81
501g - 999g	34	**2.14*** 1.34 - 3.42	22.6	**2.34*** 1.26 - 4.35	17	**3.65*** 1.62 - 8.23	25.7	**2.27*** 1.13 - 4.53
GA Categories:								
≥ 37 weeks	21.1	**Ref.**	15.8	**Ref.**	10.5	**Ref.**	9.1	**Ref.**
33 - 36 weeks	17.8	**.84** .34 - 2.12	12.2	**.77** .26 - 2.33	8.1	**.77** .19 - 3.11	13.2	**1.5** .21 - 9.85
29 - 32 weeks	17.5	**. 83** .33 - 2.09	12.2	**.77** .26 - 2.33	6.3	**.6** .15 - 2.5	15.9	**1.75** .26 - 11.81
≤ 28 weeks	27.3	**1.3** .50 - 3.37	16.7	**1.06** .33 - 3.40	10.6	**1.00** .23 - 4.45	21.4	**2.36** .33 - 16.67

* $p < .05$

WJ-Reading, chi-square tests of linear trend were significant for all outcomes. This indicates a graded relationship of BWT with these 4 outcomes. GA had no relationship to the outcomes.

Four logistic regression models were hierarchically fitted to the data. The initial models included BWT, GA, gender, race, HOME score, and interaction terms. None of the interactions that were tested were found to be significant. These were BWT by GA, BWT by race, and BWT by HOME score. In every model except that for LD the HOME score reduced the effect of race to non-significance. In the model for LD, race was not a significant factor to begin with. Even so, the inclusion of HOME in the LD model greatly reduced the beta-weight and significance of race. These findings are precisely in line with what Van Rossem et al. found with regard to IQ[10].

Table 8 presents the odds ratios for gender, BWT category by 500g increments, GA category by 3 month intervals, and HOME score quartile. The only final model in which GA was retained was the one for WJ-Reading. Thus, all the odds ratios relating the exposures to WJ-Reading are controlled for GA category, HOME score quartile, BWT Category, and gender. The other three models contain all these variables except GA category. In the IQ model, the OR's are also adjusted for age at the IQ test. Also presented for comparison are the unadjusted odds ratios.

Table 8: Adjusted† and Unadjusted Odds Ratios for BWT Categories, GA Categories, HOME Quartiles, Gender, and Race

	IQ< 85: (p-value)		WJ-M < 85 (p-value)		WJ-R < 85 (p-value) †		LD: (p-value)	
OR's	**Un-adjusted**	**Adjusted**	**Un-adjusted**	**Adjusted**	**Un-adjusted**	**Adjusted**	**Un-adjusted**	**Adjusted**
BWT Categories:	.641* p = .006	.60* p = .007	.606* p = .008	.561* p = .006	.484* p = .0015	.217* p = .0001	.60* p = .016	.55* p = .008
GA Categories:	.84 p = .27	1.07 p = .76	.912 p = .61	1.34 p = .22	.99 p = .95	2.09* p = .021	.75 p = .16	.99 p = .97
HOME quartiles:	3.02** p < .0001	3.10** p < .0001	2.01** p < .0001	2.10** p < .0001	2.67** p < .0001	2.99** p < .0001	1.55* p = .003	1.61* p = .002
Gender:	1.05 p = .84	1.42 p = .20	1.588 p = .095	2.10* p = .017	2.025* p = .049	4.424* p < .001	2.172* p = .014	2.57* p = .005
Race:	3.89** p < .0001	1.3 p = .357	2.04* p = .016	.789 p = .542	4.27** p < .0001	1.9 p = .166	1.5 p = .28	1.10 p = .60

† In the model for WJ-Reading all the exposures are adjusted for BWT Category, GA category, HOME quartile, and Gender. The effects all the exposures on all the other outcomes are controlled for all of the same variables except for GA category.
* Significant at p < .05
** Significant at p < .0001

DISCUSSION

This study followed a large, geographically determined group of LBW children to age nine. It is different from most research in this area in its inclusion of children between 1500 and 2000 grams. This inclusion provides a more complete picture of the problems associated with LBW. This is particularly true because the number of children born between 1500 and 2000 grams compose so large a portion of LBW children by comparison with those born at fewer than 1500 grams. The danger of conflating the findings for the smallest-born children with those for the larger-born children was overcome by splitting the group into the conventional BWT groups.

The findings of this study generally mirror what is seen in the literature. Birthweight has a significant effect on intellectual and academic outcomes independent of race, and environment. Also, the environment has an effect on the outcomes which is as strong as or stronger than the effect of birthweight. Unexpected findings in this study include the high achievement score means and the effect of the HOME score on race. A further anomalous finding was the significance of GA to the WJ-Reading outcome. These findings and the weaknesses of the study will be discussed here.

One of the unusual findings in these analyses is the inflation of the means for the two Woodcock-Johnson tests. Because some children were too low functioning to be tested, low scores were assigned to these children. Their scores were then excluded when calculating the means. This served to inflate the means, but it is not certain by how much. The means for the group when these children are included are 105 (SD = 19.4) for WJ-Math, 102 (SD = 15.3) for WJ-Reading, and 97.5 (SD = 15.4) for IQ. These are still higher than expected based on the literature, and whether they are too high, too low, or accurate is unknown. The means may also have been affected by the pattern of attrition. The children who were available to the 9-year follow up had higher IQ at 2 years, and it may be assumed that the study population in this paper represents the upper range of the LBW population. However, because there is no control group in this study to which the means may be compared, it cannot be known whether or not the LBW children are truly performing above average on the achievement tests.

A second interesting finding of this study may provide substantial aid to researchers who are trying to parse the distal variable, race, into the proximal variables that are relevant to cognitive and academic outcomes. The HOME score measure of home environment significantly reduced the strength of race as a predictor of intelligence or achievement. In the case of the WJ-Math score, the HOME score erased the predictive value of race in the model, and even switched the remaining non-significant odds ratio from protective for Whites to protective for Blacks. Too much could be made of this instance, but it demonstrates the degree to which race and the HOME score are intertwined. This news should be heartening to researchers for methodological reasons, quite apart from philosophical or political reasons. Race, after all, is a nebulous and ill-defined construct, whereas the HOME score provides a standardized and transparent measure which may take race's place in this domain of investigation. The political implications are too much to cover in this thesis, but the finding that the HOME score erases the effects of race should not be taken to mean that race has no effect on academic and intellectual outcomes. Insofar as opportunities are apportioned along racial lines in this country, and attitudes toward education and its benefits are influenced by ethnicity, race has an effect on achievement and IQ. The effect, however, appears to be mediated by societal thinking on race rather than by any simply genetic pathway.

The relationship of GA to the study outcomes was non-significant except in the model for WJ-Reading. Why this should be is unclear. It is additionally surprising that the effect should be opposite to what is expected, with children born earlier being at less risk of low reading achievement. This finding reflects the effect of controlling for BWT. When birthweight is not included in the model, increased GA reverts to a non-significant protective factor (OR = .92, p = .75). It most likely reflects the effects of fetal growth retardation. That is, at any low BWT, the older children are those who should have moved on to a higher BWT category.

The major problems of this study are attrition, the lack of a control group and lack of strong definitions. In the case of IQ and achievement scores, the definition of low performance suffered from the lack of a control group. As Wolke et al. demonstrated, the use of test norms from a time or region other than those of the subject group can result in errors of classification[35]. Nevertheless, the comparison of groups within the LBW group remain valid. A child born LBW is at greater risk of negative outcome if faced with an unstimulating home environment than if raised in a home which nurtures inquiry and education. This is consistent with the results of educational intervention tested by the Infant Health and Development Program[19], and suggests that, at any birthweight, environmental educational interventions should improve LBW's cognitive outcomes. Whether this holds true for NBW children is outside the ambit of this study, but this gap does not mitigate the validity of what is reported here.

In the case of LD, a comparison group could be derived from public records. However, the outcome measured is ill-defined and thus imprecise. The true extent and character of the relationship between LD and LBW will be very hard to determine until there is a better definition of the outcome. Nevertheless, insofar as the construct that was measured reflects some aspect of the truth, it appears that there is a relationship between LBW and LD. If the only problem with the definition is one of precision, the relationship between LD and birthweight should be stronger than what is reported here. Once an etiologic definition of LD's is available, the NBHS cohort will lend itself to closer study of this relationship.

The problem of attrition is a constant obstacle to epidemiologists. It serves to reduce power and possibly to introduce bias. In this study, however, the followed population is large enough that power is not greatly threatened. Also, the children who were not followed up were considerably more likely to have DQ's below normal than the children who were followed, and more likely to be Black. One would expect a bias to actually attenuate the finding that LBW children face a greater risk of LD than NBW children; though this is the case only if LD is associated with the same process that lowers DQ in LBW children. While this is not implausible, the possibility of bias should not be ignored. The gender differences in the outcomes should not be greatly affected by the attrition.

However, children in the group that was not included in the analyses are both more likely to be Black and more likely to show lower DQ. The relationship of these factors to one another is fairly clear in the followed group when IQ is substituted for DQ and the HOME score is introduced. However, there is no way to say that the same relationship would hold in the lost group.

In closing, the study outlined here shows results that are in line with the literature. Birthweight is associated with all the outcomes measured as is the child's home environment. Gender is associated with all the outcomes but IQ. Race was found to confound the relationship between the outcomes and home environment, which presents interesting possibilities for future research. The effect of bias due to attrition and exclusion seems not to be critical, and even the definitional problems affecting the outcomes do not present critical threats to the findings.

REFERENCES

REFERENCES

1 Tudehope, D.I., Burns, Y.R., Gray, P.H., Mohay, H.A., O'Callaghan, M.J., & Rogers, Y.M. Changing patterns of survival and outcome at 4 years of children who weighed 500-999g at birth. *Journal of Paediatrics and Child Health* (1995) 31:451-456.

2. Ornstein, M., Ohlsson, A., Edmonds, J., Asztalos, E. Neonatal follow-up of very low birthweight/extremely low birthweight infants to school age: A critical overview. *Acta Paediatrica Scandinavia* (1991) 80: 741-748.

3. Kiely, J. L., & Paneth, N. Follow-up studies of low-birthweight infants: Suggestion for design, analysis and reporting. *Developmental Medicine and Child Neurology* (1981) 23: 96-100.

4. Lyon, G.R. Learning Disabilities. *The Future of Children* (1996) 6,1: 54-75.

5. Kitchen, W.H., Rickards, A.L., Doyle, L.W., Ford G.W., Kelly, E.A., Callanan, C. Improvement in Outcome for Very Low Birthweight Children: Apparent or Real? *Medical Journal of Australia* (1992) 157: 154-157.

6. Lefebvre, F., Bard, H., Veilleux, A., Martel, C. Outcome at school-age of children with birthweights of 1000g or less. *Developmental Medicine and Child Neurology* (1988) 30: 170-180.

7. Damman, O., Walther, H., Allers, B., Schröder, M., Drescher, J., Lutz, D., Veelken, N., Schulte, F.J. Development of a regional cohort of very-low-birthweight children at six years: Cognitive abilities are associated with neurological disability and social background. *Developmental Medicine and Child Neurology* (1996) 38: 97-108.

8. Halsey, C.L., Collin, M.F., Anderson, C.L. Extremely low birth weight children and their peers. *Archives of Pediatrics and Adolescent Medicine* (1996) 150: 790-794.

9. Escobar, G.J., Littenberg, B., Petitti, D.B. Outcome among surviving very low birthweight infants: A meta-analysis. *Archives of Disability in Childhood* (1991) 66: 204-211.

10. Sternberg R.J. *Beyond IQ: A triarchic theory of human intelligence.* New York: Cambridge University Press, 1985.

11. Van Rossem R., Whitaker A.H., Feldman J.F., Pinto-Martin J.A. Black/white differences in cognitive abilities at age six in a low birthweight cohort: The role of social and medical risk factors. Unpublished, (1998).

12. Leonard, C.H., Clyman, R.I., Plecuch, R.E., Juster, R.P., Ballard R.A., Behle, M.B. Effect of medical and social risk factors on outcome of prematurity and very low birth weight. *Journal of Pediatrics* (1990) 116,4: 620-626.

13. Breslau, N., DelDotto, J. E., Brown, G. G., Kumar, S., Ezhuthachan, S., Hufnagle, K. G., & Peterson, E. L. A gradient relationship between low birth weight and IQ at age 6 years. *Archives of Pediatric and Adolescent Medicine* (1994) 148: 337-383.

14. Eilers, B.L., Desai, N.S., Wilson, M.A. Cunningham, M.D. Classroom performance and social factors of children with birth weights of 1,250 grams or less: Follow-up at 5 to 8 years of age. *Pediatrics* (1986) 77,2: 203-208.

15. Lloyd, B.W., Wheldall, K., Perks, D. Controlled study of intelligence and school performance of very low-birthweight children from a defined geographical area. *Developmental Medicine and Child Neurology* (1988) 30: 36-42.

16. Ross, G., Lipper, E.G., & Auld, P.A.M. Educational status and school-related abilities of very low birthweight premature children. *Pediatrics* (1991) 88,6: 1125-1134.

17. Hack, M., Taylor, H. G., Klein, N., Eiben, R., Schatschneider, C., & Mercuri-Minich, N. School-age outcomes in children with birth weights under 750g. *New England Journal of Medicine* (1994) 331,12: 753-759.

18. McCormick, M.C., Gortmaker, S.L., & Sobol, A.M. Very low birth weight children: Behavior problems and school difficulty in a national sample. *Journal of Pediatrics* (1990) 117: 687-693.

19. Cohen, S.E., Sigman, M., Parmelee, A.H., & Beckwith, L. Perinatal risk and developmental outcome in preterm infants. *Seminars in Perinatology* (1982) 6,4: 334-339.

20. The Infant Health and Development Program. Enhancing the outcomes of low-birth-weight, premature infants. *Journal of the American Medical Association* (1990) 263,22: 3035-3042.

21 . Sanford M.N. Offord D.R., Boyle M.H., Peace A, Racine Y.A. Ontario child health study: Social and school impairment in children aged 6 to 16 years. *Journal of the American Academy of Adolescent Psychiatry* (1992) 31,1: 60-66.

22. Seigel, L.S. I.Q. is irrelevant to the definition of learning disability. *Journal of Learning Disability,* (1989) 22,8: 469-486.

23. Rispens J., van Yperen, T.A., & van Duijn, G.A.. The irrelevance of IQ to the definition of learning disabilities: Some empirical evidence. *Journal of Learning Disabilities* (1991) 24,7: 406-412.

24. Fletcher, J.M., Francis, D.J., Rourke, B.P., Shaywitz, S.E., & Shaywitz, B.A. The validity of discrepancy-based definitions of reading disabilities. *Journal of Learning Disabilities* (1992) 25,9: 555-561.

25. Fletcher, J.M, Espy, K.A., Francis, D.J., Davidson, K.C., Rourke, B.P., & Shaywitz, S.E. Comparisons of cutoff and regression-based definitions of reading disabilities. *Journal of Learning Disabilities* (1989) 22,6: 334-355.

26. Kolligian, J. Jr., & Sternberg, R.J. Intelligence, information processing and specific learning disabilities: A triarchic synthesis. *Journal of Learning Disabilities* (1987) 20,1: 8-17.

27. Algozzine, B., & Ysseldyke, J.E. Questioning discrepancies: Retaking the first step 20 years later. *Learning Disability Quarterly* (1987) 10: 301-312.

28. Algozzine B., Ysseldyke J.E., & Shinn M. Identifying children with learning disabilities: When is a discrepancy severe? *Journal of School Psychology* (1982) 20,4: 299-305.

29. Evans, L.D. A conceptual overview of the regression discrepancy model for evaluating severe discrepancy between IQ and achievement scores. *Journal of Learning Disabilities* (1990) 23,7: 406-412.

30. Hammill, D.D. On defining learning disabilities: An emerging consensus. *Journal of Learning Disabilities* (1990) 23,2: 74-84.

31. Schuerholz, L.J., Harris, E.L., Baumgardner, T.L., Reiss, A.L., Freund, L.S., Church, R.P., Mohr, J., & Denckla, M.B. An analysis of two discrepancy-based models and a processing-deficit approach in identifying learning disabilities. *Journal of Learning Disabilities* (1995) 28,1: 18-29.

32. Shaw, F.S., Cullen, J.P., McGuire, J.M., & Brinckerhoff, L.C. Operationalizing a definition of learning disabilities. *Journal of Learning Disabilities* (1995) 28,9: 586-597.

33. Wilson L.R., & Cone T. The regression equation method of determining academic discrepancy. *Journal of School Psychology* (1984) 22: 95-110.

34. Cone T.E., & Wilson L.R. Quantifying a severe discrepancy: A critical analysis. *Learning Disability Quarterly* (1981) 4: 359-371.

35. Hammill D.D., Leigh J.E., McNutt, G., Larsen S.C. A new definition of learning disabilities. *Journal of Learning Disability* (1987) 22,2: 109-113.

36. Wolke D., Ratschinski G., Ohrt B., & Riegel K. The cognitive outcome of very preterm infants may be poorer than often reported: An empirical investigation of how methodological issues make a big difference. (1998).

37. Saigal S., Rosenbaum P., Szatmari P., Campbell D. Learning disabilities and school problems in a regional cohort of ELBW (<1000g) children: A comparison with term controls. *Journal of Developmental and Behavioral Pediatrics* (1991) 12,5: 294-300.

38. Vohr, B.R.& Coll, C.T.G. Neurodevelopmental and school performance of very low-birth-weight infants: A seven-year longitudinal study. *Pediatrics* (1985) 76,3: 345-350.

39. Saigal S., Szatmari P., Rosenbaum P, Campbell D., King S. Cognitive abilities and school performance of extremely low birthweight children and matched term control children at age 8 years: A regional study. *Journal of Pediatrics* (1991) 118,5: 751-760.

40. Rantakallio, P., vonWendt, L. Prognosis for low-birthweight infants up to the age of 14: A population study. *Developmental Medicine and Child Neurology* (1985) 27: 655-663.

41. Hall, A., McLeod, A., Counsell, C., Thomson, L., Mutch, L. School attainment, cognitive ability and motor function in a total Scottish very-low-birthweight population at eight years: A controlled study. *Developmental Medicine and Child Neurology* (1995) 37: 1037-1050.

42. McCormick, M.C., Brooks-Gunn, J., Workman-Daniels, K., Turner, J., Peckham, G.J. The health and developmental status of very low-birth-weight children at school age. *Journal of the American Medical Association* (1992) 267,16: 2204-2208.

43. Robertson C.M.T, Etches P.C., Kyle J.M., Eight year school performance and growth of preterm, small for gestational age infants: A comparative study matched for birth weight or for gestational age. *Journal of Pediatrics* (1990) 116,1: 19-26.

44. Pinto-Martin J., Paneth N., Witomski T., Stein I., Schonfeld S., Rosenfeld D., Rose W., Kazam E., Kairam R., Katsikiotis V., & Susser M. The central New Jersey neonatal brain haemorrhage study: design of the study and reliability of ultrasound diagnosis. *Paediatric and Perinatal Epidemiology* (1992) 6: 273-284.

45. Paneth N., Pinto-Martin J., Gardiner J., Wallenstein S., Katsikiotis V., Hegyi T., Hiatt I.M., Susser M.. Incidence and timing of germinal

matrix/intraventricular hemorrhage in low birth weight infants. *American Journal of Epidemiology* (1993) 137,11: 1167-1176.

46. Holzman, C., Paneth, N., Little, R., Pinto-Martin, J., Neonatal Brain Hemorrhage Study Team. Perinatal brain injury in premature infants born to mothers using alcohol in pregnancy. *Pediatrics* (1995) 95,1: 66 - 73.

47. Wechsler, D. *WISC-III [kit] : Wechsler intelligence scale for children* San Antonio [Tex.] : Psychological Corporation, Harcourt Brace Jovanobvich, 1991.

48. Bayley, N. *Bayley Scales of Infant Development.* San Antonio [Tex.] : Psychological Corporation, Harcourt Brace Jovanobvich, 1969

49. Pinto-Martin, J., Torre, C., Zhao, H. Nurse screening of low-birth-weight infants for cerebral palsy using goniometry. *Nursing Research* (1997) 46, 5: 284 - 287.

50. Woodcock, R. W. *Woodcock-Johnson Psycho-educational Battery--Revised.* Allen, TX : DLM Teaching Resources, 1990.

51. Streiner, D. L. & Norman, G. R. *PDQ Epidemiology - 2nd Ed.* St. Louis, MO: Mosby, 1996.

52. Kleinbaum, D. G. *Logistic Regression. A Self-Learning Text.* New York, NY: Springer, 1994.

53. New Jersey State Department of Education, Office of Special Education. Personal Communication (1998).

54. New Jersey State Data Center. Personal Communication (1998).