



Early Start

Evaluation Report



Nine year follow-up

Published October 2012
Ministry of Social Development
PO Box 1556
Wellington 6140
New Zealand
www.msd.govt.nz

ISBN 978-0-478-33563-7 (online)

Please send queries to the authors: dm.fergusson@otago.ac.nz

Early Start

Evaluation Report

Nine year follow-up

David Fergusson

Joseph Boden

John Horwood

Christchurch Health and Development Study

Department of Psychological Medicine

University of Otago, Christchurch

Christchurch, New Zealand

email: dm.fergusson@otago.ac.nz

October 2012

CONTENTS

LIST OF TABLES/ LIST OF FIGURES	III
ACKNOWLEDGEMENTS.....	IV
EXECUTIVE SUMMARY.....	1
CHAPTER 1 INTRODUCTION	4
1.1 The Early Start Programme	4
1.2 Evaluation of Early Start	5
1.3 Purpose of present report	5
CHAPTER 2 THE EARLY START SERVICE.....	6
2.1 The development of Early Start.....	6
2.2 Overview of Early Start and its principles	8
2.3 The pilot study.....	12
CHAPTER 3 JUSTIFICATION FOR AND PLANNING OF THE RANDOMISED TRIAL	13
3.1 Justification for a randomised trial of Early Start.....	13
3.2 Overview of the research design.....	14
3.3 Characteristics of children and families entering the randomised trial	15
CHAPTER 4 SUMMARY: RESULTS OF EVALUATION AT 36 MONTHS.....	19
4.1 Introduction.....	19
4.2 Assessment of child outcomes (0–3 years).....	19
4.3 Results for child outcomes.....	22
4.4 Assessment of parent and family outcomes (0–3 years).....	24
4.5 Results for parent and family outcomes.....	25
4.6 Programme outcomes for Māori and non-Māori participants.....	27
4.7 Summary and conclusions.....	28
CHAPTER 5 FINDINGS UP TO THE NINE YEAR FOLLOW-UP	31
5.1 Introduction.....	31
5.2 Methods.....	32
5.3 Outcomes for children 0–9 years.....	33
5.4 Parent and family outcomes to nine years follow-up	39
5.5 Benefits for Māori and non-Māori.....	41
5.6 Effects of sample attrition on study validity	42
5.7 Summary.....	48
CHAPTER 6 SUMMARY, CONCLUSIONS AND RECOMMENDATIONS.....	50
6.1 Introduction.....	50
6.2 Summary of child outcomes.....	50
6.3 Issues raised by the study findings.....	52
6.4 Policy implications of study findings.....	54
REFERENCES.....	56
APPENDIX: STATISTICAL METHODS USED IN CHAPTER 5.....	59

LIST OF TABLES

Table 3.1	Comparison between the Early Start and Control groups on social and demographic characteristics at baseline.....	15
Table 3.2	Comparison between the Early Start and Control groups on family socio-economic characteristics at baseline.....	16
Table 3.3	Comparison between the Early Start and Control groups on maternal childhood disadvantage.....	16
Table 3.4	Comparison between the Early Start and Control groups on parental adjustment.....	17
Table 3.5	Comparison between the Early Start and Control groups on pregnancy and childbirth characteristics.....	18
Table 4.1	Rates of participation in the Early Start trial and the Early Start service at enrolment and 6, 12, 24 and 36 months post-enrolment.....	21
Table 4.2	Child outcomes of the trial.....	23
Table 4.3	Family and parental outcomes at 3 years.....	26
Table 4.4	Ethnic status and child outcomes to 36 months.....	28
Table 5.1	Sample retention at baseline and at 3, 6, 9 years follow-up.....	32
Table 5.2	Hospital attendance for accidental injury (%), 0–9 years.....	34
Table 5.3	Severe/very severe physical assault by any parent (%), 0–9 years.....	35
Table 5.4	Mean (SD) physical punishment scores, 0–9 years.....	35
Table 5.5	Mean (SD) parenting competence scale scores, ages 5, 6 and 9 years.....	36
Table 5.6	Contact with services for child abuse/care and protection issues.....	37
Table 5.7	Mean (SD) parent reported SDQ externalizing, internalizing, and total scores, ages 5, 6 and 9 years.....	38
Table 5.8	Mean (SD) teacher reported SDQ externalizing, internalizing, and total scores, ages 5, 6 and 9 years.....	38
Table 5.9	Parent and family outcomes from 5 year to nine year follow-up.....	41
Table 5.10	Māori ethnicity and outcomes to nine year follow-up.....	42
Table 5.11	Comparison between the Early Start and Control groups observed at nine year follow-up on social and demographic characteristics at baseline.....	43
Table 5.12	Comparison between the Early Start and Control groups observed at nine year follow-up on family socio-economic characteristics at baseline.....	44
Table 5.13	Comparison between the Early Start and Control groups observed at nine year follow-up on maternal childhood disadvantage.....	44
Table 5.14	Comparison between the Early Start and Control groups observed at nine year follow-up on maternal adjustment.....	45
Table 5.15	Comparison between the Early Start and Control groups observed at nine year follow-up on pregnancy and childbirth characteristics.....	45
Table 5.16	Supplementary analyses of missing data.....	47

LIST OF FIGURES

Figure 3.1	CONSORT diagram of client recruitment and randomisation.....	19
------------	--	----

ACKNOWLEDGEMENTS

The research described in this report is the culmination of over 15 years' work by a large number of people and groups who have contributed to the development and evaluation of the Early Start programme. These include:

1. Our major funders: Child, Youth and Family, Ministry of Health, Canterbury District Health Board, Christchurch City Council, Trustbank Community Trust and the Health Research Council of New Zealand.
2. Other financial and non-financial support: Hyman Marks Trust, Mayor's Welfare Fund, McKenzie Trust, Telecom New Zealand and UMC Ltd.
3. Past and present members of the Early Start Board: Jenny Carter, David Close, Ricky Collins, Noel Doney, Jan Egan, Professor David Fergusson, Fraser Flanagan, Dr Rodney Ford, Dr Lynda Gray, Dr Mark Hunter, Kiwa Hutchen, Geoff Instone, Terehia Kipa, Shelley McCauley, Gillian Sinclair, Bernice Tainui, Sally Thompson and Dr Clive Wilson.
4. Past and present management and supervisory staff of Early Start: Hildegard Grant, Julie Armstrong, Jan Egan, Gabriella Hall, Carla Paton-Jones.
5. Past and present administration staff: Shirley Percasky, Shirley Smith and Robin Thomson.
6. The Family Support Workers who participated in this trial: Karen Austin, Carol Bartle, Donna Bennett, Anita Bradfield, Carolyn Collie, Carol Croy, Heather Davidson, Te Rangihau Dunn, Barbara Dyer, Monica Erasmuson, Teresa Foster, Clare Garrett, Lesley Glenn, Alison Gould, Robyn Graham, Annette Hearn, June Heeney, Jane Howe, Patricia Jamieson, Nikki Johnston, Kate Lacey, Deborah Lewthwaite, Glynis Matehaere, Adele Parkinson, Margaret Parks, Anne Robson, Shirley Roulston, Susan Smith, Michelle Sorenson, Maree Sullivan, Tupou Tavui, Sheryl Tuuta, Prue Urlwin, Corien van Daalen, Sharyn Wasley-Hill and Michelle Wood.
7. Finally, but most importantly, the 443 families who agreed to participate in this evaluation of the Early Start programme.

While we are indebted to all of the individuals and organisations listed above for their contribution to making Early Start a world leading family support service, the responsibility for any errors or omissions in the report is strictly that of the authors.

David M Fergusson

Joseph M Boden

L John Horwood

EXECUTIVE SUMMARY

The Ministry of Social Development commissioned this report to provide a nine year follow-up of the children and parents studied in the randomised controlled trial of the Early Start programme. It looks at the extent to which involvement in Early Start had benefits for the children and families enrolled in the programme. The report will sit alongside other reports that have evaluated the Early Start trial since it began in the mid-1990s.

Chapter 1: Introduction

Early Start is a home visiting service targeted at families with infants who are facing severe social, economic or emotional challenges. This report describes the results of an outcome evaluation of the Early Start home visiting service using a randomised trial in which 220 families receiving Early Start were compared with a Control group of 223 families not receiving the service. Both groups have been followed up over a nine year period to determine the extent to which children and families receiving Early Start gained benefits when compared with the Control group families.

The major findings of the evaluation are summarised in Chapters 2–6 of the report.

Chapter 2: The Early Start service

Early Start was set up in the mid-1990s as the result of a growing recognition within New Zealand of increasing rates of psychosocial problems in children. As a result of these concerns a consortium of Christchurch based organisations developed an intensive home visiting service that came to be known as Early Start. The initial consortium of providers included: The Christchurch Health and Development Study; the Family Help Trust; the Plunket Society; the Pegasus GP group; and Māori representatives.

Early Start provides a home visiting service that targets the most disadvantaged 15% of the population using a screening methodology based on that used in the Hawaiian Healthy Start programme.

The goals of Early Start focus on encouraging improvements in a number of areas of child and family wellbeing including: child health; maternal wellbeing; parenting skills; family economic functioning; and crisis management.

The service is provided by trained Family Support Workers who have professional qualifications in the areas of nursing, social work, teaching, or an allied profession. The extent of service delivery depends on the level of family need and varies from Level 1 (weekly home visits) to Level 4 (3-monthly home visits).

A pilot study of 51 client families found the service was well accepted by the client population and appeared to produce positive outcomes in a number of areas of child and family functioning.

Chapter 3: Justification for and planning of the randomised trial

The Board of Early Start was united in its view that the best method of evaluating the efficacy of the service was through a randomised controlled trial in which a group of families receiving Early Start was compared with a Control group of families not receiving Early Start. Available funding enabled a research design in which approximately 220 families receiving Early Start were compared with a Control group of 220 families not receiving Early Start.

Clients for the trial were recruited by Plunket nurses who screened all (4,523) clients in Christchurch over a 19 month period. The screening method provided a group of

443 client families who were eligible for Early Start and who agreed to enter the trial. These families were randomised into a group of 220 families provided with Early Start and 223 Control families.

As a group, families entering the trial were characterised by multiple disadvantages which included: limited educational achievement; low incomes; welfare dependence; adverse childhood circumstances; parental adjustment problems including crime; mental health and substance use problems; pregnancy smoking and unplanned pregnancy.

The families entering the trial were assessed using home based interviews at baseline, and again at six, 12, 24 and 36 months post-enrolment. In addition data was gathered from general practitioner and hospital records. Comparisons of the families in the Early Start and Control groups on a wide range of measures showed these groups were similar with respect to: social and demographic background; socio-economic conditions; maternal childhood; parental adjustment; and pregnancy/child birth history.

Chapter 4: Summary:

Results of evaluation at 36 months

Measures of a series of child related outcomes assessed up to 36 months post-enrolment were obtained from the data collected. These measures spanned: health outcomes; pre-school education and welfare use; parenting; child abuse and neglect; and child behaviour.

At the 36 months follow-up, children in the Early Start group had better child outcomes including:

- Greater contact with their family doctor (p <.05)
- Being more up to date with well-child checks (p <.05)
- Fewer hospital attendances for accidents/injuries and accidental poisoning (p <.05)
- Higher rates of enrolment with pre-school dental services (p <.01)
- Were enrolled in early childhood education for a longer period (p <.01)

- Higher rates of contact with community services (p <.01)
- More positive and non-punitive parenting (p <.05)
- Higher overall parenting scores (p <.01)
- Lower rates of severe physical assault by parents (p <.01)
- Lower rates of child behaviour problems (p <.05).

Effect sizes (Cohen's d) for the childhood outcomes above ranged from .19 to .31 with a median value of .26. These effect sizes fall into the range of small to moderate.

No benefits of Early Start were found for a range of parent and family related outcomes including:

- Maternal health and wellbeing
- Maternal substance use
- Family stability, family relationships and family violence
- Family economic and material wellbeing
- Family stress and adversity.

The outcomes of Early Start were similar for Māori and non-Māori families enrolled in the programme.

Chapter 5:

Findings up to the nine year follow-up

Families and children in Early Start were assessed again at ages five, six and nine years using parental interviews, hospital record data and teacher questionnaires. Up to the nine year follow-up, children of families enrolled in Early Start had:

- Lower rates of hospital attendance for non-intentional injury (accidents) (p <.01). These differences were most marked for the 0–3 year period.
- Lower rates of parental reported physical child abuse (p <.01). These differences were most marked for the 0–3 year period.
- Lower rates of parental reported punitive parenting (p <.05).
- Higher rates of parental reported competent parenting (p <.0001).
- Fewer parental reported childhood problem behaviours (p <.05).

Effect sizes (Cohen's d) for the childhood outcomes above ranged from .13 to .31 with a median of .26. These effect sizes fall into the range of small to moderate.

The outcomes were similar for Māori and non-Māori families enrolled in the Early Start programme.

There was no evidence to suggest Early Start had benefits for a range of parental and family outcomes that included: maternal depression; parental substance use; family violence; family economic circumstances; family stress and adversity.

Statistical analyses showed the differences in rates of sample retention for the Early Start and Control groups were unlikely to threaten study validity.

Chapter 6: Summary, conclusions and recommendations

The findings of the randomised trial show that, up to the three year follow-up, children in Early Start received a number of benefits including: greater use of health services; reduced rates of hospital attendance for childhood accidents; greater use of pre-school education and dental services; lower rates of parental reported childhood physical abuse; less punitive and more positive parenting; and lower rates of childhood problem behaviours.

By the nine year follow-up, there was evidence to show the children from families provided with Early Start had:

- Rates of hospital attendance for childhood accidents approximately 33% lower than those for the Control group (Early Start = 28%; Controls = 42%)
- Rates of parental reported physical child abuse more than 50% lower than those for the Control group (Early Start = 9.8%; Controls = 21.8%)
- More positive mean scores on measures of punitive parenting and parenting competence
- Lower mean scores on measures of parental reported child behaviour problems.

The evidence also showed Early Start had similar beneficial effects for Māori and non-Māori families.

While issues relating to sample selection, sample retention and measurement pose possible threats to the validity of the study findings, the weight of the evidence suggests Early Start has beneficial effects for a series of child related outcomes spanning health, pre-school education, service utilisation, child abuse, parenting, and child behaviour.

There was consistent evidence showing the provision of Early Start did not have any benefit for a wide range of parental and family outcomes. The lack of benefit of Early Start for parental and family outcomes highlights the importance of developing better links and integration between home visiting services such as Early Start and a wide range of other family related services. These services include: family planning and contraceptive advice; adult mental health services; educational and career support; family budgeting services; and family relationship services.

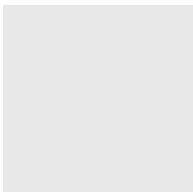
While Early Start has undergone substantial changes in both client referral methods and programme content, the weight of the evidence suggests the findings of the randomised trial are likely to apply to the present day Early Start service.

The key elements contributing to the success of Early Start in addressing child related outcomes are likely to include:

- The research base of the programme
- The use of professionally trained staff
- The development of standards and service manuals for the programme.

Whether the findings of this evaluation will apply to the services offered by Family Start providers is unknown and should be the subject of further research.

CHAPTER 1: INTRODUCTION



This report was commissioned by the Ministry of Social Development with the aim of providing a long term follow-up of the children and families studied in the randomised controlled trial of the Early Start programme. The background to the report is developed below.

1.1 The Early Start Programme

Early Start is a Christchurch based home visiting programme targeted at families with infants who are facing severe social, economic or emotional challenges. The service centres on a Home Visiting model in which families are assigned to a trained Family Support Worker who works with the family for up to five years to address family issues. The service is delivered by home visiting provided by Family Support Workers who have tertiary level qualifications in nursing, social work or a related discipline. Typically, the case loads for full-time staff contain between 10 and 20 families, depending on each family's level of need and progression through the programme.

The Early Start programme aims to produce positive changes in a wide range of outcomes spanning child and family health and wellbeing. These outcomes include:

- Improving child health
- Reducing child abuse
- Encouraging stable positive partnerships
- Improving parenting skills
- Supporting parental physical and mental health
- Encouraging family economic and material wellbeing.

A more detailed account of the development and delivery of the Early Start service is provided in Chapter 2 of this report.

1.2 Evaluation of Early Start

The evaluation of Early Start involved three stages.

1.2.1 Initial pilot research

This research involved the in-depth study of a pilot group of 51 clients over an 18 month period. The aims of this pilot were to ascertain the feasibility of setting up an intensive home visiting service in Christchurch and to assess both client outcomes and client satisfaction with the service (Fergusson et al., 1998).

1.2.2 Randomised controlled trial

This research involved a randomised controlled trial (RCT) in which 220 families receiving Early Start were compared with 223 Control families. Briefly, the study involved the following process:

- A total of 443 client families were recruited for the trial by Plunket nurses throughout the Christchurch urban region.
- Following their signed consent to enter the trial, client families were randomly assigned to Early Start or Control groups.
- Those in the Early Start group were offered the Early Start programme; those in the Control group were provided with existing child health and related services.
- The outcomes of the Early Start and Control groups were assessed by home interviews conducted with both groups at baseline, six months, 12 months, 24 months and 36 months.

The findings from this research have been summarised in a series of publications (Fergusson et al., 2005a and 2005b; Fergusson et al., 2006a and 2006b).

1.2.3 Long term follow-up

Following the initial RCT, families in both the Early Start and Control groups were interviewed and assessed when their children

were aged five, six and nine years. The aims of these assessments were to examine the extent to which any benefits of Early Start observed in the first-stage RCT were sustained in school-aged children.

1.3 Purpose of present report

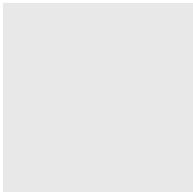
The primary purpose of this report is to present the results of the long term (up to nine years) follow-up of the families involved in the initial randomised controlled trial (RCT). These findings have not been published previously. To place these findings in context, the report also provides a review of the earlier findings of the RCT up to when the children were aged three years.

The structure of the report is as follows:

- Section 2 provides an overview and description of the Early Start programme. This account draws heavily on the presentation used in a previous report (Fergusson et al., 2005b).
- Section 3 provides a description of the overall research design including client recruitment, sample retention, measurement and assessment methods.
- Section 4 summarises the findings on client outcomes up to when the children were aged three years. This summary draws heavily on the findings presented in previous reports (Fergusson et al., 2005a and 2005b; Fergusson et al., 2006a and 2006b).
- Section 5 presents the findings from assessments made when the children were five, six and nine years.
- Section 6 provides a summary of findings, general conclusions and recommendations.

In general, the report aims to provide an overview of the extent to which involvement in the Early Start programme had benefits for children and families up to the nine year follow-up.

CHAPTER 2: THE EARLY START SERVICE



2.1 The development of Early Start

The impetus for developing the Early Start programme began in the early 1990s as a result of a growing recognition within New Zealand of the increasing rates of psychosocial problems in children. These issues spanned disruptive behaviour patterns and truancy (Report of the Education and Science Committee, 1995); adolescent substance use and abuse (Drugs Advisory Committee, 1995; Howden-Chapman et al., 1994; Public Health Group, 1996); child and adolescent mental health (McGeorge, 1995; Ministry of Health, 1994); and youth suicide (Barwick, 1992; Coggan and Norton, 1994; Ministry of Health, 1994). It became apparent these problems frequently overlapped and frequently involved a relatively small minority of children who came from disadvantaged, dysfunctional and often chaotic home environments.

These issues were highlighted in a study conducted by the Christchurch Health and Development Study that used longitudinally collected data to study the childhood history of a group of young people who had developed severe behavioural difficulties by the age of 15 years (Fergusson et al., 1994). This analysis revealed, in nearly all cases, the presence of childhood and family histories marked by a wide range of disadvantages and difficulties including socio-economic disadvantage, family conflict and instability, impaired child rearing practices, limited childhood experiences and restricted life opportunities. The most striking finding of the study was that young people reared in the most disadvantaged 5% of the cohort had

risks of severe maladjustment that were over 100 times the risks for young people in the most advantaged 50% of the cohort. This result clearly implied there was a need to address the difficulties and stresses faced by children reared in severely disadvantaged, dysfunctional or chaotic home environments, if substantial progress was to be made in addressing childhood and adolescent problems.

Traditional solutions to addressing the problems of at-risk families have largely centred around income maintenance or similar programmes that attempt to improve the economic wellbeing or material standards of high-risk families. However, an inspection of the childhoods of multiple problem children in the Christchurch Health and Development Study clearly suggested it was unlikely economic initiatives by themselves would address the many social, emotional and personal problems faced by these high-risk families. For this reason the search for solutions began to move away from focusing on the provision of traditional income support services and towards identifying programmes that provided at-risk families with direct support in the areas of parenting, child rearing and life skills.

A turning point in this process came at a conference convened by the Mental Health Foundation in 1994. At this conference, participants agreed that future programmes needed to focus on methods of home based visiting designed to meet the needs of at-risk families. It was suggested that the Hawaii Healthy Start programme provided a model that might be adapted to the New Zealand context. Healthy Start is a Hawaiian programme that has been in existence for over 20 years (Daro, 1994; Hawaii Department of Health, 1992). This programme involves two stages: population screening and service delivery. In the first stage, mothers giving birth are screened using standardised screening measures to identify at-risk families. Families meeting specified criteria are offered the Healthy Start programme. Families who accept the offer (between 80% and 90% of those eligible for the programme) are provided with intensive family support by a Family Support Worker.

In late 1994, representatives of the Family Help Trust and the Christchurch Health and Development Study met to discuss the possibility of developing a home based family support programme modelled along the lines of Healthy Start. It was agreed this would be desirable and an important first stage of programme development was to conduct a process evaluation of the programme. This was to be done by enrolling a group of 50 families into a pilot project aimed at assessing the extent to which the principles underlying Healthy Start could be adapted to a Christchurch social context. Key issues to be examined in this pilot study included:

- Could ethically acceptable methods be developed to identify at-risk families?
- Was it possible to develop an effective, culturally appropriate and non-stigmatising home visiting programme to meet the needs of at-risk families?
- How effective was this approach in leading to improvements in the wellbeing of children, including child health, parenting and life opportunities?

The group faced two hurdles in translating this plan into a viable project. The first was to find an effective method of identifying families at risk. The initial exploration of this issue suggested the most promising systematic method for identifying at-risk families was through Plunket nurses. In Christchurch, Plunket nurses see an estimated 95% of mothers shortly after birth and the Royal New Zealand Plunket Society (the Plunket Society) has developed strong links with other service providers to ensure



at-risk families are visited. For these reasons, the emerging Early Start group contacted the Southern Regional Office of the Plunket Society to enlist its cooperation in the project. After a period of negotiation, the Southern Regional Office agreed to become a member of a consortium of providers whose aims were to examine the feasibility of developing a family support service targeted at high-risk families and based on the principles of Healthy Start.

The second hurdle was to get funding for programme development. In the first instance, the consortium was successful in obtaining initial funding to support the project from the Canterbury Trustbank Community Trust. This funding enabled the consortium to develop concrete plans to develop a home based family support service for at-risk families for over 20 years, with the goal of reducing the likelihood of child maltreatment.

It was recognised that the success of any such service would depend critically on the extent to which the service was seen as culturally appropriate and relevant by Māori. To put in place mechanisms to ensure the programme was developed in a way acceptable to Māori, the Early Start consortium invited two Māori representatives (Mrs B. Tainui and Mrs T. Kipa) to join the consortium as directors. Both Mrs Tainui and Mrs Kipa had extensive experience in issues relating to Māori health, and particularly child health, and both had served as advisors and consultants to the Plunket Society. In addition to her role as a director of the Early Start programme, Mrs

Tainui was appointed as Kaumātua to the programme.

In early 1995, two representatives of the Family Help Trust visited the Hawaii Healthy Start programme to learn first-hand about the methods of screening and the service delivery used in Healthy Start. In the same year, the development of Early Start received a considerable impetus from a nationwide tour made by Dr Calvin Sia and Ms Gail Breakey from the Hawaii Healthy Start programme. In this visit, the representatives provided an overview of the Hawaiian programme and its underlying principles.

Further support was given by the Southern Regional Health Authority, who provided the Early Start programme with further funding to develop service provision in this area. In this process, the Southern Regional Health Authority also recommended the consortium be expanded to include representatives of the Pegasus Health GP group, thereby ensuring close links between the programme and general practitioners.

The net result of this process was that by mid-1995 a consortium of providers had been assembled, including the Family Help Trust, the Christchurch Health and Development Study, the Plunket Society, Māori representatives and Pegasus Health. Key staff from the Family Help Trust had received preliminary training in Hawaii and the consortium had gathered sufficient funding to support a pilot project based around a group of 50 families.

By October 1995, the consortium was in a position to recruit staff, provide staff training and to enrol families in the programme.

2.2 Overview of Early Start and its principles

While the development of Early Start was inspired by the work of Healthy Start, it is important to recognise Early Start was not an attempt to transplant an overseas programme into a New Zealand context. Rather, the aims of the Early Start consortium were to adapt the general principles of the Healthy Start programme to a Christchurch context. The key features of the Early Start programme are described below.



2.2.1 Client identification

The method of client identification used by Early Start involved a three-stage process. In the first stage, Plunket nurses applied broad and general screening criteria to identify at-risk families. Any family meeting these criteria was referred to Early Start. In the second stage, families were enrolled in Early Start for a 1 month probationary period. This period gave the family an opportunity to become acquainted with the programme and gave the programme an opportunity to learn about the family. In the third stage, an in-depth needs assessment of the family was made and families meeting pre-specified criteria were invited to join the programme on a longer term basis. At each stage of this process, signed consent was obtained from families to ensure the families were enrolled in the programme on an informed basis.

This system of client identification was designed to steer a middle course between the population based screening methods used by Healthy Start and the need for a more in-depth assessment to identify at-risk families, while treating families in an ethical and non-stigmatising way. This was achieved by developing a client identification system that combines elements of population screening, client referral and needs assessment to identify at-risk families. This approach has advantages and disadvantages when compared with the population based screening method used by Healthy Start. As noted above, the major advantage of this approach is that it avoids many of the difficulties that arise in the application of population based screening methods (such as false positive referral and possible stigmatisation or labelling). It ensures clients are enrolled in the programme on the basis of a comprehensive needs assessment rather than on the results of a screening measure. The potential disadvantage of the multi-stage process is that it provides multiple opportunities for families to decline services before they have been fully informed about these services.

2.2.2 Service provision

Early Start provides a system of home based family support and visits provided by trained Family Support Workers. The role of Family Support Workers is to support, empower and assist families to address a wide range of issues relating to child rearing, parenting and family functioning. An important feature of the programme is that the services provided to families are tailored to meet the family's particular circumstances and needs and not based on a predetermined programme that assumes one size will fit all. This flexible service provision makes it difficult to provide a concise account of the work of Family Support Workers. Nonetheless, the essential features of service provision can be summarised by noting that the work of Family Support Workers is directed at encouraging positive family change in the following areas:

- 1) Child health: Ensuring children have adequate access to and use of child health services, including immunization, preventive health care and timely visits for childhood morbidity. The key features of the service that lead to the achievement of this goal include: a) ensuring all families are enrolled with a single general practitioner who acts as the health care provider for the family; b) supporting and encouraging mothers to use child health care services; and c) developing close liaison and links with key health care providers including general practitioners, Plunket nurses and other services.
- 2) Maternal wellbeing: Ensuring the physical, social and emotional health of the child's mother is supported, protected and sustained. It is almost self-evident that good maternal functioning is a prerequisite for effective and positive child rearing. A large amount of the work of Family Support Workers involves providing social, emotional and practical support for mothers. This function spans a wide range of activities that may include support for the mother in dealing with issues of marital or partnership difficulties, family violence, substance abuse, maternal mental health problems and other sources of social and emotional stress.

- 3) **Parenting skills:** Helping mothers acquire and develop adequate parenting skills. As described in Chapter 4, many of the mothers enrolled in Early Start have experienced socially and emotionally impoverished childhoods. These childhood experiences have often provided them with limited opportunities to learn adequate parenting skills. A major role of Family Support Workers is to provide advice, support and role models to help and encourage mothers to acquire adequate parenting skills.
- 4) **Family economic functioning:** Improving family economic functioning. Poverty and/or depressed material conditions are common among families enrolled in Early Start. These difficulties appear to arise from two sets of factors that conspire to place families at risk of poverty and material hardship. First, the majority of families are on low incomes, and second, many families have limited budgeting and financial management skills. This combination of limited incomes and poor management skills makes families vulnerable to a wide range of economic problems and difficulties. An important function of the family support provided by Early Start is to help families to reduce the levels of economic stress and difficulty they face. These issues are addressed by on-going attempts by Family Support Workers to: a) encourage families to seek budget advice and to develop financial management skills; b) encourage families in debt to reduce their debt burden (and particularly hire purchase commitments); c) help families to find accommodation and household goods they can afford; and d) encourage mothers to reduce their welfare dependence and to supplement family income by part-time employment, where applicable.
- 5) **Crisis management:** Supporting and assisting with family crises. Families enrolled in Early Start are crisis prone, owing to their limited economic circumstances and personal backgrounds. An important function of Family Support Workers is to act as a source of support, advocacy and mentorship in times of

family crises. Key areas in which such crises emerge include marital relationships, family economic problems, substance abuse, family violence and difficulties with the law.

2.2.3 Case load and extent of service provision

Providing adequate family support to high-risk families is labour intensive. Owing to the demands of providing in-depth support, Family Support Workers have caseloads of approximately 15 families. The size of a caseload may vary depending on the mix of families within the caseload. The provision of family support is designed to follow a sequence in which the extent of support and assistance reduces as family change increases. The Early Start programme is aware of the need to encourage independent family functioning and of the risks of families becoming dependent on Family Support Workers. To reflect the process of transition over the course of the programme, service provision is organised into a series of levels reflecting the needs of families. These levels are:

- Level 1: All clients enter the Early Start programme at Level 1. This level requires a time allocation to the client of 2 hours a week and involves weekly home visits.
- Level 2: Clients who have spent some time in Early Start and are making progress in addressing difficulties move to Level 2. This level requires a time allocation of 1 hour a week for the client and one home visit a fortnight.
- Level 3: This level of home visiting is for families who have made substantial progress in addressing family problems and who are meeting their child's needs well. Families on this level receive a time allowance of half an hour a week and one home visit a month.
- Level 4: This level is for families who have become self-reliant and who are able to address their problems without support. Families at this level receive a home visit every 3 months to maintain contact with the programme and to confirm progress is being sustained.



In addition to the above service levels, some families facing severe crises or difficulties may be allocated to additional services that require at least 2.5 hours contact a week with the family and more than one home visit a week. This level is most commonly used in cases when families first enter the programme and where there is on-going concern children are at serious risk of child abuse or neglect.

2.2.4 Staff selection, training and supervision

An important feature of Family Support Workers is that they do not provide a specialist service such as that provided by nurses, social workers, counsellors and similar professionals. Rather, Family Support Workers act as family mentors and advocates who help a family to address the day to day problems it encounters. These job demands require Family Support Workers to have a sound training in a relevant discipline such as nursing or social work/ services coupled with the interpersonal skills and abilities to engage families in the Early Start programme. In addition, it is important workers have an understanding of the Treaty of Waitangi and an awareness of cultural issues. The Early Start programme recognises the right of Māori clients to have access to Māori Family Support Workers and encourages Māori Family Support Workers to develop links with local iwi, hapū and other relevant organisations.

The selection of Family Support Workers is conducted by a panel that includes the general clinical manager (Mrs H. Grant) and at least one of the Māori directors (Mrs B. Tainui and Mrs T. Kipa). Skills sought include: a) evidence of a relevant educational background; b)

awareness of cultural issues and obligations under the Treaty of Waitangi; c) experience in dealing with high-risk families; and d) evidence of good interpersonal skills and sound judgement.

To provide workers with a general background to their task, Early Start has devised a 4-week training programme which provides a background on a wide range of issues relevant to family support work.

The Early Start programme places considerable emphasis on the regular supervision and support of workers. There are two reasons for this emphasis. First, the task of dealing with the problems of high-risk families can often be very stressful and workers need regular supervision and support to reduce these burdens. Second, regular supervision ensures the Early Start services are delivered in a uniform way and workers are clearly advised about the boundaries of their role. To achieve these objectives each Family Support Worker receives a period of 2 hours clinical supervision a week from trained clinical supervisors. In these sessions, each case in the worker's caseload is reviewed, case notes are prepared and checked, and forward planning for each client family is discussed. In addition, these sessions provide an opportunity for workers to discuss particular issues that are of concern to them.

2.3 The pilot study

The first phase of the development of Early Start involved a pilot study in which a group of 51 families was enrolled in the programme for a period of 18 months. The findings from this pilot study have been described in a previous report (Fergusson et al., 1998). This report examined a series of issues relating to the Early Start programme, including client recruitment, service provision, client outcomes and client satisfaction. The major findings of this study were summarised as follows:

“The data gathered in this study support four major conclusions about the Early Start programme. First, that the client identification methods used by the programme produced an acceptable level of programme participation. Second, that the Early Start programme had developed an organisation and infrastructure that provided for consistent home visitation, supervision of service provision and

linkages with other provider organisations. Third, there were apparent benefits of the programme for client families, with these benefits being most evident for child health care and parenting, and least evident for family economic functioning. Fourth, that the programme was seen as supportive and culturally appropriate by its clientele. However, whilst the results of this evaluation are generally positive, it is important to note that they fall far short of demonstrating the benefits of the programme conclusively. Such evaluation requires a randomised field trial in which a group of families receiving the programme is contrasted with an equivalent group of families not receiving the programme. The present report, however, supports the view that the progress made in the development of Early Start is sufficiently promising to justify the development of such a field trial.”(Fergusson et al., 1998, p. 7.)



CHAPTER 3: JUSTIFICATION FOR AND PLANNING OF THE RANDOMISED TRIAL



This chapter provides the justification for the randomised trial and an overview of the research design employed, including the methods of randomisation, timing of assessment, and measurement methods.

3.1 Justification for a randomised trial of Early Start

At the inception of the Early Start programme, the Board of Early Start was committed to the view that a critical component of programme development was a thorough and rigorous evaluation of the effectiveness of the programme in achieving its goals. The Board considered the most compelling evidence for programme effectiveness would come from a randomised trial. In this trial a group of children and families receiving the Early Start programme would be contrasted with an equivalent group of children and families not receiving the service on a series of measures that reflected the goals of the Early Start service.

The overall research design employed was largely dictated by the funding provided to Early Start under the original Family Start¹ funding. That funding provided sufficient support for services for up to 220 families over a three-year period, thus setting the maximum size of the experimental group to 220 families. From this basis the evaluation group – the Christchurch Health and Development Study (CHDS) – devised a research design to compare the outcomes of 220 families receiving Early Start with a randomly assigned Control group of 220 families not receiving Early Start.

¹ Family Start is the Government's home based support service for families with high needs. It is operating in 31 sites across New Zealand.

This research design was submitted to the Health Research Council of New Zealand and received funding approval for assessments conducted in 1996, 2000 and 2003. The study was also assessed in 1999, 2001 and 2004 by the Canterbury Ethics Committee which gave approval for the ethical approach at each phase of the trial.

3.2 Overview of the research design

3.2.1 Recruitment of clients and randomisation to groups

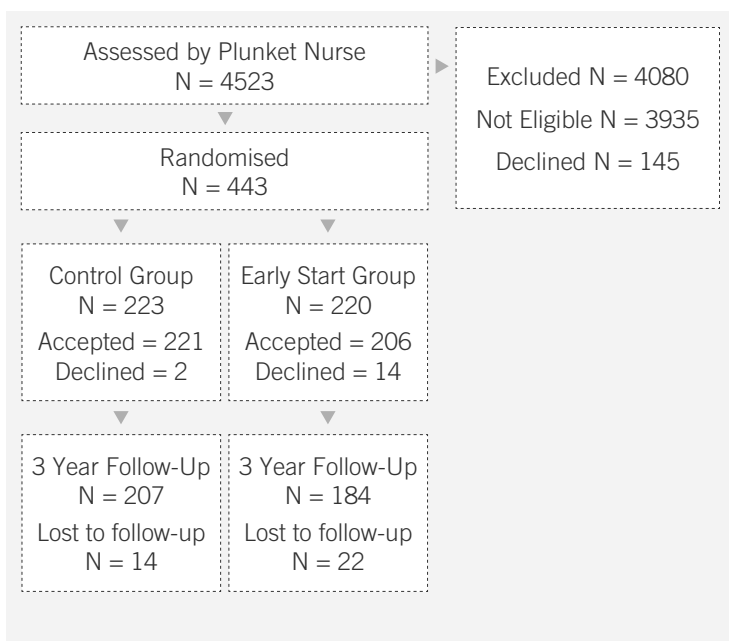
Recruitment for the randomised trial took place over a 19 month period (spanning 1999–2001). During this time Plunket nurses throughout the Christchurch area screened all families with new infants, using the screening method described in Chapter 2. Figure 3.1 provides a CONSORT diagram of the screening, recruitment and assignment process. The figure shows a total of 4,523 families were seen by Plunket nurses over the recruitment period. Of these families, 588 (13%) were deemed eligible for the trial on the basis of the screening criteria. Of those eligible, 443 (75%) agreed to participate in the trial.

Under ideal circumstances it would have been of interest to compare the characteristics of those entering and those declining to enter the trial. Such a comparison was not possible owing to the information ‘firewall’ set up between Early Start and the Plunket service. The nature of this firewall was that Plunket would not supply information to Early Start about clients who either declined or were not eligible for the service, to protect the privacy of Plunket’s clients. The net effect of this situation is that, although a substantial fraction of eligible families was recruited for the trial, there is no guarantee those families who were enrolled in the trial were representative of all clients who were eligible for the trial. It is important to recognise that any biases in the referral processes will not influence the internal validity of the comparisons between those who entered the trial and received or did not receive the Early Start service. What it will do is limit the extent to which conclusions can be drawn about the potential benefits of Early Start for all clients eligible to enter this service (as opposed to the benefits for those agreeing to enter the service).

3.2.2 Assessments

- 1) Family interviews: Clients in both groups were assessed when they enrolled in the trial and again at six, 12, 24 and 36 months post-enrolment. The assessment was done using a home based interview conducted by an interviewer (Mrs R. Deighton) employed by the CHDS. These interviews were conducted with the parent who had the greatest involvement with the child (usually the child’s mother) and lasted between 1 and 2 hours. Data collected on each family involved over 3,000 items of information describing that family over a three-year period. All interview material was quality controlled by project staff (Mr L. J. Horwood and Mrs E. Ridder) by checking completed questionnaires for clerical accuracy, consistency, and coherence in the presence of the survey interviewer. This approach provided a continuous quality control of questionnaire accuracy and content.
- 2) Information from general practitioner and hospital records: As part of the interview process, parents were asked to provide signed consent for the research group to

Figure 3.1 CONSORT diagram of client recruitment and randomisation



access general practitioner and hospital records on the health of the child enrolled in Early Start. These consents were used to access general practitioner records to verify the child's immunization status and history of well-child care. In addition, case note information was obtained for all hospital attendances/admissions.

3.3 Characteristics of children and families entering the randomised trial

Families who agreed to enter the Early Start trial were randomly assigned to either the Early Start group or the Control group. Shortly after this assignment, families in both groups were contacted by a research interviewer who conducted a baseline interview that examined a wide range of issues. (Of the Control group, 221 out of 223 families were interviewed; of the Early Start group, 206 out of 220 families were interviewed. The loss of 14 families from the Early Start group arose because families who originally entered the trial declined to continue when they were randomised to Early Start. In this instance, those withdrawing from the study also declined further interviews.)

The information gathered at the baseline interview served two functions. First, it provided an overall description of the social,

personal and related characteristics of families at the point of enrolment. Second, comparisons between the Early Start and Control groups provided a check on the extent to which randomisation to these groups led to equivalent groups of families.

3.3.1 Social and demographic background

Table 3.1 describes the Early Start and Control groups on a series of socio-demographic factors including mean parental ages, parental ethnicity, parental educational qualifications, family type, and family size. The table shows the mean age of mothers in the trial was just over 24 years; the mean age of fathers was around 27 years. The majority (>70%) of parents lacked formal educational qualifications. Between 25% and 30% of parents described themselves as Māori using questions derived from the census definition of ethnicity. The majority (>60%) of families were single parent families. Mean family size was 1.64 children. The table also shows that in all cases there were no significant differences between the Early Start and Control groups, indicating the randomisation had ensured these groups were equivalent with respect to social and demographic background.

Table 3.1 Comparison between the Early Start and Control groups on social and demographic characteristics at baseline

Measure	Controls (N = 221)	Early Start (N = 206)	p
Maternal Factors			
Mean age of mother at enrolment	24.4	24.6	.67 ¹
% Māori	26.7	24.8	.65 ²
% Lacked educational qualifications	69.9	70.6	.88 ²
Paternal Factors (Biological father)			
Mean age	26.6	27.3	.36 ¹
% Māori	25.4	30.7	.22 ²
% Lacked educational qualifications	72.3	77.8	.23 ²
Family Factors			
% Single parent family	63.8	64.6	.87 ²
Mean family size	1.6	1.6	.99 ¹

¹ t-test for independent samples

² chi squared test

Table 3.2 Comparison between the Early Start and control groups on family socio-economic characteristics at baseline

Measure	Controls (N = 221)	Early Start (N = 206)	p
% Welfare dependent	90.1	88.4	.57 ²
Mean family income (\$ per week)	\$398	\$433	.56 ¹
Mean amount of debt (excl. mortgage)	\$1,662	\$1,515	.59 ¹
% Family income inadequate/very inadequate	32.9	40.4	.11 ²

1 t-test for independent samples
2 chi squared test

Table 3.3 Comparison between the Early Start and control groups on maternal childhood disadvantage

Measure	Controls (N = 221)	Early Start (N = 206)	p ¹
% Raised in single parent family	50.2	55.3	.29
% Interparental conflict/assault	57.0	50.5	.18
% Child abuse	41.6	44.7	.53
% Impoverished family circumstances	45.7	46.1	.93
% Unhappy/very unhappy childhood	27.2	31.9	.29

1 chi squared test

3.3.2 Socio-economic background

Table 3.2 describes the socio-economic background of the study sample. It is clear from this table that those entering the trial tended to be relatively economically disadvantaged and/or impoverished. As would be expected from the high rate of single parenthood, the great majority (approximately 90%) of families were dependent on welfare benefits for the major or only source of income. This level of welfare dependence was reflected in the mean family income level of \$415 a week (assessed in 2000 and 2001).

On average, families had debts (excluding mortgages) of over \$1,500, and over a third of families described their income as inadequate or very inadequate to meet day to day living costs. There were no significant differences between the Early Start and Control groups with respect to the socio-economic factors.

3.3.3 Maternal childhood

To develop an account of the family's social circumstances, mothers were asked a series of questions about the extent to which they had been exposed to disadvantage during childhood. These results are summarised in Table 3.3, which compares the two groups on measures of maternal childhood. The table shows that women in both groups reported what appear to be relatively high levels of exposure to childhood adversity with: a) over half reporting being reared in a single parent family or witnessing interparental family violence; and b) over 40% reporting being the victim of physical or sexual abuse in childhood.

Approximately 45% of mothers said they had been reared in impoverished family circumstances and nearly 30% described their childhood as being unhappy or very unhappy. There were no significant differences between the Early Start and Control groups with respect to measures of maternal childhood.

3.3.4 Parental adjustment

Information on maternal childhood was supplemented by further measures of the social adjustment of the mother during adolescence and at the time of the interview. Data was also gathered on the characteristics of the woman's current partner, if applicable. Partners included both those who were living with the mother and non-resident partners. Table 3.4 compares the Early Start and Control groups on these measures. The table shows mothers had often experienced difficulties during adolescence: over 40% reported running away from home; over a third said they had been in trouble with the police; one in five reported problems with alcohol and over one third had used illicit drugs; one in seven had appeared in the Youth Court; and a similar proportion had

become pregnant before age 16. At the time of the interview, rates of alcohol and other drug use among mothers were low, but one in six mothers reported being depressed since the birth of the study child.

The information on current male partners suggested relatively high levels of criminality, substance use, and violence within this group. Over 50% of current male partners were described as having been in trouble with the law, between 9% and 17% had current problems with alcohol or drugs, between a quarter and a third were described as having problems with aggression, and a similar proportion had assaulted their current partner. In all cases there were no significant differences between the characteristics of the parents in the Early Start and Control groups.

Table 3.4 Comparison between the Early Start and control groups on parental adjustment

Measure	Controls	Early Start	p ¹
Maternal Adolescence	(N = 221)	(N = 206)	
% Ran away from home	49.8	42.7	.14
% In trouble with the Police	34.8	33.5	.77
% Problems with alcohol	21.3	20.4	.82
% Used illicit drugs	34.4	34.5	.99
% Appeared in Youth Court	14.0	16.0	.56
% Became pregnant before age 16	12.7	14.6	.57
Maternal Psychological Adjustment	(N = 221)	(N = 206)	
% At least weekly alcohol use	8.1	6.3	.47
% Weekly/daily cannabis use	7.2	8.3	.70
% Depression	16.7	18.9	.55
Adjustment of Current Male Partner	(N = 121)	(N = 116)	
% Alcohol problems	14.1	17.2	.50
% Cannabis and other drug problems	9.1	12.1	.46
% Aggression problems	24.8	33.6	.14
% In trouble with the law	55.4	55.2	.98
% Assaulted partner	25.6	35.0	.11

¹ chi squared test

3.3.5 Pregnancy and childbirth

Table 3.5 compares the Early Start and Control groups on a series of measures relating to past and current pregnancy and childbirth. The table shows the mean age at which women had first become pregnant was 19 years. Approximately one in seven women enrolled in the trial who had a previous child reported the child had entered foster care. In the great majority of cases (>80%) the current pregnancy was unplanned and the majority of mothers (>60%) reported smoking cigarettes during pregnancy, with just over one in six using cannabis in pregnancy. Just over a quarter of the women had been admitted to hospital during pregnancy, with one in seven children being admitted to intensive care following birth. The mean weight of infants at birth was just over 3.2 kg and almost 85% of mothers reported breast-feeding their children. As with other comparisons, there were no significant differences between the Early Start and Control groups.

3.3.6 Overall conclusions

The series of comparisons shown in Tables 3.1 to 3.5 leads to two general conclusions.

First, as a group, all the families entering the trial were subject to disadvantages in a number of areas that spanned: a) socio-economic deprivation and problems; b) adverse maternal childhood experiences; c) maternal and partner adjustment problems; and d) adverse pregnancy history features including youth at first pregnancy, high rates of unplanned pregnancy, and high rates of smoking during pregnancy.

Second, in all (40) comparisons, there were no statistically significant ($p < .05$) differences observed between the Early Start and Control groups. In two comparisons (use of cannabis during pregnancy and child admitted to intensive care), marginally significant ($p < .10$) differences were observed. These results are consistent with what would be expected from the randomisation to groups. It would be expected that, if the participants were randomised to groups, between one and two comparisons would have been significant at the .05 level and between three and four comparisons significant at the .10 level. The findings thus provide considerable reassurance the assignment to the Early Start and Control groups produced equivalent groups of families.

Table 3.5 Comparison between the Early Start and control groups on pregnancy and childbirth characteristics

Measure	Controls (N = 221)	Early Start (N = 206)	p
Previous Pregnancy			
Mean age at first ever pregnancy	19.4	19.3	.94 ¹
% Previous pregnancy, child in foster care	13.6	13.6	.99 ²
Pregnancy/Childbirth Characteristics			
% Pregnancy unplanned	82.3	80.1	.57 ²
% Smoked cigarettes during pregnancy	62.9	63.1	.96 ²
% Used cannabis during pregnancy	14.5	21.4	.06 ²
% Admitted to hospital during pregnancy	29.9	24.3	.19 ²
% Baby admitted to intensive care	16.7	11.2	.10 ²
Mean birth weight (grams)	3207	3258	.45 ¹
% Mother breast-fed child	83.7	85.4	.62 ²

1 t-test for independent samples
2 chi squared test

CHAPTER 4: SUMMARY: RESULTS OF EVALUATION AT 36 MONTHS



4.1 Introduction

This chapter provides a summary of the findings of the evaluation of Early Start at 36 months. It is based on material in three previous reports/scientific papers (Fergusson et al., 2005a and 2005b; Fergusson et al., 2006b) and addresses the following research questions:

1. To what extent did the provision of Early Start lead to increased positive outcomes for children enrolled in the programme when compared with the Control group?
2. To what extent did the provision of Early Start lead to increased positive outcomes for the parents and families enrolled in Early Start when compared with the Control group?
3. To what extent did any benefits of Early Start differ between Māori and non-Māori participants?

4.2 Assessment of child outcomes (0–3 years)

Child outcomes were assessed at baseline, and again at six months, 12 months, 24 months and 36 months from when the child was enrolled in the trial. The following methods were used.

4.2.1 Parent interviews

At baseline, six, 12, 24 and 36 months, client families were assessed on a structured interview conducted in the clients' homes by a trained survey interviewer. Interviews typically lasted between 45 minutes and an hour.

4.2.2 Medical record data

Interview data was supplemented by general practitioner information on immunization and well-child visits and by hospital record data on attendances made by the child enrolled in the trial.

These assessments were used to construct the following measures of childhood health, education, use of community services and wellbeing.

4.2.3 Health outcomes

A series of measures was used to assess the health outcomes of the trial. These measures included:

- The number of visits made by the child to the family doctor by 36 months.
- Whether the child was up to date with all immunizations at 36 months.
- Whether the child had received all well-child checks provided by the family doctor by 36 months.
- Rates of hospital attendance for accidents/injuries and accidental poisoning up to 36 months.
- Whether the child was enrolled with pre-school dental services or a dentist at 36 months.

Information on immunizations, well-child checks and injuries/poisoning was obtained from medical records (following signed parental consent), and other measures were based on parental reports.

4.2.4 Pre-school education and community service utilization

To assess the extent to which families used non-medical community services, two measures of service use were developed:

- The duration of the child's attendance at pre-school education services by 36 months.
- The number of community service agency contacts the family had made up to 36 months.

4.2.5 Parenting

At 36 months, parents were given a 49-item parenting questionnaire that contained items derived from the Child Rearing Practices Report (Block, 1981; Dekovic et al., 1991), and the Adult-Adolescent Parenting Inventory (Bavolek and Keene, 1999; Hanson, 1990). Factor analysis of this item set revealed the test items measured two general factors. These factors were:

- Positive parenting: Those scoring high on this factor tended to agree with statements suggesting they found parenting a rewarding task.
- Non-punitive parenting: Those scoring high on this factor tended to disagree with statements implying the use of physical punishment was the most effective way of managing child behaviour.



4.2.6 Child abuse and neglect

This was assessed using two measures. The first was parental reporting of severe punishment of the child by either parent, based on the severe/very severe assault subscales of the Parent-Child Conflict Tactics Scale (Straus et al., 1998) assessed at 12, 24 and 36 months. These subscales comprise eight items that measure severe punitive behaviours (e.g. 'hit him/her with a fist or kicked him/her hard', 'grabbed him/her around the neck and choked him/her'). Parents were classified as engaging in severe physical assault if they reported at least one item over the assessment period. The second was parental reporting of contact with the Child, Youth and Family (CYF) service for issues relating to child abuse and neglect.

4.2.7 Child behaviour

At 36 months, child behaviours were assessed using 50 items from the Infant Toddler Social and Emotional Assessment Scale (Briggs-Gowan and Carter, 1998). These measures spanned a series of behavioural dimensions, which were then

categorized into two overall scores:

- Externalizing behaviours: Those children scoring high on these behaviours tended to demonstrate such behaviours as (over) activity, aggression/defiance, peer aggression, and emotional negativity.
- Internalizing behaviours: Those children scoring high on these behaviours tended to show more inhibition/separation and depression/withdrawal problems.

4.2.8 Participation in service delivery and research assessments

Table 4.1 shows the number of clients enrolled in the Early Start service who were actively in receipt of services at the six, 12, 24 and 36 months follow-up. The table also shows that: a) at 36 months just under 60% of clients were receiving the Early Start service (the average duration in the programme was 24 months); b) the numbers in the Early Start and Control groups who were assessed at six, 12, 24 and 36 months; and c) at 36 months just over 90% of those enrolled in the trial were interviewed.

Table 4.1 Rates of participation in the Early Start trial and the Early Start service at enrolment and six, 12, 24 and 36 months post-enrolment

Period	Research Participation				Service Participation	
	Controls		Early Start		Active in Early Start	
	N	%	N	%	N	%
Enrolment	223	100	220	100	220	100
Baseline	221	99.1	206	93.6	206	93.6
six months	216	96.9	196	89.1	174	79.1
12 months	216	96.9	198	90.0	162	73.6
24 months	211	94.6	187	85.0	143	65.0
36 months	207	92.8	184	83.6	131	59.5

4.3 Results for child outcomes

Table 4.2 compares the Early Start and Control groups on a series of measures that summarize outcomes relating to child health and health care, pre-school education, agency use, parenting, child abuse and neglect, and behavioural outcomes. Each comparison is tested for statistical significance using t-tests, and chi square tests as applicable. The size of effect is described by the standardized difference (Cohen's d) between means or proportions. The association between treatment group and outcomes is described by the odds ratio for dichotomous outcomes and the correlation ratio (eta) for continuous outcomes.

The table shows:

- Child health: Table 4.2a shows the children in the Early Start group had greater contact with their family doctor ($p < .05$), were more up to date with well-child checks ($p < .05$), had fewer hospital attendances for accidents/injuries and accidental poisoning ($p < .05$), and had higher rates of enrolment with pre-school dental services ($p < .05$). However, those in the Early Start group did not have a significantly ($p > .80$) higher rate of immunization than those in the Control group. Effect sizes ranged from .02 to .25 (median = .22).
- Pre-school education and use of other community services: Table 4.2b shows

that those in Early Start were enrolled in early childhood education for a longer period ($p < .01$) and had higher rates of contact with community services ($p < .01$) than those in the Control group. Effect sizes ranged from .22 to .31.

- Parenting: Table 4.2c shows the parents in the Early Start group reported significantly higher positive ($p < .01$) and non-punitive parenting ($p < .05$) and had higher overall parenting scores ($p < .01$) than those in the Control group. Effect sizes ranged from .22 to .27.
- Child abuse and neglect: Table 4.2d shows the parents in the Early Start group reported a significantly lower rate of severe physical assault ($p < .01$) than the Control group parents. There were, however, no differences in rates of agency contact for child abuse and neglect. During the course of the trial, seven children were admitted to hospital for child abuse and neglect. Of these, five came from the Control group and two from the Early Start group. Effect sizes ranged from .04 to .26.
- Behavioural adjustment: Table 4.2e shows the Early Start group had a marginally significant ($p < .07$) lower rate of externalizing problems and significantly lower overall rates of internalizing problems ($p < .01$) and overall problems ($p < .05$) than the Control group. Effects sizes ranged from .19 to .26.



Table 4.2 Child outcomes of the trial

Measure	Controls (N = 207)	Early Start (N = 184)	p	Association ^a (95% CI)	Effect (d) (95% CI)
Table 4.2a Child Health					
Mean number of GP visits (0–36 months)	20.7	23.4	<.05	.11 (.01,.21)	.24 (.02,.41)
% Up to date with immunizations (0–36 months)	91.9	92.5	.83	1.09 (0.51, 2.32)	.02 (-.19,.22)
% Up to date with well-child checks (0–36 months)	30.1	41.9	<.05	1.70 (1.11, 2.59)	.25 (.06,.46)
% Attended hospital for accident/injury or accidental poisoning (0–36 months)	26.3	17.5	<.05	0.59 (0.36, 0.98)	.22 (.02,.41)
% Enrolled with dental nurse/dentist at 36 months	62.8	72.3	<.05	1.54 (1.01, 2.37)	.20 (.01,.40)
Table 4.2b Service utilization					
Mean duration of early childhood education, months (0–36 months)	13.6	16.4	<.05	.11 (.01,.21)	.22 (.02,.42)
Mean number of community service contacts (0–36 months)	7.7	8.7	<.01	.16 (.06,.26)	.31 (.13,.51)
Table 4.2c Maternal parenting attitudes^b					
Mean positive parenting attitudes (36 months)	9.88	10.14	<.01	.13 (.03,.23)	.26 (.06,.47)
Mean non-punitive attitudes (36 months)	9.90	10.12	<.05	.11 (.01,.21)	.22 (.02,.42)
Mean parenting score (36 months)	9.87	10.14	<.01	.13 (.03,.23)	.27 (.07,.47)
Table 4.2d Child abuse and neglect					
% Parental report of severe physical assault (0–36 months)	11.7	4.4	<.01	0.35 (0.15, 0.80)	.26 (.07,.48)
% In contact with agencies for child abuse or neglect (0–36 months)	21.3	19.6	.39	0.91 (0.55, 1.48)	.04 (-.15,.25)
Table 4.2e Child behavioural Adjustment^b					
Mean externalizing score (36 months)	10.09	9.90	<.07	.09 (-.01,.19)	.19 (-.01,.39)
Mean internalizing score (36 months)	10.12	9.86	<.01	.13 (.03,.23)	.26 (.06,.47)
Mean total behaviour score (36 months)	10.11	9.87	<.05	.12 (.02,.22)	.24 (.04,.44)

a Association between treatment group and outcomes is described by the odds ratio for dichotomous measures and the correlation ratio for continuous measures.

b For ease of interpretation, all parenting and child behaviour scores have been standardized to a mean of 10 and a standard deviation of 1.

4.4 Assessment of parent and family outcomes (0–3 years)

This section provides an account of the differences between the parental outcomes experienced by families enrolled in Early Start when compared with the families in the Control group. This account draws heavily on the research reported by Fergusson et al., (2006b).

4.4.1 Methods

On the basis of parental interviews conducted up to when the child was aged three years, the following measures of parental and family outcomes were constructed.

- 1) Maternal health and wellbeing: This was assessed using the following measures:
 - Contraceptive use: At each assessment, parents were asked whether they were currently using any form of contraception.
 - Subsequent pregnancy: At the 24 month follow-up, parents were asked whether they had ever become pregnant since the study child was born. At the 36 month follow-up parents were asked whether they had become pregnant since the last interview.
 - Maternal depression: At six, 12, 24 and 36 months, parents were questioned about their depressive symptoms. Items from the Composite International Diagnostic Interview (CIDI) (World Health Organization, 1993) were used to determine whether parents met the DSM-IV diagnostic criteria for major depression over the period since the previous assessment.
- 2) Maternal substance use (tobacco, alcohol and other drugs): At each assessment, parents were questioned about cigarette smoking, their use of alcohol and other drugs and their experience of problems associated with alcohol and/or drug use since the previous assessment. Parents were asked whether they smoked cigarettes and, if so, how many cigarettes they smoked each day. Questions concerning alcohol and drug related problems were based on items from the CIDI relating to DSM-IV diagnostic criteria for alcohol and drug abuse and dependence. On the basis of this information, parents were classified as having alcohol and/or substance use problems if they reported any abuse or dependence item over each 12 month follow-up period.
- 3) Family stability, family relationships and family violence: The following measures were used to assess key aspects of the child's family environment:
 - Family stability: To determine whether the child was living in a single parent family, parents were asked to describe their current living situation at each point of observation and whether a partner was present in the household. At each assessment, parents were also questioned about changes in family structure since the previous assessment, including parental separation, reconciliation, remarriage, placement with foster parents, and any other changes of parents. In addition, information was obtained on the number of changes of residence since the previous assessment.
 - Family violence: At each assessment, parents were questioned about partner violence using the revised Conflict Tactics Scale (CTS2) (Straus et al., 1996). Parents were classified as being assaulted by a partner if they reported any incident of physical assault by any partner over each 12 month follow-up period.
- 4) Family economic and material wellbeing: Family economic circumstances were assessed using the following measures:
 - Welfare dependence: At each assessment parents were asked whether either parent was currently in receipt of a social welfare benefit. Families were defined as welfare dependent if they currently relied on a social welfare benefit.
 - Family income: Family income was assessed in New Zealand dollars, net of tax, and included all income from

welfare benefits, paid employment and other sources.

- Parental workforce participation: At the 12, 24 and 36 month assessments parents were asked if they currently worked in paid employment, including any part-time work. This was recorded separately for mothers and partners.
 - Economic hardship: Economic hardship factors were recorded at each assessment from parental reporting of those factors experienced by the family since the last assessment. These economic hardship factors included 'borrowed money from family or friends,' 'unable to pay the bills,' 'unable to pay rent,' and 'postponed visits to the doctor or dentist'. A family hardship score was calculated by summing the number of hardship factors reported at each assessment period.
- 5) Family exposure to stress and adversity: At each point of assessment, using a 45-item questionnaire, respondents were asked about their exposure to stressful and adverse life events. These events were categorised into four dimensions of stress and adversity: illness and death; economic and financial problems and crises; family or social relationship problems; and victimisation. In addition, an overall score of stressful and adverse life events was calculated by summing the number of events reported by parents over each 12 month follow-up period.

4.5 Results for parent and family outcomes

Table 4.3 compares the Early Start and Control groups on a series of measures of maternal health, family functioning, family economic circumstances and exposure to adverse life events. These were assessed up to the 36 month follow-up. Each comparison is tested for statistical significance using chi square tests for dichotomous outcomes and t-tests for continuously scored measures.

The table shows a consistent lack of association between maternal and family

outcomes and group status. Specifically, there was no evidence to suggest the Early Start group experienced benefits in the areas of maternal health (Table 4.3a), family functioning (Table 4.3b), family economic functioning (Table 4.3c) or exposure to stressful life events (Table 4.3d). This lack of association is manifested in two ways. First, all comparisons fail to reach statistical significance. Second, the differences between groups show no systematic trend for one group to fare better than the other. In some comparisons (further pregnancy, welfare dependence, income, maternal employment, economic stress) families in the Early Start group fared better than families in the Control group. In others (use of contraceptives, alcohol problems, parental separation, maternal assault and victimisation events), families in the Control group fared better than families in the Early Start group. This pattern of results is clearly consistent with the conclusion the Early Start service offered families no consistent benefits in the areas of maternal health, family functioning, family economic circumstances and exposure to stress and adversity.



Table 4.3 Family and parental outcomes at three years

Measure	Controls (N = 207)	Early Start (N = 184)	p
Table 4.3a Maternal health			
% Currently using contraception at 36 months	54.1	50.0	.42
% Ever pregnant to 36 months	47.6	42.9	.35
% Major depression (past 12 months)	15.9	16.9	.81
% Mother smoked cigarettes at 36 months	62.3	62.0	.94
% Alcohol use problems (past 12 months)	9.7	14.1	.17
% Substance use problems (past 12 months)	5.8	4.9	.69
Table 4.3b Family functioning			
% Single parent family at 36 months	57.0	62.0	.32
% Any relationship separation (0–36 months)	26.3	31.1	.31
Mean number of family changes (0–36 months)	1.1	1.0	.64
% Mother physically assaulted by partner (past 12 months)	7.3	8.7	.60
Mean number of residential changes (0–36 months)	2.5	2.5	.98
Table 4.3c Family economic functioning			
% Welfare dependent at 36 months	71.5	70.1	.76
Mean family income per week at 36 months (in NZD)	492	499	.64
% Mother in paid employment at 36 months	26.6	31.5	.28
% With partner in paid employment at 36 months	30.4	27.2	.48
Mean number of hardship factors (past 12 months)	4.2	4.5	.32
Table 4.3d Exposure to stressful life events			
% Any illness/death (past 12 months)	41.6	39.7	.71
% Any economic problem (past 12 months)	80.2	74.5	.18
% Any family/social relationship problem (past 12 months)	47.3	54.9	.14
% Any victimisation (past 12 months)	5.3	10.3	.06
Mean number of stressful life events (past 12 months)	3.5	3.6	.51

All parenting and child behaviour scores have been standardized to a mean of 10 and a standard deviation of 1.

4.6 Programme outcomes for Māori and non-Māori participants

There have been on-going public policy debates about how beneficial mainstream or generic programmes are for Māori (Durie et al., 2010; Smith, 1999). For these reasons it is important to examine the extent to which the outcomes of Early Start were similar or different for Māori and non-Māori population groups.

Table 4.4 compares the Early Start and control groups on a series of outcome measures found to be significant in Table 4.2. For each comparison, the table shows separate results for Māori and non-Māori families. In constructing the table, Māori families were classified as families in which either parent figure was described as Māori using questions based on the New Zealand census questions. For each comparison the

table reports the test of the treatment group x ethnicity interaction. In testing effects for this table, continuous measures were analysed using two way analysis of variance; for dichotomous outcomes logistic regression was used. The table leads to the following conclusions.

As a general rule, in most comparisons there were no significant treatment group x ethnicity interactions, indicating that most outcome results were similar for both Māori and non-Māori. In two comparisons there was a significant treatment group x ethnicity interaction. First, for general practitioner visits the differences for Māori were smaller than for non-Māori, suggesting the programme benefits were weaker for Māori in this instance. Second, internalising behaviour differences for Māori were greater than for non-Māori suggesting that, in this instance, programme benefits were stronger for Māori.



Table 4.4 Ethnic status and child outcomes to 36 months

Measure	Māori		Non-Māori		Group x Ethnicity p
	Controls (N = 75)	Early Start (N = 76)	Controls (N= 131)	Early Start (N= 105)	
Child health					
% Up to date with well-child checks by 36 months	32.0	38.2	29.0	47.6	.23
Mean GP visits by 36 months	23.2	22.0	19.3	24.2	<.05
Hospital attendance					
% Attended hospital for accident/injury by 36 months	26.7	25.0	26.0	15.2	.25
Pre-school dental care					
% Enrolled with dental service at 36 months	60.0	68.4	64.9	75.2	.77
Early Childhood Education					
Mean duration of early childhood education (months)	12.2	15.8	14.5	16.8	.63
Parenting					
Mean positive parenting attitudes (36 months)	9.63	10.04	10.02	10.20	.25
Mean non-punitive attitudes (36 months)	9.76	10.08	9.98	10.14	.45
Mean total parenting score (36 months)	9.65	10.07	10.01	10.19	.25
Child abuse and neglect					
% Severe/very severe physical assault	12.0	2.6	11.5	5.7	.35
Child behavioural adjustment					
Mean externalising score (36 months)	10.28	9.98	9.97	9.82	.49
Mean internalising score (36 months)	10.41	9.84	9.96	9.88	<.05
Mean total overall score (36 months)	10.36	9.93	9.98	9.82	.19

On closer inspection, the table shows a general trend for the differences for Māori to be slightly larger than for non-Māori in the areas of early childhood education, parenting, child abuse and neglect and child behavioural adjustment. This suggests programme benefits for Māori tended to be as good as if not better than those for non-Māori. This conclusion is supported by a comparison of the effect sizes (Cohen's d) for Māori and non-Māori for each outcome. For Māori, the range of d values was -.10 to .55, with a median of .29; for non-Māori the range of d values was .08 to .40, with a median of .21.

Collectively these findings suggest the benefits of Early Start were similar for Māori and non-Māori but, if anything, effect sizes were slightly larger for Māori.

4.7 Summary and conclusions

This chapter has presented a summary and overview of the key findings of the assessment of Early Start at the 36 month follow-up. In terms of the three questions raised in the Introduction to this chapter, the analysis leads to the following conclusions.

4.7.1 To what extent did the provision of Early Start lead to increased positive outcomes for children enrolled in the programme when compared with the outcomes of the Control group?

The evaluation shows the Early Start programme produced benefits in a number

of areas of childhood functioning. These included:

- **Child health:** Children in the Early Start group had greater contact with family doctors, higher rates of well-child care, greater use of pre-school dental services and lower rates of childhood accidents. Each of these outcomes reflects areas targeted by the programme. While children in the Early Start group had a high rate of immunization they did not differ in this respect from the Control group. The high rates of immunization in the Control group can be explained by the fact the local general practitioner organization (Pegasus Health) instituted a campaign to encourage immunization at the same time the trial was being conducted.
- **Pre-school education and service utilisation:** Children in the Early Start group had greater involvement in pre-school education and their families made higher use of local community services.
- **Positive parenting:** Parents in the Early Start group had higher scores on measures of non-punitive and positive parenting at the 36 month assessment.
- **Child abuse and neglect:** Parents in the Early Start group reported a substantially lower rate of severe child assaults than the parents in the Control group did (4% vs. 11%). This suggests the programme was effective in reducing rates of physical child abuse. However, this trend was not reflected in rates of contact with official agencies for concerns relating to child abuse and neglect. The absence of an association with agency contact may reflect the fact the Early Start group was under regular surveillance by Family Support Workers and would be expected to have greater agency contact for abuse and neglect concerns.
- **Behavioural adjustment:** Finally, those in the Early Start group showed small reductions in rates of both externalizing and internalizing behaviours at the 36 month assessment.

A feature of the results that merits comment concerns the size of effects. These were in the small to moderate range. This tendency

for evaluations of home visiting programmes to show modest effects has been commented on in several reviews (Gomby et al., 1999; Olds and Kitzman, 1993). We believe there is a straightforward explanation for this finding. This explanation centres around two features of the randomised trial evaluation of home visits that differ from the conditions prevailing in standard clinical trials. First, those entering trials of home visits are not a homogeneous population experiencing a common set of problems. Rather, they are a heterogeneous group experiencing a wide range of issues and difficulties. Second, those participating in trials of home visiting do not receive a standard method of treatment but rather a programme of home visits designed to meet their needs. Reflection on these features suggests it would not be expected that a varying treatment applied to a heterogeneous population would produce large specific effects. What one would expect to find is that the effective programmes would show the pattern of small but pervasive benefits evident in this evaluation.

4.7.2 To what extent did the provision of Early Start lead to increased positive outcomes for the parents and families enrolled in Early Start when compared with the outcomes of the Control group?

Despite the clear benefits of Early Start for the parenting and child related outcomes described above, there was no evidence to suggest the programme had a positive impact on a wide range of family related outcomes including maternal health, family functioning, family economic circumstances and susceptibility to family stress. These findings may have some important implications for understanding the ways family support works.

There is some research that suggests targeting family risk factors for child abuse may provide an important strategy for improving the effectiveness of home visiting programmes (Chaffin, 2004; Duggan et al., 2004; Windham et al., 2004). The present results do not support these conclusions: the

results of the evaluation of Early Start showed a range of positive changes in parenting and child related outcomes in the absence of corresponding family level change. These findings support a conclusion drawn on the basis of an earlier evaluation of the Early Start service that this programme appears to work by promoting 'new learning' in areas relating to child health, education and parenting rather than by changing long standing family difficulties (Fergusson et al., 1998). What these findings may suggest is that home visiting programmes are most effective when they focus on providing parents with new skills, insights and approaches to the complex task of parenthood; they are least effective when they attempt to change long standing family problems and difficulties.

4.7.3 To what extent did the benefits of Early Start differ between Māori and non-Māori participants?

The findings of the evaluation suggested that overall there were few detectable differences in the outcomes for Māori and non-Māori families. There was a small but pervasive tendency for Māori families to receive greater benefits than non-Māori families. This was evident in a median effect size of .29 for Māori compared to a median effect size of .21 for non-Māori. These comparisons lead to the view that programme benefits were similar for Māori and non-Māori but if anything Māori received greater benefits than non-Māori.

These results are clearly relevant in the context of on-going debates about the relative contributions of generic or mainstream programmes and Māori

programmes. There have been strong claims made that generic or mainstream programmes have limited effectiveness in addressing issues for Māori and for this reason greater investments should be made in programmes owned, developed and run by Māori for Māori (Durie, 1998; Fanslow et al., 2000; Ministry of Health, 1998). The results of this evaluation of Early Start are not consistent with this view and suggest mainstream programmes may, in fact, deliver similar if not greater benefits to Māori clients.

The benefits of Early Start for Māori may reflect a number of features of the design and implementation of the Early Start programme. When the Early Start programme was set up, considerable emphasis was placed on developing an organisation, training environment and programme that was sensitive to issues relating to Māori. This was achieved by an on-going process that involved: a) initial consultation with Māori about the programme design and directions; b) the establishment of a Board on which approximately 50% of its members were Māori; c) investment in cultural training for all workers; and d) the employment of Māori staff. It would appear this combination of processes resulted in an organisational environment that produced outcomes for Māori families that were as good as, if not better than, the outcomes for non-Māori families. These results clearly suggest mainstream programmes can deliver effective outcomes for Māori providing these programmes make an adequate investment in addressing issues relating to Māori consultation, representation and service delivery.

CHAPTER 5: FINDINGS UP TO THE NINE YEAR FOLLOW-UP



5.1 Introduction

Following the initial assessment of outcomes up to three years described in Chapter 4, families in both the Control and Early Start groups were followed up again at five, six and nine years after the children were enrolled in the trial. At all of these times data was collected from parental interviews and hospital records, supplemented by teacher report information for the children at ages five, six and nine years.

The availability of this data makes it possible to examine the extent to which participation in Early Start had longer term benefits for children and families when compared with the outcomes experienced by the Control group. This chapter reports the findings from this research and addresses three general research questions:

- 1) To what extent did the provision of Early Start lead to positive outcomes for children enrolled in the programme when compared with the Control group at up to nine years follow-up?
- 2) To what extent did the provision of Early Start lead to positive outcomes for parents and families enrolled in Early Start when compared with the Control group at up to nine years follow-up?
- 3) To what extent did any benefits of Early Start to nine years follow-up vary between Māori and non-Māori families?

More generally, the aims of this chapter are to examine the extent to which the overall findings reported in Chapter 4 persisted when the samples of Early Start and Control families were studied over a nine year follow-up period.

5.2 Methods

5.2.1 Data collection

Families in the Early Start and Control groups were studied at five, six and nine years after they enrolled in the programme. The following methods of data collection were used:

- **Parental interviews:** At each point of study, parents were given a structured interview conducted by a trained survey interviewer, lasting approximately 1 hour. The majority of interviews were done in the respondent's home, but telephone interviews were conducted for those living outside the Canterbury region. Interviews spanned a number of domains of childhood wellbeing and family functioning.
- **Hospital records:** At each point of study, parents were asked to give signed consent for the research group to access the hospital records for the child enrolled in the study. Over 99% of parents provided this consent.
- **Teacher questionnaires:** At the five, six and nine year follow-up, the child's class teacher was asked to respond to a questionnaire describing the child's behaviour and progress at school. This data collection was subject to signed consent from the child's parents. Over 99% of parents provided this consent.

All aspects of data collection were subject to the review and approval of the Upper South Ethics Committees administered by the Ministry of Health.

5.2.2 Sample retention

Table 5.1 shows the number of sample members studied at baseline, three years, six years and nine years and the percentage of the original samples assigned to the Early Start and Control groups. The table shows that overall there was a good rate of sample retention, with 370 of the 443 families (83%) entering the trial being followed up over a 9 year period. There is also clear evidence sample retention for the Control group was better than that for the Early Start group. At nine years, 89% of the Control group was studied, whereas 78% of the Early Start group was studied. This difference in retention rates is highly statistically significant ($p < .001$).

The reasons for the differential sample loss are not entirely clear. Notably, 14 (29%) of the 49 Early Start families lost to follow-up had declined before the trial began. Further, there was a small tendency for higher rates of dropout in the Early Start group over the nine year follow-up. Collectively, the 14 families declining to participate in Early Start coupled with the higher refusal rate for the Early Start group meant the response rate for the Control group was noticeably and significantly higher than that for the Early Start group. We conjecture that these differences in study participation are largely a consequence of the dual burden of service participation and research interviewing imposed on the Early Start families. This respondent burden was far greater for the Early Start families than for the Control group families, and contributed to the

Table 5.1 Sample retention at baseline and at three, six, nine years follow-up

Assessment Period	Controls (N = 223)		Early Start (N = 220)		Total (N = 443)	
	N	%	N	%	N	%
Baseline	221	99.1	206	93.6	427	96.3
3 Years	207	92.8	184	83.6	391	88.3
6 Years	205	91.9	180	81.8	385	86.9
9 Years	199	89.2	171	77.7	370	83.5

decrease in the number of Early Start families studied to the nine year follow-up.

Irrespective of the reasons for the differential sample losses, they pose complications for the analysis of the nine year follow-up. It could be suggested that differential sample loss in the Early Start and Control groups poses a threat to study validity. Depending on the nature of these losses they could act to obscure or to inflate differences between the groups (Hernán et al., 2004; Schulz and Grimes, 2002).

There is no completely satisfactory solution to the problems of differential sample loss in randomised trials (Schulz and Grimes, 2002). To address the issues raised by differential sample loss, the following approach was used:

- In the main analyses reported in this chapter, the samples sizes identified in Table 5.1 were used to compare the Early Start and Control groups.
- To examine the effects of differential sample loss, a series of supplementary analyses was conducted. These analyses included: a) comparisons of the family, social, and behavioural characteristics of those studied to the nine year follow-up, to examine the equivalence of the studied samples; and b) the use of sample selection correction methods to take account of differential sample loss. The aim of these supplementary analyses was to examine the extent to which differential sample loss was likely to have posed a threat to study validity.

5.2.3 Statistical methods

The data reported in this chapter examines the outcomes of the Early Start and Control groups up to the nine year follow-up. To test for differences between the Early Start and Control groups, a unified analysis methodology was used by fitting a generalised estimating equation model to each data table having repeated measures. This statistical method is described in the Appendix. The analysis approach estimated: a) the effects of treatment (experimental vs. control) on the outcome measure; b) the effects of time of measurement on the outcome measure; and c) the treatment by

time interaction. The treatment by time interaction tests whether the effects of the treatment varied with the time of measurement.

5.3 Outcomes for children 0–9 years

5.3.1 Hospital attendance for non-intentional injuries (accidents)

Table 5.2 shows the percentages of children attending hospital for non-intentional injuries (accidents) over the follow-up periods 0–3 years, 4–6 years and 7–9 years. The data in the table was obtained from a review of the hospital records of the study children, subject to the signed consent of their parents. Over 99% of parents participating in the study agreed to allow access to their children’s hospital records. The table shows that, at all ages, children in the Control group had higher rates of accidents than those in the Early Start group. These differences were most marked over the 0–3 years follow-up. By the nine year follow-up, an estimated 42% of children in the Control group had attended hospital on at least one occasion for a childhood accident compared to 28% of those provided with the Early Start service. The results of statistical modelling (see Appendix) showed there was a statistically significant difference between the rates of hospital admission for accidents ($p < .05$). Cohen’s d was .29, suggesting the provision of Early Start had moderate effects on the risks of hospital admission for childhood accidents which were most evident from 0–3 years but which persisted for up to nine years.

There was a significant main effect for time of observation ($p < .0001$), reflecting the clear tendency for rates of accidents to decline with increasing age. However, there was no significant treatment by time interaction ($p > .20$), implying the benefits of Early Start were present throughout the observation period.

Table 5.2 Hospital attendance for accidental injury (%), 0–9 years

Assessment period	Controls		Early Start	
	%	N	%	N
0–3 years	26.3	216	17.5	197
4–6 years	14.4	211	11.2	183
7–9 years	8.0	199	7.6	171
Ever by 9 years	42.1		28.3	

Main effect for treatment ($p < .05$); main effect for time ($p < .0001$); treatment x time interaction effect ($p > .20$); Cohen's $d = .29$ (95% CI: .09-.49).

5.3.2 Parent reports of physical child abuse and punishment

Parents were interviewed on the Parent-Child Conflict Tactics Scale (PC-CTS) (Straus et al., 1998) at each assessment from 1–6 years follow-up and again at nine years follow-up. The PC-CTS provides measures of the extent to which the parent reports using a series of disciplinary methods including physical punishments, psychological threats, and rewards.

Table 5.3 shows the percentages of the samples who reported the use of harsh or abusive methods of punishment when the study children were aged 1–3, 4–6 and nine years, and the pooled population-averaged estimate for the period 0–9 years. For the purposes of this table, the definition of harsh or abusive punishment was the reporting of at least one of a series of methods of punishment within each measurement period. The methods of punishment included in this definition were: shaking; hitting with an object; hitting with a fist or kicking; grabbing by the neck/choking; hitting as hard as possible; burning/scalding; throwing/knocking down; and slapping on the head.

The table shows that, at all assessments, children in the Control group had higher rates of parental reported harsh punishment/abuse than those in the Early Start group. These differences were most evident over the 0–3 years follow-up. The net result of these trends was that, by nine years, approximately 20% of those in the Control group were reported to have been exposed to harsh/abusive punishment compared with approximately 10% of those in the Early Start group. The results of statistical analysis (see Appendix

showed this difference was statistically significant ($p < .01$). Cohen's d was .29, suggesting the provision of Early Start had moderate effects on the risks of severe/very severe physical assault which were most evident from 0–3 years but which persisted for up to nine years.

There was a significant main effect for time of observation ($p < .05$), reflecting the clear tendency for rates of severe/very severe physical assault to decline with increasing age. However, there was no significant treatment by time interaction ($p > .70$), implying the benefits of Early Start were present throughout the observation period.

The analysis in Table 5.3 was extended to consider the use of all forms of physical punishment by Early Start and Control group parents. This was done by computing a diversity score measure of the reported use of physical punishment. This score was constructed by summing the number of different types of physical punishment methods reported by the parents over the periods 0–3 years, 4–6 years and at the nine year follow-up. The methods of physical punishment considered spanned: smacking the child on the bottom; slapping the child on a limb; and the harsh/abusive punishment methods noted in Table 5.3 above.

Table 5.4 shows the mean numbers of physical punishment methods used by Early Start and Control group parents. The table shows that, at all ages, parents in the Early Start group reported less use of physical punishment than parents in the Control group ($p < .05$).

Although parents in the Early Start group reported significantly lower rates of punishment than the Control group parents, the effect size ($d = .13$) for this comparison was smaller than for the measure of harsh/abusive punishment reported in Table 5.3 ($d = .29$). The reason for this is that, while parents in the Early Start group reported avoiding the use of severe methods of physical punishment, their use of other methods of physical punishment remained relatively high. For example, by the nine year follow-up, 91.4% of Early Start families reported using physical punishment methods on at least one occasion compared with 92%

of the Control group. These findings suggest Early Start reduced rates of harsh and abusive punishment for families participating in the programme, although it did not eliminate the use of physical punishment for the great majority of families.

There was a significant main effect for time of observation ($p < .0001$), reflecting the clear tendency for self reported physical punishment scores to decline with increasing age. However, there was no significant treatment by time interaction ($p > .80$), implying the benefits of Early Start were present throughout the observation period.

Table 5.3 Severe/very severe physical assault by any parent (%), 0–9 years

Assessment period	Controls		Early Start	
	%	N	%	N
1–3 years	11.1	216	4.1	197
4–6 years	8.1	211	4.9	183
9 years	5.0	199	1.8	171
Ever by 9 years	20.1		9.8	

Main effect for treatment ($p < .01$); main effect for time ($p < .05$); treatment x time interaction effect ($p > .70$); Cohen's $d = .29$ (95% CI: .09-.49).

Table 5.4 Mean (SD) physical punishment scores, 0–9 years

Assessment period	Controls		Early Start	
	Mean (SD)	N	Mean (SD)	N
1–3 years	1.85 (1.06)	216	1.66 (1.06)	197
4–6 years	1.85 (1.14)	211	1.68 (1.06)	183
9 years	0.56 (1.06)	199	0.43 (0.73)	171
Overall mean (SD)	1.44 (1.30)		1.29 (1.13)	

Main effect for treatment ($p < .05$); main effect for time ($p < .0001$); treatment x time interaction effect ($p > .80$); Cohen's $d = .13$ (95% CI: .01-.24).

5.3.3 Parenting competence

The analysis of physical punishment and child abuse in section 5.3.2 was supplemented by a more general analysis of parenting using a measure of parental competence based on The Parenting Scale (Arnold et al., 1993). This measure was a 30-item scale spanning a number of domains of parenting behaviours, including laxness of discipline, parental overreactions and parental verbosity. Items were reverse-scored so that higher scores reflected more positive parenting practices. Factor analysis of the present data showed these items formed a uni-dimensional factor reflecting the extent of parenting competence, as suggested by the authors of the scale (Arnold et al., 1993). The scale's reliability (Cronbach's alpha) ranged from alpha = .82 to .83. Table 5.5 shows mean scores for Early Start and Control parents when the children were aged five, six and nine years. For ease of interpretation, The Parenting Scale has been scored to have a mean of 10 and a standard deviation of 1.

The table shows that, at all assessment points, parents in the Early Start group had higher mean parental competence scores (implying greater parental competence) than those in the Control group ($p < .05$). The overall effect size was .25, implying the provision of Early Start had moderate effects on overall parenting competence, with these benefits lasting up to the age of nine years.

There were no significant main effects for time ($p > .90$), or for treatment by time interaction ($p > .30$), implying the benefits of Early Start remained constant throughout the observation period.

5.3.4 Contact with services for child abuse/care and protection issues

Table 5.6 compares the Early Start group with the Control group on two measures of contact with services at any point to age 9 for issues relating to child abuse or care and protection.

These measures were:

- The number of families reporting they had been in contact with agencies regarding physical child abuse. This was based on questionnaire items in which families were asked about contact with a range of services because of physical child abuse.
- The number of families reporting contact with the Child, Youth and Family (CYF) service because of concerns about the treatment or wellbeing of their children. This was based on questionnaire items in which families were asked about contacts with CYF.

The results in the table show an inconsistent pattern of findings:

- Parental reports of agency contact for physical child abuse showed families in the Control group reported nearly twice the rate of such contact as those from the Early Start group. This difference is statistically significant ($p < .05$; $d = .20$), suggesting that, consistent with the results in sections 5.3.2 and 5.3.3, children from families provided with Early Start had less exposure to physical child abuse.
- A comparison of rates of reported CYF contact showed these rates were not significantly different for the Early Start

Table 5.5 Mean (SD) parenting competence scale scores, ages five, six and nine years

Assessment period	Controls		Early Start	
	Mean (SD)	N	Mean (SD)	N
5 years	9.86 (1.02)	207	10.16 (0.96)	180
6 years	9.89 (1.08)	205	10.12 (0.88)	180
9 years	9.90 (1.05)	199	10.11 (0.93)	171
Population-averaged mean (SD)	9.88 (1.05)		10.13 (0.92)	

Main effect for treatment ($p < .01$); main effect for time ($p > .90$); treatment x time interaction effect ($p > .30$); Cohen's $d = .25$ (95% CI: .13-.37)

Table 5.6 Contact with services for child abuse/care and protection issues

Outcome and age	Controls		Early Start		p ¹	Effect size (d) and 95% CI
	%	n	%	n		
% Agency contact for child abuse/neglect to age 9	6.3	13	2.2	4	<.05	.20 (.00-.40)
% CYF contact to age 9	30.9	63	27.2	49	>.20	.08 (-.12-.29)

1 Chi-square test of independence

group and the Control group ($p > .20$; $d = .08$). However, there was a small tendency for those in the Early Start group to have fewer CYF contacts.

The reasons for the inconsistencies between the measures of hospital admission and agency contact for physical child abuse on the one hand, and the rates of CYF contact and referral on the other, are not clear. We have examined the records of those reporting contacts and can find no clear explanation of the differences.

Finally, data was gathered on hospital attendances for child abuse and neglect. This produced results which were consistent with the findings suggesting the Early Start group had lower rates of exposure to child abuse. Of the seven children admitted to hospital for child abuse and neglect, five came from the Control group. Because of the small numbers involved, statistical testing of this difference is not justified.

5.3.5 Parental reports of child behaviour, five, six and nine year follow-up

At the five, six and nine year follow-up, parents were questioned about their child's behaviour using the Strengths and Difficulties Questionnaire (SDQ) (Goodman, 1997). The SDQ is a 30-item parental report questionnaire designed to assess a number of child behaviour domains including:

- Externalizing behaviours (conduct problems and hyperactivity/inattention): Children scoring high on this dimension are characterised by elevated rates of oppositional,

aggressive and anti-social behaviours; and by high rates of restless, inattentive and hyperactive behaviours.

- Internalizing behaviours (emotionality and peer difficulties): Children scoring high on this dimension are characterised by high rates of anxious, withdrawn and depressive behaviours; and by difficulties relating to peers.
- In addition, an overall score representing total difficulties can be computed by summing the externalizing and internalizing behaviour subscales. Coefficients alpha for parent reports ranged from .70 to .85.

Table 5.7 shows mean scores on each of the two behaviour domains, and the overall score at ages five, six and nine years follow-up. For ease of comparison, all measures have been scaled to an overall mean of 10 and standard deviation of 1. The table shows that, at all assessment points, parents in the Early Start group reported lower SDQ externalizing, internalizing, and total scores than those in the Control group ($p < .05$). The differences between the Early Start group and the Control group had overall effect sizes ranging from .15 to .17, indicating a modest effect of treatment on parent reported SDQ externalizing, internalizing, and total scores.

There were no significant main effects for time (all p values $> .80$), or for treatment by time interaction (all p values $> .20$), implying the benefits of Early Start remained constant throughout the observation period.

Table 5.7 Mean (SD) parent reported SDQ externalizing, internalizing, and total scores, ages five, six and nine years

Outcome/assessment period	Controls		Early Start	
	Mean (SD)	N	Mean (SD)	N
Externalizing five years	10.08 (1.03)	207	9.91 (0.95)	180
Externalizing six years	10.09 (1.05)	205	9.89 (0.93)	180
Externalizing nine years	10.02 (1.05)	199	9.97 (0.94)	171
Population-averaged mean (SD)	10.07 (1.04)	--	9.92 (0.94)	--
Main effect for treatment ($p < .05$); main effect for time ($p > .90$); treatment x time interaction effect ($p > .20$); Cohen's $d = .15$ (95% CI: .03-.27)				
Internalizing five years	10.07 (1.05)	207	9.92 (0.93)	180
Internalizing six years	10.07 (1.07)	205	9.92 (0.91)	180
Internalizing nine years	10.07 (1.09)	199	9.92 (0.89)	171
Population-averaged mean (SD)	10.07 (1.07)	--	9.92 (0.91)	--
Main effect for treatment ($p < .05$); main effect for time ($p > .80$); treatment x time interaction effect ($p > .90$); Cohen's $d = .15$ (95% CI: .03-.27)				
Total score five years	10.09 (1.04)	207	9.90 (0.94)	180
Total score six years	10.10 (1.08)	205	9.89 (0.90)	180
Total score nine years	10.05 (1.07)	199	9.94 (0.91)	171
Population-averaged mean (SD)	10.08 (1.06)	--	9.91 (0.91)	--
Main effect for treatment ($p < .05$); main effect for time ($p > .90$); treatment x time interaction effect ($p > .30$); Cohen's $d = .17$ (95% CI: .06-.29)				

Table 5.8 Mean (SD) teacher reported SDQ externalizing, internalizing, and total scores, ages five, six and nine years

Outcome/age	Controls		Early Start	
	Mean (SD)	N	Mean (SD)	N
Externalizing five years	10.01 (1.06)	207	9.99 (0.92)	180
Externalizing six years	10.02 (1.01)	205	9.97 (0.99)	180
Externalizing nine years	9.98 (1.01)	199	10.02 (0.99)	171
Population-averaged mean (SD)	10.01 (1.03)	--	9.99 (0.97)	--
Main effect for treatment ($p > .70$); main effect for time ($p > .90$); treatment x time interaction effect ($p > .90$); Cohen's $d = .02$ (95% CI: -.10-.14)				
Internalizing five years	10.00 (1.00)	207	10.00 (1.00)	180
Internalizing six years	9.98 (1.01)	205	10.03 (0.99)	180
Internalizing nine years	9.86 (0.97)	199	10.16 (1.02)	171
Population-averaged mean (SD)	9.95 (1.00)	--	10.06 (1.00)	--
Main effect for treatment ($p > .10$); main effect for time ($p > .90$); treatment x time interaction effect ($p > .10$); Cohen's $d = .11$ (95% CI: .01-.23)				
Total score five years	10.01 (1.06)	207	9.99 (0.93)	180
Total score six years	10.01 (1.04)	205	9.99 (0.96)	180
Total score nine years	9.92 (0.99)	199	10.09 (1.00)	171
Population-averaged mean (SD)	9.98 (1.03)	--	10.03 (0.97)	--
Main effect for treatment ($p > .70$); main effect for time ($p > .80$); treatment x time interaction effect ($p > .10$); Cohen's $d = .05$ (95% CI: .07-.17)				

5.3.6 Teacher reports of child behaviour, five, six and nine year follow-up

Parallel to parental reports of child behaviour, reports of child behaviour at school were obtained by asking the child's classroom teacher to complete a classroom based assessment of the child based on the SDQ. Coefficients alpha for teacher reports ranged from .76 to .89.

Table 5.8 shows mean scores on teacher reported externalizing behaviours, internalizing behaviours, and total SDQ score at five, six and nine years follow-up. The results of GEE modelling show that:

- For both externalizing and internalizing behaviours, and for the overall score, children from families provided with Early Start did not differ significantly (all p-values >.10) from those in the Control group.
- There was no evidence of statistically significant effects of time (all p values >.80) and treatment x time interactions (all p values >.10), implying a lack of association between treatment and outcomes at all points of observation.

The differences in the trial outcomes for parental reported and teacher reported behaviours raise some important and perplexing questions about the reasons for these differences. It should be noted that these findings are not entirely unexpected. There has been a large literature showing parent and teacher reports of child behaviour are only modestly correlated ($r = .33$ approximately) (Achenbach et al., 1987; McCombs Thomas et al., 1990; Spiker et al., 1992; Touliatos and Lindholm, 1981). In the earlier literature on this topic, the weak correlation between parent and teacher reports was interpreted as evidence of the lack of validity or reliability of these reports. Other research has suggested this lack of correlation may be due to children behaving differently in different social contexts, with the result that parent and teacher reports of behaviour do not show strong agreement (Fergusson et al., 2009; Fergusson and Horwood, 1989). The most plausible interpretation of the study findings is that:

- Parent reports of improvements in child behaviour are a result of the Early Start programme teaching parents how to manage child behaviours within the home context.
- However, the effect of this training does not generalise to the school context, with the result that Early Start appears to have little or no effect on behavioural adjustment at school.

5.4 Parent and family outcomes to nine years follow-up

Parallel to the analysis of child outcomes given in sections 5.3.1 to 5.3.6, an analysis was conducted of a wide range of parental and family outcomes. This analysis compared measures from the following domains of parental and family functioning obtained at the five, six and nine year follow-up:

- Maternal depression
- Parental substance use
- Family violence
- Family economic circumstances
- Family life stress.

In all, a total of 10 comparisons were made between the outcomes of the Early Start and Control families. No comparison was found to be statistically significant, and only one comparison was found to be marginally significant ($p < .10$). These findings strongly suggest the provision of Early Start had no material effect on either parental or family functioning up to the age of nine years. This is consistent with the findings for the 3 year follow-up reported in Chapter 4.

Table 5.9 shows the findings for each of the five domains of parent/family functioning. The table provides: a) a description of the outcome measures considered; b) the mean scores or percentages for both the Early Start and Control groups for each outcome pooled over the assessments at five, six and nine years; c) the overall statistical significance of the comparison; and d) the values of Cohen's d . The table shows:

- 1) Maternal depression: This was assessed at the five, six and nine year follow-up using questions from the CIDI (World Health Organization, 1993) measuring DSM-IV

symptoms of major depression. The measure of depressive symptomatology consisted of the total number of symptoms reported by participants at each assessment point. At all ages there were no significant differences between the Early Start and Control groups. A repeated measures GEE model applied to the data on maternal depression gave an overall value of $p > .20$ and a value of $d = .05$. Both estimates imply an absence of association between group membership (Early Start vs. Control) and rates of maternal depressive symptoms.

- 2) Parental substance use: Parental substance use was assessed via three dichotomous measures: a) parental cigarette smoking at the five, six and nine year follow-up; b) parental alcohol problems in the 12 months before the five, six and nine year follow-up; and c) parental use of cannabis or other illicit drugs at the five, six and nine year follow-up. Parental cigarette smoking was assessed via two items questioning whether the child's mother or her partner smoked cigarettes currently. Parental alcohol problems were assessed using items from the CIDI relating to DSM-IV symptoms of alcohol abuse/dependence. Parents were classified as having alcohol problems if they reported any abuse or dependence item in the 12 month period. Parental cannabis use/other illicit drug use was assessed via custom-written survey items asking whether the mother or her partner had used cannabis or any other illicit drug in the 12 months before the interview. A repeated measures GEE model applied to the data on substance use showed there were no significant

differences between the Early Start group and the Control group for parental smoking and alcohol problems ($p > .40$ and $.60$, respectively). However, the analyses suggested the parents in the Early Start group were marginally significantly ($p < .10$) more likely to report having used cannabis or another illicit substance in the 12 months before the assessment. Overall, there was little evidence to suggest there were systematic differences between families in the Early Start and Control groups with respect to substance use issues.

- 3) Family violence: Intimate partner violence (IPV) was assessed for the 12 months before the five, six and nine year follow-up via questions from the Revised Conflict Tactics Scale (Straus et al., 1996). The questions were asked of the child's primary caregiver from the perspective of both the victim and the perpetrator. The selected items spanned the domains of minor psychological aggression, severe psychological aggression, minor physical assault, and severe physical assault. The items from each domain were scored as present/absent and summed to create a measure reflecting the extent of victimisation and perpetration of aggression in the context of an intimate partner relationship. A repeated measures GEE model fitted to the data showed no significant differences between the Early Start and Control groups on either IPV victimisation or perpetration ($p > .60$ and $.80$ respectively), suggesting the absence of an association between group membership and IPV.
- 4) Family economic circumstances: Family economic circumstances at the five, six and nine year follow-up were assessed via three measures: a) a dichotomous measure of whether either parent was in receipt of any welfare benefit at each point of assessment; b) the total amount of debt reported by families at the time of each assessment; and c) the number of hardship items endorsed by parents based on items derived from the Economic Living Standards Index (Jensen et al., 2005). Repeated measures GEE models applied to data for each of the three outcomes showed no significant differences



between the Early Start and Control groups on any of the three measures of family economic circumstances ($p > .50$, $.20$ and $.80$, respectively). These findings suggest there was no association between group membership and family economic circumstances.

- 5) Adverse life events: A measure of the family's exposure to adverse life events obtained at the five, six and nine year follow-up was adapted from the Holmes and Rahe Social Readjustment Rating Scale (Holmes and Rahe, 1967). The scale consisted of 47 items representing a range of possible stressful or significant life events. Participants were asked to rate the extent to which they had been affected by each event on a five-point scale ranging from 'no event' to 'very upset'. These scores were then summed to create an overall measure of stressful life events exposure since the previous assessment. Scores on the measure were scaled to a mean of 10 and a standard deviation of 1. Repeated measures GEE models applied to the data showed no significant differences between the Early Start and Control groups ($p > .40$), suggesting no association between group membership and exposure to adverse life events.

5.5 Benefits for Māori and non-Māori

An important question raised by the preceding analysis concerns the extent to

which the benefits of Early Start were similar for the children of Māori and non-Māori families. As shown in the analysis of the 3 year follow-up, there was no evidence to suggest the benefits for Māori were different from those for non-Māori. This analysis was extended to the nine year follow-up by testing each comparison for ethnicity (Māori/non-Māori) by group (Early Start/Control) interactions. Evidence of a significant ethnicity x group interaction is indicative of ethnic differences in study outcomes.

The results of this analysis are shown in Table 5.10 which reports for each comparison in the previous sections (5.3.1 to 5.3.5): a) the rates or mean scores for each outcome for Māori and non-Māori participants in the Early Start and Control groups; and b) the test of the ethnicity by group interaction. Table 5.10 shows:

- In some cases (accidental injury; severe physical assault; physical punishment) the differences between the Early Start and Control groups were somewhat smaller for Māori than for non-Māori participants. However, for parenting competence scores and parent reported behavioural adjustment, the differences between the Early Start and Control groups were larger for Māori than for non-Māori participants.
- In all cases there was no evidence of a significant ethnicity x group interaction.
- These findings are consistent with the view that the benefits of Early Start were similar for both Māori and non-Māori participants.

Table 5.9 Parent and family outcomes from 5 year to nine year follow-up

Outcome	Controls	Early Start	p
Mean (SD) depression symptoms	1.39 (2.97)	1.55 (3.12)	>.20
% Reporting cigarette smoking	63.3	65.9	>.40
% Reporting alcohol problems	7.7	8.5	>.60
% Used cannabis/other illicit drugs	18.5	24.1	<.10
Mean (SD) intimate partner violence (IPV) victimisation score	9.98 (0.84)	10.02 (1.15)	>.60
Mean (SD) IPV perpetration score	9.99 (0.96)	10.01 (1.04)	>.80
% Welfare dependent	56.8	59.5	>.50
Mean (SD) debt (in NZD)	4,492 (8,524)	5,248 (9,140)	>.20
Mean (SD) number of hardship factors	3.46 (3.56)	3.52 (3.44)	>.80
Mean (SD) adverse life events score	9.97 (0.98)	10.03 (1.02)	>.40

Table 5.10 Māori ethnicity and outcomes to nine year follow-up

Outcome	Māori		Non-Māori		Test of ethnicity x group interaction	
	Controls (n = 81)	Early Start (n = 78)	Controls (n = 137)	Early Start (n = 120)	2 (1 df)	p
% Hospital attendance for accidental injury	46.6	36.5	39.6	23.2	0.96	>.30
% Severe physical assault	17.3	12.8	21.2	7.5	1.81	>.10
Mean (SD) physical punishment scale scores	1.54 (1.54)	1.53 (1.28)	1.65 (1.54)	1.38 (1.34)	1.22	>.20
Mean (SD) parenting competence scale scores	9.65 (1.14)	10.04 (0.96)	10.02 (0.97)	10.19 (0.89)	1.31	>.20
Mean (SD) parent reported SDQ total scores	10.26 (1.15)	9.94 (0.96)	9.97 (0.99)	9.89 (0.89)	1.63	>.20

There are two important caveats on these conclusions. The first is that the sample of Māori families studied is relatively small (159). It is possible that, if a larger sample of Māori families had been studied, Māori/non-Māori differences in the benefits of Early Start may have been found. The second caveat is the Māori/non-Māori differences have been examined using quantitative methods of statistical analysis rather than through a Kaupapa Māori methodology. It is also possible that research from a Te Ao Māori perspective could reach different conclusions from those drawn above.

In short, the evidence suggests Early Start had similar benefits for Māori and non-Māori participants. However, this conclusion needs further confirmation using larger samples of Māori families and a consideration of the benefits of Early Start from a Kaupapa Māori perspective.

5.6 Effects of sample attrition on study validity

A potential problem with the study design concerns the higher rates of sample attrition among the Early Start group when compared to the Control group; at nine years, 78% of the Early Start group was studied, compared to 89% of the Control

group. This differential sample attrition was largely due to 14 families assigned to the Early Start group withdrawing at the beginning of the trial. Under a strict intention to treat design, these families were counted as part of the Early Start group. In addition, rates of sample attrition were slightly higher among the Early Start group. The net effects of these trends was that, by the age of nine years, 22% of the Early Start group was lost to follow-up, compared to 11% of the Control group.

The differential sample loss poses a threat to study validity. It may be suggested that, had all the families in both the Early Start and Control groups been studied over the 9 year period, the study results may have differed from those found in this report. To address the issues raised by sample attrition, two approaches to examining the possibility of sample selection bias were used.

5.6.1 Comparison of the equivalence of the Early Start and Control groups at nine years

One way of examining the extent to which sample losses may have biased the results is to compare the characteristics of the

samples of Early Start and Control families studied at the nine year follow-up to the baseline measures. If sample biases were present, one might expect these to be evident in the profile of social, demographic and related characteristics observed at the baseline measurement of those studied at the nine year follow-up.

Tables 5.11 to 5.15 compare the 171 Early Start families studied at the nine year follow-up with the 199 Control families studied at the nine year follow-up on a series of 35 measures assessed at baseline. These measures included:

maternal and paternal social and demographic characteristics; socio-economic characteristics; maternal childhood disadvantage; maternal adjustment; and pregnancy and childbirth characteristics.

Table 5.11 compares the two groups on a series of measures of social and demographic factors assessed at baseline. These factors include: maternal and paternal age; ethnicity; educational qualifications; single parenthood; and family size. The table shows there were no significant differences between the two groups on these measures.

Table 5.11 Comparison between the Early Start and Control groups observed at nine year follow-up on social and demographic characteristics at baseline

Measure	Controls (N = 199)	Early Start (N = 171)	p ¹
Maternal factors			
Mean age of mother at enrolment	24.4	24.7	>.60
% Māori	25.1	23.2	>.60
% Lacked educational qualifications	66.8	66.1	>.80
Paternal factors (biological father)			
Mean age	26.6	27.1	>.50
% Māori	24.6	29.2	>.30
% Lacked educational qualifications	69.0	75.8	>.10
Family factors			
% Single parent family	63.3	64.3	>.80
Mean family size	1.6	1.5	>.70

1 t-test for independent samples for comparisons of means, Chi square test of independence for comparison of proportions

Table 5.12 compares the Early Start and Control families studied at the nine year follow-up on a series of measures of socio-economic circumstances, including: welfare dependence; mean family income; debt; and income adequacy assessed at baseline. In all cases there were no significant differences between the two groups.

Table 5.13 compares the Early Start and Control families studied at the nine year follow-up on a series of measures of maternal childhood disadvantage, including: being raised in a single parent family; being exposed to interparental conflict or assault; being exposed to child abuse; living in impoverished circumstances; and having an unhappy childhood. In all cases there were no significant differences between the groups.

Table 5.12 Comparison between the Early Start and Control groups observed at nine year follow-up on family socio-economic characteristics at baseline

Measure	Controls (N = 199)	Early Start (N = 171)	p ¹
% Welfare dependent	89.9	87.5	>.40
Mean family income (\$ per week)	\$341	\$343	>.80
Mean amount of debt (excl. mortgage)	\$1,659	\$1,520	>.60
% Family income inadequate/very inadequate	31.1	39.3	>.10

1 t-test for independent samples for comparisons of means, Chi square test of independence for comparison of proportions

Table 5.13 Comparison between the Early Start and Control groups observed at nine year follow-up on maternal childhood disadvantage

Measure	Controls (N = 199)	Early Start (N = 171)	p ¹
% Raised in single parent family	51.2	54.8	>.50
% Interparental conflict/assault	55.8	54.8	>.80
% Child abuse	36.2	38.1	>.70
% Impoverished family circumstances	45.2	48.8	>.40
% Unhappy/very unhappy childhood	26.1	32.7	>.10

1 Chi square test of independence

Table 5.14 Comparison between the Early Start and Control groups observed at nine year follow-up on maternal adjustment

Measure	Controls (N = 199)	Early Start (N = 171)	p ¹
Maternal adolescence			
% Ran away from home	49.2	41.7	>.10
% In trouble with the Police	33.7	34.5	>.80
% Problems with alcohol	22.1	20.2	>.60
% Used illicit drugs	34.2	33.7	>.70
% Appeared in Youth Court	12.1	16.1	>.20
% Became pregnant before age 16	12.1	13.7	>.60
Maternal psychological adjustment at baseline			
% At least weekly alcohol use	8.0	6.0	>.40
% Weekly/daily cannabis use	7.0	8.9	>.50
% Depression	21.6	16.1	>.10

1 t-test for independent samples

Table 5.15 Comparison between the Early Start and Control groups observed at nine year follow-up on pregnancy and childbirth characteristics

Measure	Controls (N = 199)	Early Start (N = 171)	p ¹
Previous pregnancy			
Mean age at first ever pregnancy	19.3	19.4	>.80
% Previous pregnancy, child in foster care	14.6	14.3	>.90
Pregnancy/childbirth characteristics			
% Pregnancy unplanned	81.4	82.1	>.80
% Smoked cigarettes during pregnancy	62.8	64.9	>.60
% Used cannabis during pregnancy	13.1	19.6	<.10
% Admitted to hospital during pregnancy	25.0	27.6	>.50
% Baby admitted to intensive care	16.6	10.1	<.10
Mean birth weight (grams)	3192	3263	>.30
% Mother breast-fed child	83.4	84.5	>.70

1 t-test for independent samples for comparisons of means, Chi square test of independence for comparison of proportions

Table 5.14 compares the Early Start and Control families studied at the nine year follow-up on a series of measures of maternal adolescent behaviour and psychological adjustment at baseline, including: running away from home; getting in trouble with the police; having problems with alcohol; using illicit drugs; appearing before the Youth Court; becoming pregnant before age 16; using alcohol at least weekly at baseline; using cannabis at least weekly at baseline; and meeting criteria for major depression at baseline. As in previous comparisons there were no significant differences between the two groups.

Finally, Table 5.15 compares the Early Start and Control groups studied at the nine year follow-up on a series of measures of pregnancy and childbirth, including: age at first pregnancy; the percentage of children from earlier pregnancies that were in foster care; whether the pregnancy was unplanned; maternal smoking during pregnancy; maternal cannabis use during pregnancy; mother admitted to hospital during pregnancy; whether the infant was admitted to intensive care after birth; birth weight; and whether the child was breast-fed. There were no significant differences between the two groups.

In summary, Tables 5.11 to 5.15 report 35 comparisons between the Early Start and Control groups studied at the nine year follow-up. In all cases there were no significant differences between the two groups. This is the result that would be expected in a study design in which sample

losses occurred at random for both the Early Start and Control groups, with the rate of this loss being greater for the Early Start group.

5.6.2 Missing data estimation

An alternative approach to examining the effects of sample attrition and missing data is to use methods of missing data estimation to estimate the outcomes that would have been observed for those not followed up at nine years. In this case, the missing data estimation was conducted using the IMPUTE procedure provided by STATA 10.0 (StataCorp, 2007) using 31 variables describing the sample at baseline. These variables spanned: maternal and paternal age; ethnicity; educational qualifications; single parenthood; family size; welfare dependence; family income; debt; inadequacy of family income; being raised in a single parent family; being exposed to interparental conflict or assault; being exposed to child abuse; living in impoverished circumstances; having an unhappy childhood; running away from home; getting in trouble with the police; having problems with alcohol; using illicit drugs; appearing before the Youth Court; becoming pregnant before age 16; using alcohol at least weekly at baseline; using cannabis at least weekly at baseline; meeting criteria for major depression at baseline; age at first pregnancy; the percentage of children from earlier pregnancies that were in foster care; whether the pregnancy was unplanned; maternal smoking during pregnancy; and maternal cannabis use during pregnancy. For those missing baseline observations the outcome measures were set to the overall mean or percentage values. This approach gives a full intention to treat design in which there are observations for all members of both the Early Start and Control groups. The missing data estimation was conducted using two assumptions:

- The first and more conservative assumption was that those lost to follow-up in the Early Start group received no benefit from Early Start.
- The second and more liberal assumption was that Early Start families lost to



follow-up received the same benefit as those Early Start families studied to the nine year follow-up.

Table 5.16 reports the effect size estimates (Cohen's d) and significance levels for three approaches to analysing the data, for the outcome measures that showed significant benefits for Early Start. The first approach reports the observed findings as presented earlier in this chapter. The second approach uses the conservative missing data estimation methods that assume Early Start families lost to follow-up did not receive any benefit from the programme. The third approach uses the liberal assumption that those lost to follow-up in the Early Start group gained similar benefits to those followed up to nine years.

The table shows that, irrespective of which approach to the estimation of the treatment effects is used, the answers are very similar. Specifically:

- For the approach using the observed data, the values of d range from .13 to .29 with a median value of .25. In all cases, these values are statistically significant ($p < .05$).
- For the approach in which missing data was imputed with the assumption that those lost to follow-up in the Early Start group did not benefit from the intervention, values of d ranged from .12 to .28 with a median of .22. In three cases these differences were statistically significant ($p < .05$) and in two cases they were marginally significant ($p < .10$).
- For the approach that assumed those lost to follow-up in the Early Start group gained a similar benefit from the programme to those followed up, values of d ranged from .13 to .30, with a median value of .24.

These findings clearly support the view that losses to follow-up did not pose a major threat to the validity of the trial.

Table 5.16 Supplementary analyses of missing data

Outcome measure	Observed data		Adjusted for missing data (Model 1) ¹		Adjusted for missing data (Model 2) ²	
	p	d (95%CI)	p	d (95%CI)	p	d (95%CI)
% Hospital attendance for accidental injury	<.01	.29 (.09-.49)	<.01	.27 (.08-.46)	<.01	.29 (.10-.48)
% Severe/very severe physical assault by any parent	<.01	.29 (.09-.49)	<.01	.27 (.08-.45)	<.01	.29 (.11-.48)
Physical punishment scale score	<.05	.13 (.01-.24)	=.052	.13 (.02-.23)	<.05	.13 (.03-.24)
Parenting competence scale score	<.0001	.25 (.13-.37)	<.01	.21 (.11-.32)	<.001	.25 (.14-.35)
Parent reported SDQ total score	<.05	.17 (.06-.29)	=.071	.15 (.04-.26)	<.05	.17 (.06-.28)

1 Model 1 assumes those lost to follow-up received no treatment benefit (conservative imputation)

2 Model 2 assumes those lost to follow-up but who were in the Early Start group received treatment benefit (less conservative imputation)

5.6.3 Summary of sample attrition and missing data estimation

On the basis of both comparisons of the Early Start and Control group studied to nine years on their characteristics at baseline and through the use of methods of missing data estimation, there is generally consistent evidence the differential loss to follow-up for the Early Start and Control groups did not pose a detectable threat to study validity. Specifically:

- The comparison of the Early Start and Control groups studied to nine years showed no significant differences between these groups on 35 measures assessed at baseline.
- The application of missing data estimation methods showed the findings of the trial at nine years were similar irrespective of whether the observed or imputed data was analysed.

While differential sample loss must always raise concerns about trial validity, the weight of the evidence clearly suggests the differing rates of follow-up for the Early Start and Control groups were unlikely to pose a serious threat to trial validity.

5.7 Summary

In this chapter, we have extended the analysis of the original Early Start trial to include a nine year follow-up of those enrolled in the trial. The key findings of this analysis may be summarised as follows.

5.7.1 The longer term benefits of Early Start

Findings from the nine year follow-up suggested a number of child related benefits for those families enrolled in Early Start. Children whose families were enrolled in Early Start had fewer hospital attendances for childhood accidents and lower rates of parental reported physical child abuse. In addition, parents in the Early Start group reported more competent and less punitive parenting styles than those in the Control group. In terms of child behaviour, parents in the Early Start group reported lower rates of behaviour problems up to the nine year

follow-up. There were no differences between the two groups in terms of teacher reported behaviour problems, suggesting that changes in childhood behaviour at home did not generalise to the school setting.

Overall, the effect sizes for child related outcomes were relatively modest, ranging from .13 to .29, with a median value of .25. These findings are similar to the results obtained at the 3 year follow-up where there was evidence of a series of small to moderate benefits of Early Start in the areas of child health, child abuse, parenting, and child behaviour.

5.7.2 Early Start and family functioning

While the results suggest benefits for child related outcomes, there was no evidence to suggest Early Start had any benefits for a wide range of outcomes relating to family and parental functioning. These outcomes include: maternal depression; parental substance use; family violence; family economic circumstances; welfare dependence; and the family's exposure to stress. These findings also parallel the results found at the 3 year follow-up, where similar findings were reported.

There are at least two explanations for the apparent failure of the Early Start programme to produce family level change. The first is that the child-centred focus of Early Start means most of the efforts of the programme were focused on producing positive changes for children rather than for the adults in the family. This emphasis may explain why the major benefits of the programme are related to the child rather than to the family or parent. A second feature that may have contributed to these results is that addressing many of the family related issues (e.g. maternal depression; parental substance use; family violence) requires the availability of expert therapeutic resources that are beyond the skills of Family Support Workers. The net result is that to address many parental and family problems required families to be referred to various agencies. The limited changes seen in parental and family related outcomes may reflect the limited availability of high quality

evidence based services with the capacity to address issues such as maternal mental health, parental substance use, family violence, and similar outcomes.

Discussions with Early Start senior staff members suggested three possible reasons the programme did not produce improvements in parental outcomes.

First, in some cases parents facing problems were reluctant to seek help even when offered support to do so by their Family Support Worker. This was clearly the case with maternal depression, for example.

Second, services were sometimes not available in a timely way, with these delays reducing parent commitment to change. This was particularly the case with drug and alcohol services.

Finally, in some cases, parents receiving advice from a service did not follow this advice. A clear example of this was in the use of budget advice services.

These three factors (parental reluctance; lack of availability of services; and problems with compliance) are likely to explain the lack of benefits for parental outcomes.

Following the results of this randomised trial, Early Start has made a number of organisational changes to address some of these issues. These innovations include:

- Regular screening for depression, family planning and family violence.
- The development of performance benchmarks for parental outcomes.
- Increased supervision and monitoring of Family Support staff.

5.7.3 Ethnic differences in outcomes

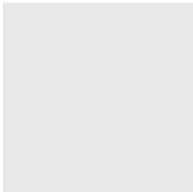
There have been on-going debates about the extent to which generic services such as Early Start have the capacity to address the needs of Māori families (Durie et al., 2010). The findings of this study suggest the outcomes for Māori and non-Māori participants in Early Start were similar, supporting the idea that well-run and culturally appropriate services like Early Start are beneficial for Māori and non-Māori

alike. However, there are important caveats on these conclusions. The first is that the number of Māori families studied (N = 159) was relatively small. Before firm conclusions can be drawn there is a need to study larger samples of Māori and non-Māori families. In addition, the outcomes of the Early Start service have been evaluated using a Western Science model based around a randomised controlled trial and methods of quantitative statistical analysis. The extent to which Early Start can be shown to have benefits on the basis of research using a Kaupapa Māori perspective is not known.

5.7.4 Methodological Issues

An unfortunate feature of the study was that families in the Early Start group dropped out of the research at a greater rate than the families in the Control group did. The principal reason for this differential sample loss was that 14 families randomised to the Early Start service declined to enter the service and participate in the evaluation. In addition the rate of sample loss in the Early Start group was slightly higher than that in the Control group. This difference was probably due to the greater burden for participation placed on the Early Start group. Irrespective of the reasons for the higher rate of drop out of the Early Start group, sample loss poses a potential threat to study validity. A series of checks on this threat using both baseline comparison data and missing data estimation suggested the differential sample loss was unlikely to pose a major threat to study validity. Nonetheless, the possibility the trial may have been biased by differential sample loss cannot be completely discounted.

CHAPTER 6: SUMMARY, CONCLUSIONS AND RECOMMENDATIONS



6.1 Introduction

This report has examined the extent to which children and families provided with Early Start showed beneficial outcomes when compared with a Control group of families not provided with Early Start. These assessments were conducted up to nine years following their enrolment in the trial.

6.2 Summary of child outcomes

6.2.1 Findings to three years

Up to the 3 year follow-up, the outcomes for children in families provided with Early Start showed a number of benefits when compared to the outcomes for children in the Control group. These benefits included:

- Greater use of early health care services
- Reduced rates of hospital attendance for childhood accidents
- Greater use of pre-school education and dental services
- Lower rates of parental reported childhood physical abuse
- Less punitive and more positive parenting
- Lower rates of childhood problem behaviours.

While statistically significant differences were found for these outcomes, the effect sizes for benefits of Early Start were modest. The values of Cohen's d were in the range of .19 to .31, with a median value of .26.

These findings suggest Early Start produced small but pervasive benefits for children. These benefits spanned the areas of: healthcare; injury; use of pre-school services;

reduced risks of physical child abuse; increased positive parenting; and improvements in behavioural adjustment.

6.2.2 Findings to nine year follow-up

The findings to the nine year follow-up showed the number of differences between the Early Start and Control groups persisted up to the 9 year assessment. Specifically, by nine years post-enrolment, children in the Early Start group had:

- Rates of hospital attendance for childhood accidents approximately 50% lower than those for the Control group. These benefits largely reflected a lower rate of accidents in the Early Start group during the pre-school (0–3 year) period.
- Rates of parental reported physical child abuse 50% lower than children in the Control group. Again, these benefits largely reflected the lower rates of physical child abuse during the pre-school (0–3 year) period.
- More positive mean scores on measures of punitive parenting and parenting competence.
- Lower mean scores on measures of parental reported child behaviour problems.

As with the results up to three years, the effect sizes for the benefits of Early Start by the nine year follow-up were modest. The values of Cohen's *d* ranged from .13 to .29 with a median value of .25. The findings to nine years are consistent with the view that the major benefits for children were: reduced risks of childhood accidents; reduced risks of parental reported physical child abuse; more positive parenting; and reductions in rates of problem behaviours.

However, not all findings on child related outcomes to nine years showed positive benefits for Early Start participants. Specifically:

- On the basis of teacher reports of school behaviours and related outcomes, there were no differences between the outcomes of Early Start and Control group families. These results clearly suggest any benefits of Early Start were

largely confined to the home setting and did not generalise to the school setting.

- While families in the Early Start group reported lower rates of agency contact for physical child abuse, reported rates of contact with Child, Youth and Family services were similar to those for the Control group.

6.2.3 Outcomes for families

Despite extensive assessment of parent and family outcomes, there was no evidence to suggest the provision of Early Start had benefits for a wide range of parent and family outcomes at any point of the study. The outcomes considered included:

- Parental smoking
- Contraceptive practice
- Parental substance use
- Maternal depression
- Welfare dependence
- Family income
- Family violence
- Parental separation and conflict.

As a general rule, the families provided with Early Start were a group of families facing multiple disadvantages. The findings of this study suggest this general level of disadvantage was not mitigated in any way by the provision of Early Start. As we have pointed out previously, the major benefits of Early Start appear to be improvements in child related outcomes including: health care; education; child abuse; parenting; and behaviour. Beyond these outcomes the Early Start service appeared to have little or no impact on parental or family functioning. These findings were robust and general, and were evident at all points of observation.

6.2.4 Benefits for Māori

An important policy issue raised by these findings concerns the extent to which Early Start had benefits for Māori. In particular, there have been growing concerns expressed by Māori researchers and academics that mainstream or generic programmes may fail to deliver benefits for Māori (Durie et al., 2010; Smith, 1999). For this reason, efforts have been

made to ensure Early Start was delivered in a culturally appropriate way (see Chapter 2) and to compare the outcomes for both Māori and non-Māori families. The findings of these comparisons showed the outcomes for Māori and non-Māori were similar. These findings support the view that well delivered and culturally appropriate programmes, such as Early Start, may have similar benefits for Māori and non-Māori alike.

A potential limitation of these conclusions is that the evaluation of Early Start has been conducted on a Western Science model centred on a randomised trial. It could be suggested a parallel evaluation using a Kaupapa Māori methodology may have led to a greater insight into both the benefits and limitations of Early Start from a Māori perspective.

6.3 Issues raised by the study findings

6.3.1 Methodological issues

All research studies have limitations which relate to issues of sampling, research design and measurement. The present study is no exception. While the Early Start trial has been well conducted in comparison to many other studies in this area, it has a number of limitations. These include:

- 1) Sampling and recruitment: Ideally, it would have been desirable to base the present trial on a representative sample of families meeting the requirements for entry into Early Start. The study design attempted to address this issue by recruiting families via the Plunket Society, which saw an estimated 95% of all families with infants. However, as part of the recruitment process only 75% of families eligible for the Early Start trial agreed to participate. Because of the Plunket Society's privacy conditions, the characteristics of those declining to enter the trial are unknown. The implications of this feature of the study design are that, while the findings show Early Start had benefits for the families participating in the trial, the extent to which these benefits apply to all families eligible to enter Early Start is unknown.
- 2) Sample retention: A further feature of the research design was that a number of the trial participants dropped out for various reasons over the course of the 9 year study. The result was that, by nine years, just under 84% of the families enrolled in the trial were studied. The dropout rates were substantially higher for those enrolled in Early Start (22%) than for those in the Control group (11%). The reason for the higher rate of dropout in the Early Start group was largely due to 14 families randomised to the Early Start service declining this service. Had these losses not occurred, the rate of sample retention for the Early Start group (83%) would not have differed significantly from the rate of sample retention for the Control group (89%).
To address this threat, a number of methods of missing data estimation were used to adjust for any non-random sample losses. The findings from these adjustments show the study conclusions did not appear to be affected by the differential sample losses in the Early Start and Control groups. Nonetheless, the possibility the trial results may have been biased by the differential dropout of Early Start and Control families cannot be entirely discounted.
- 3) Measurement: The data gathered in this study largely relies on parental self-reporting. Since the parents involved in the study could not be 'blind' to the treatment condition to which they were assigned, there is a possibility the study findings could be influenced by reporting biases. To some extent, the possibility of these biases was addressed by obtaining data from sources that were blind to the treatment group to which the family belonged. These sources included both official hospital records and teacher reporting data. The results from these sources were mixed. In particular, hospital record data showed children in the Early Start group had substantially lower rates of hospital attendance for childhood accidents, suggesting the results for this outcome were not distorted by recall bias. On the other hand, measures of behaviour based on

parental and teacher reporting led to different conclusions. Parental reported outcomes suggested benefits for Early Start, whereas teacher reports failed to show this. These findings could reflect a bias in parental reporting, but they could also reflect that parents and teachers were describing behaviours in different contexts. In short, the evidence from the study of hospital admissions suggests the positive findings of the trial cannot be entirely dismissed as being due to response bias: the results from parent and teacher reports suggest that the possibility some findings have been influenced by response bias cannot be dismissed.

The methodological problems associated with Early Start reflect problems common to many RCTs conducted in 'naturalistic' settings using human participants (Hernán et al., 2004; Schulz and Grimes, 2002). While these problems need to be recognised and taken into account, the weight of the evidence from the trial clearly suggests the Early Start programme had a number of short term and long term benefits. These benefits included: improved access to early health care; increased use of pre-school education and related services; reduced rates of hospital attendance for childhood accidents; reduced rates of parental reported physical child abuse; increases in positive parenting; and increases in parental reported child behaviours.

6.3.2 Comparison with previous research

There has been a large amount of research into the benefits of home visiting programmes such as Early Start. While there has been considerable enthusiasm for this approach in the literature, the results of rigorous trials of home visiting have been disappointing. Many studies have failed to find benefits for this approach (Daro and Harding, 1999; Gomby, 2000; MacMillan et al., 2005; Olds et al., 1999). Two programmes that are exceptions to this trend are the Nurse Family Programme (NFP) developed by Olds and his associates, and the Early

Start programme. Findings from the Nurse Family Programme have reported benefits in a number of areas including: lower levels of child abuse; lower levels of maternal maladaptive behaviour; reduced welfare dependence; increased levels of child safety; and reduced risk of childhood injury. In addition, the effects of the programme were evident for juvenile antisocial behaviour, offending and substance use outcomes assessed 15 years following enrolment in the programme (Olds et al., 1997; Olds et al., 1998; Olds et al., 1994). These findings clearly suggest that well designed and well-implemented home visiting programmes can have both short term and long term benefits for the children participating in these programmes.

Early Start and the NFP share three features which may explain their success when compared with other home visiting models.

- 1) Research base: Both programmes were evolved in a research context and were motivated to develop evidence based approaches to addressing the multiple problems faced by disadvantaged families.
- 2) Programme standardisation: In the development of both programmes, considerable efforts have been made to ensure programme standardisation and fidelity of programme delivery through the development of service manuals and methods of worker supervision.
- 3) Employment of professional workers: Both programmes use professionally trained staff to deliver the programme. The NFP uses trained nurses, and Early Start uses workers with professional qualifications in the areas of nursing, social work and education.

These parallels between the Nurse Family Programme and Early Start suggest that home visiting programmes lead to beneficial short term and long term consequences for children, providing the following conditions are met:

- The home visiting programme is well founded on research.
- Substantial investments are made in ensuring the fidelity of programme delivery, including the development and

use of service manuals, staff training and adequate worker supervision.

- The programme is provided by trained workers with tertiary level qualifications.

6.4 Policy implications of study findings

6.4.1 Implications for family and parental outcomes

The findings of this 9 year evaluation of Early Start show the programme is effective in improving outcomes for children but it does not have any detectable benefits for parental outcomes or family functioning. These findings suggest that, to address the range of problems facing disadvantaged families, home visiting services such as Early Start need to be set in the context of a broader range of effective evidence based services aimed at meeting the various psychological, social, emotional and economic needs of these families. Such services include:

- Family planning and contraceptive advice
- Parental mental health services
- Support and treatment services for parental substance use
- Educational and career support
- Family budgeting services
- Family relationship services.

Ideally, the best approach to addressing these issues would be through the development of an integrated Child, Parent and Family service that has the capacity to address the range of psychological, social and economic needs of families facing multiple challenges. Home visiting programmes such as Early Start can make positive contributions to improving childhood outcomes but they are not substitutes for well-integrated sets of evidence based services for families. The next generation of research and policy development needs to consider ways of aligning, integrating and coordinating the diverse array of agencies involved in child, parent and family issues so these services can provide better outcomes for children, parents and families. There is now a growing literature which shows efforts to develop integrated child and family services improves

outcomes for both families and children (Ofsted, 2009). The importance of developing greater integration and collaboration of child, family and related services has also been recognised in a New Zealand context in the Strengthening Families (Ministry of Health, 1996) and Whanau Ora policy frameworks (Durie et al., 2010; Ihimiera, 2007).

In short, the findings of the present study clearly suggest the benefits of home visiting programmes such as Early Start could be strengthened if these services were located within an integrated system of evidence based child, parent and family services.

6.4.2 Will the results from Early Start apply to Family Start?

An important policy issue for New Zealand concerns the extent to which the findings from Early Start can be applied to the Family Start service. The Family Start service was set up in 1998 with a broadly similar agenda to Early Start to address the needs of disadvantaged families facing stress and difficulty. In comparing Early Start and Family Start, (Cribb, 2009) notes that the two programmes have similar goals and approaches, but they differ in a number of ways:

“The two programmes are similar in that they are home visiting models targeted to the most at-risk families. The differing history has led to differences in emphasis. Early Start, initially created as a randomised control trial programme, maintains a strong link to research, has set benchmarks for itself and focuses on consistent delivery and ‘fidelity’ to the programme design. Family Start developed from the Government’s Strengthening Families Strategy and focused on contracting appropriate providers and localised responses, while also identifying key aspects that needed to be delivered.” (p. 6)

Given the differences between Family Start and Early Start, it is unclear whether or not the benefits found for Early Start will apply to Family Start. Further, because of the variability in the services provided by Family Start providers, this issue will need to be

addressed on a site by site basis. There is a clear need for further research to examine the extent to which the outcomes achieved by specific Family Start programmes are similar to the outcomes achieved by Early Start.

6.4.3 Do the findings apply to the present day Early Start Service?

The present study was based on the Early Start service as it was provided in 1998–1999. Since that time a number of important changes have occurred which have implications for the effectiveness of the service. These changes have been both positive and negative.

On the positive side, the service has been strengthened by: a) the development of more extensive manuals and staff guidelines; b) increased staff supervision; c) the inclusion of evidence based programmes including Partners in Parenting Education (PIPE) (Robinson et al., 1997) and the Triple P and Incredible Years Parenting programmes (Reid et al., 2004; Sanders, 1999); and d) the introduction of more targeted services for depression, contraception, family violence and school readiness. These developments have been underwritten by a collaborative relationship between Early Start and the Christchurch Health and Development Study, in which the CHDS has provided feedback from the research process to strengthen the service provided by Early Start. These innovations should have strengthened the quality control and evidence base of the programme and made it more effective.

On the other hand there have been changes to the referral system which mean Early Start addresses a more challenging clientele than was the case in 1999. Specifically, as a result of requests from funders, Early Start has stopped using the client screening system provided by Plunket described in this report (see Figure 3.1). It now operates a system in which multiple providers including Plunket nurses, midwives, GPs, CYF staff and others may refer families to the service. In the view of Early Start staff members, this change in

the referral system has led to families who face a greater range of challenges than the families studied in this report being referred to the service.

On balance, it seems likely the improvements in the Early Start service since 1999 coupled with the service's changing clientele have produced a situation in which the outcomes of the service today are similar to those described in this report. These conclusions are supported by a series of benchmarks Early Start uses to quality control service outcomes. These benchmarks are based on the results of the randomised trial results up to 36 months, described in Chapter 4. The monitoring of Early Start outcomes using these benchmarks has shown that the outcomes of the present day Early Start service are as good, if not better, than the outcomes for the randomised trial. These findings provide considerable reassurance that the findings of the randomised trial described in this report are likely to apply to the present day Early Start service.



References

- Achenbach, T. M., McConaughy, S. H., & Howell, C. T. (1987). Child/adolescent behavioral and emotional problems: Implications of cross-informant correlations for situational specificity. *Psychological Bulletin*, 101, 213-232.
- Arnold, D. S., O'Leary, S. G., Wolff, L. S., & Acker, M. M. (1993). The Parenting Scale: A measure of dysfunctional parenting in discipline situations. *Psychological Assessment*, 5(2), 137-144.
- Barwick, H. (1992). Youth Suicide Prevention Project. Workshop report and literature review. Wellington: Department of Health.
- Bavolek, S. J., & Keene, R. (1999). *Adult-Adolescent Parenting Inventory-2: Revised and reformed*. Park City, UT: Family Development Resources, Inc.
- Block, J. H. (1981). The Childrearing Practices Report: A set of Q items for the description of parental socialization attitudes and values. Berkeley, CA: Unpublished manuscript, Institute of Human Development, University of California.
- Briggs-Gowan, M. J., & Carter, A. S. (1998). Preliminary acceptability and psychometrics of the Infant-Toddler Social and Emotional Assessment (ITSEA): A new adult-report questionnaire. *Infant Mental Health Journal*, 19, 422-445.
- Chaffin, M. (2004). Is it time to rethink Healthy Start/Healthy Families? *Child Abuse & Neglect*, 28, 589-595.
- Coggan, C., & Norton, R. (1994). Reducing self-directed harm (suicide and attempted suicide) among young people: A public health approach? *Community Mental Health in New Zealand*, 8, 26-31.
- Cribb, J. (2009). Review of Family Start and Early Start. Wellington: Ministry of Social Development.
- Daro, D. (1994). Healthy Families America: A guide for evaluating Healthy Families America efforts. Chicago: National Committee to Prevent Child Abuse.
- Daro, D. A., & Harding, K. A. (1999). Healthy Families America: Using research to enhance practice. *Future Child*, 9(1), 152-176.
- Dekovic, M., Janssens, J. M. A. M., & Gerris, J. R. M. (1991). Factor structure and construct validity of the Block Child Rearing Practices Report (CRPR). *Journal of Consulting and Clinical Psychology*, 3(2), 182-187.
- Drugs Advisory Committee (1995). Cannabis and Health in New Zealand. Wellington: Drugs Advisory Committee.
- Duggan, A., Fuddy, L., Burrell, L., Higman, S. M., McFarlane, E., Windham, A., et al. (2004). Randomized trial of a statewide home visiting program to prevent child abuse: Impact in reducing parental risk factors. *Child Abuse & Neglect*, 28(6), 623-643.
- Durie, M., Cooper, R., Snively, S., Grennell, D., & Tuaine, N. (2010). Whānau Ora: Report of the Taskforce on Whānau-Centred Initiatives. Wellington: Ministry of Social Development.
- Durie, M. H. (1998). *Whaiora. Maori Health Development* (2nd ed.). Auckland: Oxford University Press.
- Fanslow, J., McGregor, K., Coggan, C., Bennett, S., & McKenzie, D. (2000). Research into programmes to prevent intentional injury and violence to children. *Centre Report Series* (Vol. 52). Auckland: Injury Prevention Research Centre, University of Auckland.
- Fergusson, D. M., Boden, J. M., & Horwood, L. J. (2009). Situational and generalised conduct problems and later life outcomes: Evidence from a New Zealand birth cohort. *Journal of Child Psychology & Psychiatry*, 50(9), 1084-1092.
- Fergusson, D. M., Grant, H., Horwood, L. J., & Ridder, E. (2005a). Randomized trial of the Early Start Program of home visitation. *Pediatrics*, 116(6), e803-e809.
- Fergusson, D. M., Grant, H., Horwood, L. J., & Ridder, E. M. (2006a). Client satisfaction with the Early Start service. *Social Policy Journal of New Zealand*, 28, 179-196.

- Fergusson, D. M., Grant, H., Horwood, L. J., & Ridder, E. M. (2006b). Randomized trial of the Early Start Program of home visitation: Parent and family outcomes. *Pediatrics*, 117(3), 781-786.
- Fergusson, D. M., & Horwood, L. J. (1989). Estimation of method and trait variance in ratings of conduct disorder. *Journal of Child Psychology & Psychiatry & Allied Disciplines*, 30(3), 365-378.
- Fergusson, D. M., Horwood, L. J., & Grant, H. (1998). The Development of Early Start. Christchurch: The Early Start Consortium.
- Fergusson, D. M., Horwood, L. J., Grant, H., & Ridder, E. (2005b). Early Start Evaluation Report. Christchurch: Early Start Project Ltd. <http://www.msd.govt.nz/work-areas/social-research/early-start.html>.
- Fergusson, D. M., Horwood, L. J., & Lynskey, M. T. (1994). The childhoods of multiple problem adolescents: A 15-year longitudinal study. *Journal of Child Psychology & Psychiatry*, 35(6), 1123-1140.
- Gomby, D. S. (2000). Promise and limitations of home visitation. *Journal of the American Medical Association*, 284(11), 1430-1431.
- Gomby, D. S., Culross, P. L., & Behrman, R. E. (1999). Home visiting: Recent program evaluations - analysis and recommendations. *Future Child*, 9(1), 4-26.
- Goodman, R. (1997). The Strengths and Difficulties Questionnaire: A research note. *Journal of Child Psychology and Psychiatry*, 38(5), 581-586.
- Hanson, R. A. (1990). Initial parenting attitudes of pregnant adolescents and a comparison with the decision about adoption. *Adolescence*, 15(99), 629-643.
- Hawaii Department of Health (1992). Healthy Start. Report to the 16th Legislature. State of Hawaii: Hawaii Department of Health.
- Hernán, M. A., Hernández-Díaz, S., & Robins, J. M. (2004). A structural approach to selection bias. *Epidemiology*, 15(5), 615-625.
- Holmes, S. J., & Rahe, R. (1967). The Social Readjustment Rating Scale. *Journal of Psychosomatic Research*, 11, 213-218.
- Howden-Chapman, P., Bushnell, J., & Carter, C. (1994). Prescription and Permission: An analysis of possible philosophical bases for Drug and Alcohol Policy in New Zealand. Wellington: Health Services Research Centre.
- Ihimaera, L. V. (2007). Whakarato Whanau Ora: Whanau wellbeing is central to Maori wellbeing. In T. R. Matatini (Ed.), (pp. 35). Palmerston North.
- Jensen, J., Spittal, M., & Krishnan, V. (2005). ELSI Short Form: User manual for a direct measure of living standards. Wellington: Ministry of Social Development.
- MacMillan, H. L., Thomas, H., Jamieson, E., Walsh, C. A., Boyle, M. H., Shannon, S., et al. (2005). Effectiveness of home visitation by public-health nurses in prevention of the recurrence of child physical abuse and neglect: A randomised controlled trial. *The Lancet*, 365, 1786-1793.
- McCombs Thomas, A., Forehand, R., Armistead, L., Wierson, M., & Fauber, R. (1990). Cross-informant consistency in externalizing and internalizing problems in early adolescence. *Journal of Psychopathology and Behavioral Assessment*, 12(3), 255-262.
- McGeorge, P. (1995). Child, Adolescent and Family Mental Health Services. Wellington: Ministry of Health.
- Ministry of Health (1994). Report and Recommendations of the Steering Group on Youth Mental Health and Suicide Prevention. Wellington: Ministry of Health.
- Ministry of Health (1996). Child Abuse Prevention: The Health Sector's Contribution to the 'Strengthening Families' Initiative. Wellington, New Zealand: Ministry of Health.
- Ministry of Health (1998). Child Health Programme Review. Wellington, New Zealand: Ministry of Health.
- Ofsted (2009). The impact of integrated services on children and their families in Sure Start children's centres. London: Ofsted.
- Olds, D. L., Eckenrode, J., Henderson, C. R., Jr., Kitzman, H., Powers, J., Cole, R., et al.

- (1997). Long-term effects of home visitation on maternal life course and child abuse and neglect. Fifteen-year follow-up of a randomized trial. *Journal of the American Medical Association*, 278(8), 637-643.
- Olds, D. L., Henderson, C. R., Jr., Cole, R., Eckenrode, J., Kitzman, H., Luckey, D., et al. (1998). Long-term effects of nurse home visitation on children's criminal and antisocial behavior: 15-year follow-up of a randomized controlled trial. *Journal of the American Medical Association*, 280(14), 1238-1244.
- Olds, D. L., Henderson, C. R., Jr., & Kitzman, H. (1994). Does prenatal and infancy nurse home visitation have enduring effects on qualities of parental caregiving and child health at 25-50 months of life? *Pediatrics*, 93(1), 89-98.
- Olds, D. L., Henderson, C. R., Jr., Kitzman, H., Eckenrode, J. J., Cole, R. E., & Tatelbaum, R. C. (1999). Prenatal and infancy home visitation by nurses: Recent findings. *Future Child*, 9(1), 44-65.
- Olds, D. L., & Kitzman, H. (1993). Review of research on home visiting for pregnant women and parents of young children. *The Future of Children*, 3(3), 53-92.
- Public Health Group (1996). Cannabis: The Public Health Group's Advice to the Minister of Health 1995-1996. Wellington: Public Health Group.
- Reid, M. J., Webster-Stratton, C., & Baydar, N. (2004). Halting the development of conduct problems in head start children: The effects of parent training. *Journal of Clinical Child & Adolescent Psychology*, 33(2), 279-291.
- Report of the Education and Science Committee (1995). Inquiry into Children in Education at risk through Truancy and Behavioural Problems. Wellington: New Zealand House of Representatives.
- Robinson, J. L., Emde, R. N., & Korfmacher, J. (1997). Integrating an emotional regulation perspective in a program of prenatal and early childhood home visitation. *Journal of Community Psychology*, 25(1), 59-75.
- Sanders, M. R. (1999). Triple P-Positive Parenting Program: Towards an empirically validated multilevel parenting and family support strategy for the prevention of behavior and emotional problems in children. *Clinical Child and Family Psychology Review*, 2(2), 71-90.
- Schulz, K. F., & Grimes, D. A. (2002). Sample size slippages in randomised trials: exclusions and the lost and wayward. *The Lancet*, 359(9308), 781-785.
- Smith, L. T. (1999). *Decolonizing methodologies: Research and indigenous peoples*. London: Zed Books.
- Spiker, D., Kraemer, H. C., Constantine, N. A., & Bryant, D. (1992). Reliability and validity of behavior problem checklists as measures of stable traits in low birth weight, premature preschoolers. *Child Development*, 63(6), 1481-1496.
- StataCorp (2007). Stata Statistical Software: Release 10.0. Texas: Stata Corporation, College Station.
- Straus, M. A., Hamby, S. L., Boney-McCoy, S., & Sugarman, D. B. (1996). The Revised Conflict Tactics Scales (CTS2). Development and preliminary psychometric data. *Journal of Family Issues*, 17(3), 283-316.
- Straus, M. A., Hamby, S. L., Finkelhor, D., Moore, D. W., & Runyan, D. (1998). Identification of child maltreatment with the Parent-Child Conflict Tactics Scales: Development and psychometric data for a national sample of American parents. *Child Abuse and Neglect*, 22(4), 249-270.
- Touliatos, J., & Lindholm, B. W. (1981). Congruence of parents' and teachers' ratings of children's behavior problems. *Journal of Abnormal Child Psychology*, 9(3), 347-354.
- Windham, A. M., Rosenberg, L., Fuddy, L., McFarlane, E., Sia, C., & Duggan, A. K. (2004). Risk of mother-reported child abuse in the first 3 years of life. *Child Abuse & Neglect*, 28(6), 645-667.
- World Health Organization (1993). *Composite International Diagnostic Interview (CIDI)*. Geneva, Switzerland: World Health Organization.

APPENDIX: STATISTICAL METHODS USED IN CHAPTER 5

The data reported in Chapter 5 examines the outcomes of the Early Start and Control groups up to the nine year follow-up (Tables 5.2 to 5.5, 5.7 and 5.8). These tables report on a series of repeated measures analyses in which the outcomes of the Early Start and Control groups are compared on an outcome (Y) observed at multiple time points. To test for differences between the Early Start and Control groups a unified analysis methodology was used by fitting the model in EQ 1 to each outcome measure:

$$f(Y_t) = B_0 + B_1X + B_2 t + B_3(X.t) + U \quad (\text{EQ 1})$$

where Y_t is the measure of the outcome Y observed at time t; X is a dichotomous variable representing treatment status (Early Start/Control); and (X.t) is an interaction term. This model addresses three questions:

- 1) Treatment effects: The parameter B1 estimates the overall treatment difference in the outcome Y over the observation period.
- 2) Time effects: The parameter B2 represents changes in the expected value of Y over time.

- 3) Time by treatment interaction: The parameter B3 represents the extent to which any effects of treatment on the outcome Y varied with time.

This model was fitted using population-averaged Generalised Estimating Equation (GEE) models. In these models the term $f(Y_t)$ denotes the link function. Following the usual conventions for the generalised linear model, the link function is logistic for dichotomous outcomes, log for count measures, and identity for means.

The reason for using this analytic methodology is that it provided a consistent approach to examining treatment, time and treatment by time interaction in the data gathered up to the nine year follow-up.

The tables below provide estimates of the model parameters, standard errors and tests of significance for the data presented in Tables 5.2 to 5.5, 5.7 and 5.8.

Statistical analyses

Table 5.2 Parameter estimates for GEE models (logistic link function) of hospital attendance for accidental injury, 0–9 years

Effect	Parameter estimate (B)	SE	p
Treatment (B1)	0.18	.09	<.05
Time (B2)	-0.62	.11	<.0001
Treatment x Time (B3)	-0.12	.11	>.20

Table 5.3 Parameter estimates for GEE models (logistic link function) of severe/very severe physical assault by any parent, 0–9 years

Effect	Parameter estimate (B)	SE	p
Treatment (B1)	0.43	.15	<.05
Time (B2)	-0.38	.17	<.05
Treatment x Time (B3)	-0.04	.17	>.70

Table 5.4 Parameter estimates for GEE models (log link function) of mean physical punishment scores, 0–9 years

Effect	Parameter estimate (B)	SE	p
Treatment (B1)	0.12	.06	<.05
Time (B2)	-0.48	.03	<.0001
Treatment x Time (B3)	-0.01	.03	>.80

Table 5.5 Parameter estimates for GEE models (identity link function) of mean parenting competence scale scores, ages five, six and nine years

Effect	Parameter Estimate (B)	SE	p
Treatment (B1)	-0.25	.09	<.01
Time (B2)	0.00	.02	>.90
Treatment x Time (B3)	0.02	.02	>.30

Table 5.7 Parameter estimates for GEE models (identity link function) of mean parent reported SDQ externalizing, internalizing, and total scores, ages five, six and nine years

Effect	Parameter estimate (B)	SE	p
Externalizing Treatment (B1)	0.17	.07	<.05
Externalizing Time (B2)	0.00	.02	>.90
Externalizing Treatment x Time (B3)	-0.03	.02	>.20
Internalizing Treatment (B1)	0.17	.09	<.05
Internalizing Time (B2)	0.00	.02	>.80
Internalizing Treatment x Time (B3)	0.00	.02	>.90
Total score Treatment (B1)	0.17	.08	<.05
Total score Time (B2)	0.00	.02	>.90
Total score Treatment x Time (B3)	-0.02	.02	>.30

Table 5.8 Parameter estimates for GEE models (identity link function) of mean teacher reported SDQ externalizing, internalizing, and total scores, ages five, six and nine years

Effect	Parameter estimate (B)	SE	p
Externalizing Treatment (B1)	0.02	.09	>.70
Externalizing Time (B2)	-0.00	.03	>.90
Externalizing Treatment x Time (B3)	-0.00	.03	>.90
Internalizing Treatment (B1)	-0.11	.07	>.10
Internalizing Time (B2)	0.00	.03	>.90
Internalizing Treatment x Time (B3)	-0.04	.03	>.10
Total score Treatment (B1)	-0.03	.09	>.70
Total score Time (B2)	-0.00	.03	>.80
Total score Treatment x Time (B3)	-0.04	.03	>.10



EARLY START