Early Start

Evaluation Report

Early Start
Evaluation Report

David Fergusson
John Horwood
Elizabeth Ridder
Christchurch Health and Development Study
Department of Psychological Medicine
Christchurch School of Medicine and Health Sciences
Christchurch, New Zealand
email: david.fergusson@chmeds.ac.nz

Hildegard Grant
Early Start Service
Christchurch, New Zealand

November 2005

CONTENTS

LIST OF TABLES

LIST OF FIGURES

ACKNOWLEDGEMENTS

FOREWORD

The Ministry of Social Development is pleased to support the publication of the Early Start Evaluation Report as part of the 10-year anniversary celebrations for the Christchurch Early Start project. One of the most important tasks of government is making provision for the healthy development of children, especially those in vulnerable circumstances. One means of doing so is through early intervention initiatives aimed at supporting parents in nurturing their development. In developing programmes for this purpose, it is equally important that we endeavour to understand whether they are achieving these goals.

We welcomed the opportunity extended earlier this year by Professor David Fergusson to social sector agencies to provide input into the development of an evaluation report on the Early Start programme. We believe this report makes a significant landmark contribution to our knowledge about "what works" in this important area and underlines the value of using robust evaluation designs, particularly randomised controlled trials, in providing clear answers about the impacts of well designed programmes.

We congratulate everyone involved in developing and delivering the Early Start programme to the high standard necessary for the outcomes reported. We especially acknowledge the Early Start Service, its partner provider agencies and the Christchurch Health and Development Study team.

Marcel Lauzière
Deputy Chief Executive
Social Development Policy & Knowledge
Ministry of Social Development

Allen Gomes
Family, Child, Youth & Community,
Research & Evaluation Unit
Ministry of Social Development

EXECUTIVE SUMMARY
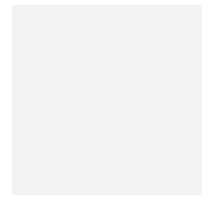
This report provides an account of the development, implementation and evaluation of the Early Start programme. The report is presented in a series of chapters that set the background to the development of the Early Start programme, the design of the randomised trial and the outcomes of the evaluation of the Early Start programme.

Chapter 1
Introduction

Chapter 1 sets the background for the development and evaluation of the Early Start programme and provides a brief review of the home visitation literature.

In the early 1990s, a consortium of providers including the Family Help Trust and the Christchurch Health and Development Study came together in Christchurch to develop a home-based family support programme. This programme became known as Early Start and, following an initial pilot study, funding was provided by the Family Start initiative. This enabled the Early Start consortium to establish a randomised trial of the service.

To set the background to this trial, the report provides a brief literature review of the results of randomised trials of home visitation. The evidence from home visitation programmes shows that:

- Overall, the results of the impact of home visitation are mixed and inconclusive.
- Many programmes fail to produce positive effects while other programmes have done so.
- There is an urgent need to identify the factors that make programmes work.
- There is a need to develop home visitation programmes that are well designed and well implemented to produce positive outcomes for at-risk children and families.

Chapter 2
Background to the Early Start Programme

Chapter 2 sets the background to the Early Start programme, providing a detailed description of the Early Start service, the rationale and processes for client recruitment, and the principles of service delivery. A detailed case history is provided to illustrate the process of service provision to a family enrolled in Early Start.

The delivery of Early Start centred around a process of home visitation aimed at achieving a series of goals in the areas of: improvements in child health; reduction of child abuse; improvements in parenting skills; supporting parental physical and mental health; encouraging family economic and material well-being; and encouraging stable positive partnerships.

The philosophy of the programme was to build collaborative, trusting and problem-solving partnerships between clients and family support workers to build on strengths and eliminate deficits to maximise the health and well-being of children and their families.

Chapter 3
The Justification and Planning of the Randomised Trial

Chapter 3 discusses the justification and planning of the randomised trial. A critical component of programme development was to thoroughly and rigorously evaluate the effectiveness of the Early Start programme to achieve its goals.

To achieve this goal, a randomised controlled trial was used to compare the outcomes of 220 families involved in Early Start with a randomly assigned control group of 223 families not involved in Early Start.

Families in both groups were interviewed at the point of trial enrolment and at six, 12, 24 and 36 months post-enrolment using a home-based interview. Information was also obtained from general practitioner and hospital records.

Chapter 4
Client Recruitment, Characteristics and Retention

Chapter 4 provides statistical information on the recruitment process, the characteristics of the Early Start and control families at the point of enrolment in the randomised trial, the patterns of participation in the Early Start service and the patterns of participation in the data collection process.

The information gathered at the baseline showed that:

- As a group, the families entering the randomised trial were subject to disadvantages in a number of areas.
- In all comparisons, there were no statistically significant differences between the Early Start and control groups, suggesting that the assignment to groups produced equivalent groups of families.

With the passage of time there was a decline in the number of clients actively receiving the Early Start service. At 36 months post-enrolment, 59% of clients were receiving the programme. However, there was no association between the extent of service delivery and a range of social, family and personal factors, suggesting that losses to the service tended to occur in a non-systematic way.

Overall, the analysis suggested high rates of research participation with nearly 90% of those enrolled in the trial being studied for 36 months. Dropout from the research assessment was unrelated to a wide range of social, individual, family and related factors.

Chapter 5
Child-Related Outcomes

Chapter 5 examines the extent to which children in the Early Start group showed improved outcomes in a number of areas relating to health, early childhood education, parenting, child abuse and neglect and behavioural adjustment. Parents were interviewed at six, 12, 24 and 36 months. Health information was cross-validated with general practitioner and hospital records.

Compared to the control group, children receiving Early Start had:

- Improved health care and health outcomes (greater use of general practitioners, higher rates of well-child checks, fewer hospital attendances for accidents, injuries or poisonings, and greater use of preschool dental services).
- Increased exposure to early childhood education.
- Increased exposure to positive parenting practices (positive and non-punitive parenting).
- Lower rates of severe/very severe physical assault by parents.
- Reduced rates of externalising and internalising behaviour problems.

Chapter 6
Maternal and Family Outcomes

Chapter 6 examines the extent to which the Early Start programme showed positive outcomes in the areas of: maternal health and well-being; family stability, family relationships and family violence; family economic and material well-being; and family exposure to stress and adversity.

There was a consistent lack of association (p > .05) between maternal and family outcomes and group status, and there was no systematic trend for one group to fare better than the other group.

Comparison with the results from Chapter 5 suggests that while the Early Start programme led to positive changes in parenting and child outcomes the programme showed little apparent benefits in the areas of maternal and family outcomes.

Chapter 7
Further Analysis

Chapter 7 provides an extension to the analyses reported in previous chapters. The chapter focuses on two key issues: effect modification and treatment of missing data.

Family Features and Programme Benefits

It could be proposed that the benefits of the programme may vary with the type of family to which it is supplied. A series of analyses extended the research design to examine the extent to which the programme benefits were similar among different types of families, including Māori families, and families facing multiple sources of stress and difficulty. The findings suggested that, in general, the Early Start programme had similar benefits for client families regardless of ethnicity, family size, maternal age or level of family disadvantage.

Missing Data

Missing data estimation methods were used to estimate the study results that would have been observed had all study participants been observed at all time periods. Adjustment for missing data led to very similar conclusions to the main analysis, suggesting that missing data did not pose a threat to study validity.

Chapter 8
Summary, Conclusions and Recommendations

Chapter 8 provides an overview of the findings of this randomised trial of Early Start. First, the chapter provides a summary of the overall findings of the trial and considers threats to trial validity. Second, the findings of the trial are compared with results from both international and New Zealand studies of home visitation. This shows that, in comparison to other evaluations of home visitation programmes, Early Start performed better than all the other programmes reviewed, with perhaps one exception (the Nurse Family Partnership Program). Finally, the chapter provides an examination of issues for the future development and evaluation of Early Start.

CHAPTER 1:
INTRODUCTION
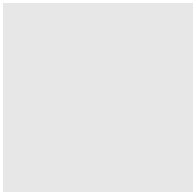
1.1 Overview

This report describes an evaluation of the Early Start programme. Early Start is a home visitation programme that has been developed in Christchurch since 1995 by a consortium of providers that has included: the Family Help Trust; the Christchurch Health and Development Study; the Southern Regional Office of the Royal New Zealand Plunket Society; the Pegasus Health GP group; representation from the Department of Child, Youth and Family Services, Māori representatives and community members. The aims of this consortium have been to develop an intensive home-based family support system to meet the needs of high-risk families and their children.

This report describes the findings from a randomised trial aimed at providing an outcome evaluation of the programme. In this trial, the outcomes for 220 families receiving the programme were compared with the outcomes for a control group of 223 families not receiving the programme. This chapter sets the background for this evaluation and examines the following issues:

- The background history that led to the development of the Early Start programme.
- The general principles on which the programme was based.
- The results of the pilot evaluation.
- An overview of the randomised trial.
- A brief review of the literature on home visitation.

1.2 The Development of the Early Start Programme

The impetus for the development of the Early Start programme began in the early 1990s as a result of growing recognition within New Zealand of increasing rates of psychosocial problems in children. These problems 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, Bushnell, & Carter, 1994; Public Health Group, 1996); child and adolescent mental health (McGeorge, 1995; Ministry of Health, 1994) and youth suicide (Barwick, 1992; Coggan & Norton, 1994; Ministry of Health, 1994). It became increasingly apparent that 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, Horwood, & Lynskey, 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. The clear implication of this result was that if substantial progress was to be made in the area of addressing childhood and adolescent problems there was a need to address the difficulties and stresses faced by children reared in severely disadvantaged, dysfunctional or chaotic home environments.

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 well-being or material standards of high-risk families. However, inspection of the childhoods of multiple-problem children in the Christchurch Health and Development Study clearly suggested it was unlikely that 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 a focus on the provision of traditional income support services and towards the identification of programmes that provided at-risk families with direct support in the areas of parenting, childrearing 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 upon methods of home-based visitation designed to meet the needs of at-risk families. It was also suggested at this meeting that the Hawaii Healthy Start Program provided a model that might be adapted to the New Zealand context. Healthy Start is an 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 then offered

the Healthy Start Program. Families who accept the offer (between 80% and 90% of those eligible for the programme) are then provided with intensive family support provided 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 that this would be desirable and that an important first stage of programme development was to conduct a process evaluation of the programme 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 visitation programme to meet the needs of at-risk families?

• How effective was this approach in leading to improvements in the well-being 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. Initial exploration of this issue suggested that 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 Plunket Society has developed strong linkages with other service providers to ensure that at-risk families are visited. For these reasons, the emerging Early Start group contacted the Southern Regional Office of the Royal New Zealand 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 upon the principles of Healthy Start.

The second hurdle was obtaining funding for programme development. In the first instance, the consortium was successful in obtaining funding from Canterbury Trustbank Community Trust. The acquisition of this funding placed the consortium in a position to develop concrete plans to develop a home-based family support service for at-risk families. 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, 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.

To gain further background in this area, two representatives of the Family Help Trust visited the Hawaii Healthy Start Program in early 1995 to learn first-hand about methods of screening and to study the service delivery used in Healthy Start.

The development of Early Start received considerable impetus in 1995 as a result of a nationwide tour made by Dr Calvin Sia and

Ms Gail Breakey from the Hawaii Healthy Start Program. In this visit, the representatives toured New Zealand providing an overview of the Hawaiian programme and its underlying principles.

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

The net result of this process was that by mid-1995 a consortium of providers – including the Family Help Trust, the Christchurch Health and Development Study, the Plunket Society, Māori representatives and Pegasus Health – had been assembled. 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 50 families.

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

1.3 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 that 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 Program to a Christchurch context. The key features of the Early Start programme are described below.

1.3.1 Client Identification

A detailed account of the application of this client-identification system is given later in this report. Briefly, this system 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 into Early Start for a one-month probationary period. This period gave the family an opportunity to become acquainted with the programme and also 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 prespecified 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 that families were enrolled into 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 demands of treating families in an ethical and non-stigmatising way. This has been achieved by the development of a client-identification system that combines elements of population screening, client referral and needs assessment. 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) and ensures that clients are enrolled into 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.

1.3.2 Service Provision

Early Start comprises a system of home-based family support and visitation provided by trained family support workers. Their task 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 each family's particular circumstances and needs rather than being based on a predetermined programme. This flexibility of the service provision, however, 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 that children have adequate access to and use of child health services including: immunisation; preventive health care and timely visits for childhood morbidity. The key features of the service that lead to the achievement of this goal include: ensuring that all families are enrolled with a single general practitioner who acts as the health care provider for the family; the support and encouragement of mothers in their use of child health care services; and the development of close liaison and linkages with key health care providers including general practitioners, Plunket nurses and other services.

2. *Maternal well-being*
   Ensuring that 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. Accordingly, 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; mental health problems; and other sources of social and emotional stress.

3. *Parenting skills*
   Helping mothers acquire and develop adequate parenting skills. As is 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 assist and encourage mothers in the development of 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 dependent on welfare benefits, and second, many families have limited budgeting and financial management skills. This combination of limited income 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 assist families in reducing the level of economic stress and difficulties they face. These issues are addressed by ongoing attempts by family support workers to: encourage families to seek budget advice and to develop financial management skills; encourage families in debt to reduce their debt burden (and particularly hire purchase commitments); assist families to find accommodation and household goods they can afford; and encourage, where applicable, mothers to reduce their welfare dependence and to supplement family income by part-time employment.

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

1.3.3 Caseloads and Extent of Service Provision

Providing adequate family support to high-risk families is labour intensive, and owing to the demands of providing in-depth support, family support workers have a caseload of approximately 15 families. The size of caseloads varies depending on the mix of families. The provision of family support is designed to follow a sequence in which, with increasing family change, the extent of support and assistance reduces. 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 two hours per week and involves weekly home visitation.

- 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 one hour per week for the client and one home visit per fortnight.

- Level 3 – This level of home visitation 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 per week and one home visit per month.

- Level 4 – This level is for families who have become self-reliant and are able to address their problems without support. Families at this level receive a home visit every three months to maintain contact with the programme and to confirm that 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 per week with the family and more than one home visit per week. This level is most commonly used in cases when families first enter the programme and where there is ongoing concern that children are at serious risk of abuse or neglect.

1.3.4 Staff Selection, Training and Supervision

An important feature of family support workers is that these workers do not provide specialist services such as those provided by nurses, social workers, counsellors and similar professionals but rather they act as family mentors and advocates who assist the family in addressing the day-to-day problems it encounters. These job demands require that family support workers 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 that workers have an understanding of the Treaty of Waitangi and an awareness of cultural issues. The Early Start programme also recognises the right of Māori clients to have access to Māori family support workers and also encourages Māori family support

workers to develop linkages with local iwi, hapu and other relevant organisations.

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: evidence of relevant educational background; awareness of cultural issues and obligations under the Treaty of Waitangi; experience in dealing with high-risk families; and evidence of good interpersonal skills and sound judgement.

To provide workers with a general background to their task, Early Start has devised a four-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 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 prove to be very stressful and workers are in need of regular supervision and support to reduce these burdens. Second, regular supervision ensures that the Early Start services are delivered in a uniform way and that workers are clearly advised about the boundaries of their role. To achieve these objectives each family support worker receives two hours' clinical supervision per 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 ample opportunity for workers to discuss particular issues that are of concern to them.

1.4 The Pilot Study

The first phase of the development of Early Start involved a pilot study in which 51 families were enrolled in the programme for a period of 18 months. The findings from this pilot study have been described in a previous report (Fergusson, Horwood, & Grant, 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 infra structure 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." (page 7).

1.5 Background to the Randomised Trial

The work conducted in the pilot study provided the necessary background to developing a randomised trial of Early Start. The development of such a trial became feasible following the announcement of the Family Start Initiative in 1997 by the Minister of Social Welfare. In this announcement, the Minister set aside a budget of $20 million over a three-year period to support the work of home visitation services to be developed in Rotorua, West Auckland, Whangarei and Christchurch. The availability of this funding made it possible for the Early Start consortium to secure funding for a three-year period to provide services for up to 220 families. This funding in turn laid the foundations for a randomised trial in which it was proposed to contrast the outcomes of 220 families receiving Early Start with a

control group of 223 families not receiving this service. In addition, the Christchurch Health and Development Study was successful in securing funding from the Health Research Council of New Zealand to evaluate the results of this trial. The study was ethically reviewed by the Canterbury Ethics Committee and received ethical approval in 1999. The net result of this process was that by mid-1999, the Early Start consortium was in a position to begin a randomised trial of the outcomes of the Early Start programme. A detailed account of the design of this trial is provided in Chapter 3. Briefly, the study involved the following:

- 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 experimental or control groups.

- Those in the Early Start group were offered the Early Start programme whereas 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. All interviews were conducted by trained interviewers employed by the Christchurch Health and Development Study.

- Information gathered by home interview was supplemented by information from GP and hospital records, subject to the signed consent of the family.

The research design was thus one in which 443 families (assigned to experimental and control groups) were studied on five occasions over a three-year period.

1.6 Brief Literature Review

There is now substantial and increasing literature on the design and implementation of family support programmes (see, for example, Chaffin, 2004; Daro & Harding, 1999; Gomby, 1999; Olds, 2002). Two journals have also published special issues of home visitation (*The Future of Children* in 1993 and 1999, and the *Journal of Community Psychology* in 1998). A complete review of this literature is beyond the scope of this report. To set the background for the present study, we confine our review of this literature to an examination of the results of randomised trials of home visitation programmes and to reporting results from New Zealand home visitation programmes. Our task in preparing this review has been aided very considerably by two comprehensive reviews of family support programmes that have appeared in the journal *The Future of Children*. The first, by Olds and Kitzman (1993) examined the state of knowledge in this area in the early 1990s, whereas the second updated this work to the end of the 1990s (Gomby et al, 1999).

1.6.1 Reviews of Home Visitation Programmes

1. Olds and Kitzman (1993)

The first comprehensive review of the research on home visitation was that of Olds and Kitzman (1993). They reviewed only those studies which were methodologically rigorous, in which families were usually randomised to one of two groups: the intervention group (those families receiving the home visitation programme) and a control group (standard care). Outcomes included: promotion of children's cognitive development; enhancement of maternal life-course development; prevention of child abuse and neglect and promotion of children's behavioural functioning and health.

Results were inconsistent across all of the programmes and outcomes including intellectual functioning, parental caregiving, maternal life-course development and rates of child abuse. Some studies found benefits from the programmes in these outcomes while other studies showed no significant differences in outcomes between intervention and control groups. For example, in one trial, those children who were home visited "made much better use of preventive health services (well-child care), had fewer hospitalisations overall, and a lower proportion of cases with severe monilial diaper rash" (page 81). However, "none of the six trials that sought to use home visiting to prevent child abuse and neglect demonstrated overall decreases in maltreatment …" (page 81).

Further, "programs which employ

professionals (especially nurses) and are based on a more comprehensive service model stand a greater chance of influencing qualities of parental caregiving and the child's intellectual functioning than do narrowly focussed programs staffed by paraprofessionals" (page 79). Positive benefits were observed for most trials of intellectual functioning when professionals were employed to provide the intervention. Conversely, when paraprofessionals were employed, five out of seven trials failed to show positive effects for cognitive development.

Olds and Kitzman concluded that "… data on program effectiveness for particular populations range from the spectacular to the disappointing" (page 85); however, "the problems faced by vulnerable families in our society are so immense and the costs of failing to address these problems so great that we cannot wait for a definitive body of research before we begin to take action" (page 89).

2. Gomby, Culross and Behrman (1999)

Gomby and colleagues reviewed six home visitation programmes active during the 1990s across the United States. These programmes included the Nurse Home Visitation Program (NHVP, now known as the Nurse Family Partnership Program); Hawaii's Healthy Start; Parents as Teachers (PAT); the Comprehensive Child Development Program (CCDP); Healthy Families America (HFA) and the Home Instruction Program for Preschool Youngsters (HIPPY). They are among the best studied and each has been rigorously evaluated using randomised controlled trials.

The review showed that none of the programmes reported consistent benefits in health-related outcomes including immunisation rates, well-child visits and dental and medical visits compared to control groups. Changes in children's developmental and behavioural outcomes also showed inconsistent results among all of the programmes; "none found significant effects on all or even a majority of the measures employed, and many revealed no positive effects at all" (page 12). Similar inconsistencies were found for rates of child

abuse, maltreatment and neglect. Further, the authors found wide variability in benefits across models and across programme sites implementing the same model, suggesting that programmes should develop strategies to improve the implementation and quality of services.

In evaluating these results, Gomby and colleagues (1999) concluded that:

> *"In 1993 we cautioned that research findings concerning home visitation programmes were not uniformly positive across outcomes or families and that the magnitude of benefits was modest ... The ensuing six years have brought more research and much more is ongoing ... Such scrutiny suggests that no home visiting model produces impressive or consistent benefits in child development or child health. Several models produce some benefits in parenting and perhaps in the prevention of child abuse and neglect, but only on some of the measures used to assess these outcomes."* (page 24).

The authors went on to state that:

> *"... any new expansion of home visiting programs should be reassessed in light of these findings ... Change is necessary to improve the home visiting services that are currently in place, to adapt existing home visiting models, and to try new service strategies ... The findings indicate that home visiting services are not a silver bullet for all that ails families and children, but then no single program or service can be ... It is up to us to strengthen existing services and craft new approaches to meet the needs of families and children."* (page 24).

3. Recent Research

It is clear from the above reviews that with the passage of time the initially enthusiastic views of home visitation have been modified by experimental evidence to present a far more cautious and less optimistic view of what may be achieved by home visitation. Since the publication of the 1999 review by Gomby and colleagues, a further series of results from the Nurse Family Partnership Program and the Hawaii Healthy Start Program have been published. The findings from these trials only serve to add to the confused picture in this area, with some trials suggesting some positive benefits and others failing to find such benefits.

Ongoing results of the randomised controlled trial of the Nurse Family Partnership Program (Olds, Kitzman et al, 2004) have shown a number of positive outcomes. Children who were visited by nurses during infancy were more likely than those in a comparison group to be enrolled in formal out-of-home care, have higher intellectual functioning and fewer clinical or borderline behavioural problems at a six-year follow-up. Olds et al concluded that their programme was able to improve aspects of children's functioning, leading to better intellectual and behavioural adjustment at the age of school entry.

Olds and colleagues (2004) have also used a three-armed design (controls, nurse home visits, paraprofessional home visits) to examine the effects of professional versus paraprofessional home visitation on mother and child outcomes. Although results showed positive effects on maternal life course, home environments and mother–child interactions, no significant effects of paraprofessional home visits were found for child outcomes. Those families visited by nurses had home environments that were more supportive of children's early learning, and children had better executive functioning, language and intellectual development than controls. The results provided further support for the effectiveness of the programme on child development. Olds and colleagues believe their results do not warrant the continuation of support for a paraprofessional version of home visitation.

Long-term outcomes of the Nurse Family Partnership Program have also been examined (Olds et al, 1998). A 15-year follow-up of the programme showed that by age 15, children who had participated in the programme reported fewer arrests and convictions, fewer sexual partners, fewer behavioural problems due to alcohol and drug use and ran away from home fewer times than their control group counterparts. Mothers who participated in the programme had fewer subsequent pregnancies, received welfare support for fewer months and had fewer arrests and convictions than women in the comparison group (Olds, Eckenrode et al, 1997).

A series of papers by Duggan and colleagues (Duggan, Fuddy et al, 2004; Duggan, McFarlane et al, 2004; Gray, 2001) reported the outcomes of a randomised controlled trial of Hawaii Healthy Start on the prevention of child abuse and neglect in families assessed as being at risk of child abuse. The outcomes were measured by parental self-report (using annual maternal interviews), observations of parenting behaviours, hospitalisation data and official information. No significant differences were found between the Healthy Start and control groups on self-reported abusive behaviours and home visitors were rarely concerned about the possibility of abuse. There were small differences between groups on self-reported measures of neglect; Healthy Start mothers were less likely to report they were "too caught up in their problems to show their child love" (page 611), or they were unable to access needed medical care. There were no differences between groups on other measures of neglect or maternal responsiveness. The authors concluded that there was "little program impact in preventing child abuse".

4. Summary of Research Findings

Table 1.1 provides a comprehensive outline of the findings of the home visitation programmes reviewed by Gomby and colleagues (1999).

a) Findings from two randomised trials of the Nurse Family Partnership Program (Kitzman, Olds et al, 1997; Kitzman et al, 2000; Olds, Henderson, Tatelbaum, & Chamberlin, 1988; Olds, Henderson, & Kitzman, 1994; Olds et al, 1999; Olds, Kitzman et al, 2004) are reported in Table 1.1. Families were followed-up from pregnancy to four years in Elmira, New York, and from pregnancy to six years in Memphis, Tennessee.

b) Findings are reported from the Hawaii Healthy Start Program (Duggan, Fuddy et al, 2004; Duggan et al, 1999; Duggan et al, 2000), which followed 643 families for three years following enrolment and random assignment to either the programme or a control group.

c) Evolving from Hawaii's Healthy Start, Healthy Families America (Daro &

Harding, 1999) spans multiple sites across the United States. Families are followed from birth to age five years. Evaluations of the programme involve multiple designs, therefore only the overall main group findings from the randomised controlled trials are reported in the table.

d) Findings were reported from two randomised trials of the Parents as Teachers Program (Wagner, Cameto, & Gerlach-Downie 1996; Wagner & Clayton, 1999; Wagner, Clayton, Gerlach-Downie, & McElroy, 1997). The programme was provided for up to three years to predominantly Latin American parents in Northern California, and to teen parents in Southern California.

e) Main group findings were reported for the Comprehensive Child Development Program (St Pierre & Layzer, 1999), which followed 4,410 families for up to five years across 21 project sites.

To address the wide variation between programmes in the type of outcomes measured and how the outcomes were measured, outcomes were contrasted between programmes in six general domains: child abuse and neglect; child health; parenting; early childhood education; child behaviour and maternal outcomes. In this table, outcomes are described by a simple scoring system in which + indicates that the programme reported at least one statistically significant benefit for the outcome domain, - indicates that no benefits were found and * indicates that the outcome was not assessed.

The table shows that, with the exception of the Nurse Family Partnership Program, home visitation programmes have failed to show clear or consistent benefits.

Table 1.1 Findings from randomised trials of home visitation

Programme	Child Abuse and Neglect	Child Health	Parenting	Utilisation of Preschool Education	Child Behaviour	Maternal Life Course
Outcomes						
Nurse Family Partnership						
Elmira[1]	+	+	+	*	*	+
Memphis[2]	+	-	+	+	-	+
Hawaii Healthy Start[3]	-	+	-	*	*	-
Healthy Families America						
Overall findings[4]	-	*	+	*	*	+
Parents as Teachers[5]						
Northern California	*	-	-	*	*	-
Southern California	+	+	-	*	*	-
Comprehensive Child Development Program[6]						
Overall findings	*	-	-	*	-	-

[1] Olds et al, 1988; Olds et al, 1994; Olds et al, 1999.
[2] Kitzman, Olds et al, 1997; Kitzman et al, 2000; Olds et al, 1999; Olds, Kitzman et al, 2004.
[3] Duggan, Fuddy et al, 2004; Duggan et al, 1999; Duggan et al, 2000.
[4] Daro & Harding, 1999.
[5] Wagner et al, 1996; Wagner & Clayton, 1999; Wagner et al, 1997.
[6] St Pierre & Layzer, 1999.

1.6.2 Further Research

In a randomised controlled trial in Queensland, Australia, Fraser, Armstrong, Morris and Dadds (2000) investigated the effectiveness of a home visitation programme over a period of 12 months for families at risk of child abuse, neglect or other forms of child maltreatment. The effectiveness of the programme varied across outcomes. No significant differences were found for completed immunisation or use of other community services. Small differences were detected between the intervention and control groups for a number of outcomes, including parenting and child abuse potential during the early stages of the intervention, but these differences had disappeared by 12 months and at an 18-month follow-up. It was concluded that, although there are few longer-term effects of home visitation, home visitation in the immediate postnatal period is likely to demonstrate significant short-term effects. Fraser and colleagues (2000) emphasised the need for the analysis of longer-term outcomes of home visitation.

Bugental et al (2002) added a cognitive retraining component to the home visitation intervention for Latin American families at moderate risk of child abuse and maltreatment. They compared three groups: those families who were randomly assigned to a control condition; those who received a programme modelled on Healthy Start; or those who received home visitation as well as a cognitive component, whereby parents had "repeated experience in finding new ways (directed away from self- or child-blame) of explaining problems and in finding new ways of resolving those problems" (page 247).

The authors found that mothers who received cognitive retraining were significantly less likely to physically abuse their children than those mothers in the home visitation or control groups. There were no significant differences between the home visitation and control groups on this measure. Further, the greatest benefits to child health were gained from cognitive retraining. Children's health was also enhanced in the home visitation group compared to the control group, suggesting "an incremental process in which children's health outcomes are increasingly enhanced as program features are added" (page 255).

1.6.3 New Zealand Research

There has been comparatively little New Zealand research conducted on home visitation programmes but the available evidence has not been consistent or highly promising.

The most rigorous evaluation study was a randomised trial of the Parents as First Teachers (PAFT) programme. This programme is a version of the Parents as Teachers programme described in Table 1.1 above, which is based on the belief that parents are the child's first and most important teachers. Over 25,000 families have been enrolled in the programme since 1994. Pilot projects of the PAFT programme were evaluated using a multi-centre randomised controlled design in which children attending PAFT and a control group were compared on a number of outcome measures. The results of these evaluations were reported in a series of analyses that suggest either no benefits for the PAFT programme (Campbell & Silva, 1997) or only small benefits (Boyd, 1997a, 1997b). These conclusions are generally consistent with recent findings from randomised trials of PAT in the United States (see Table 1.1). A review of the findings by Livingstone (1998) concluded that, while the findings of the trials were "bland", there were a number of deficiencies in the implementation of the programme or the conduct of the evaluations that call the findings of the trials into question.

Farquhar (2003), in a subsequent analysis of the PAFT programme, used a combination of qualitative and quantitative methods to evaluate PAFT. Information was gathered from interviews and surveys with PAFT families, parent educators, programme co-ordinators and community professionals. Contrary to the results found in the randomised trials, Farquhar's analysis found that PAFT had benefits in multiple areas of child and family functioning. These included: parents taking a greater interest in the child's learning; improvement in child safety and standard of care; enhanced child health; strengthened parenting knowledge and practices; utilisation of support services; changes to parent behaviours and lifestyle and parental engagement in further learning. Families also reported a high level of satisfaction with the programme. However, lack of a control group limits the value of the findings considerably.

The Family Start programme has recently been subject to two important evaluations. In 2003, the Evaluation Management Group gathered information contained in a number of reports to provide a process evaluation of the Family Start programme (Evaluation Management Group, 2003). More recently, a final outcome/impact evaluation of Family Start has been published by the Centre for Child and Family Policy Research (2005).

The Family Start programme was set up in 1998 to provide a programme of service delivery to improve the outcomes of children living in at-risk families. Using a "strengths-based" approach, Family Start aimed to improve child health and development, parental life course and family functioning, and parenting skills and practices.

Because of opposition to the use of a randomised trial by service providers, the Evaluation Management Group employed a process evaluation approach aimed at examining the adequacy of the implementation of Family Start. This evaluation identified a number of weaknesses across the programme sites in the ways in which the Family Start programme was implemented. These weaknesses included: governance structures and challenges establishing working relationships with other service providers; differences in standards

and levels of service; significant attrition of client families; difficulties in employing staff with the required range of knowledge and skills and insufficient training of staff; and the timing and appropriateness of the delivery of PAFT (Evaluation Management Group, 2003).

The aims of the outcome/impact evaluation of Family Start were to: determine the short-term outcomes for children and their families participating in Family Start; determine the impact of Family Start on community agencies, and the influence of community agencies on Family Start; and identify ways that Family Start could be altered or improved. Again, because of opposition to the use of a randomised trial by service providers, a single group longitudinal design was implemented.

Information was gathered from the national database of Family Start, interviews with families who were still in the programme and with families who had left the programme, Family Start staff interviews and surveys, and interviews with external agencies. Participant retention in the programme across centres was low (31%–44%). Overall, the results of the evaluation of Family Start showed that "Family Start is moving towards the achievement of its goal to improve outcomes for New Zealand's most at-risk families" (page 105). However it is noted that "the degree to which programme goals are being achieved varies". A number of goals had not been met. For example, breastfeeding and immunisation goals had not been met, the rate of smoking among caregivers did not decrease, referrals to the Department of Child, Youth and Family did not differ between areas with Family Start and areas without Family Start, there was no decrease in welfare dependency across the assessment periods, and there were lower rates of attendance in early childhood education when compared with national figures.

A limitation of the designs used in the evaluations of both PAFT (specifically, Farquhar's evaluation) and Family Start was that both evaluations failed to compare the outcomes of those receiving the programme with an equivalent group of families not receiving the programme. There is increasing

evidence to suggest uncontrolled evaluations of this type tend to over estimate programme benefits (Chaffin, 2004; Gomby, 1999; Gomby et al, 1999). For example, in their review of the CCDP programme Gomby et al noted that:

> "In the CCDP study ... the group that received program services improved in a number of domains from the baseline, but so did the no-treatment control group – and by an equivalent amount. Without a control group, the program might have appeared effective when it was not" (page 19).

In summary, existing evaluations of home visitation programmes in New Zealand have produced somewhat inconclusive results. The evaluation of PAFT using a randomised trial approach suggested few benefits for this programme, while a further evaluation using a less rigorous descriptive approach to evaluation that did not include an adequate comparison group suggested possible benefits. The evaluations of the Family Start programme suggested difficulties in implementation, which made adequate programme evaluation difficult, and revealed that many goals of the programme had not been achieved. In many respects these findings mirror the results obtained in US-based programmes where: studies using randomised trials have often failed to show benefits for home visitation programmes (Gomby et al, 1999) and studies using less rigorous designs have often concluded that such programmes are beneficial despite the lack of evidence from randomised trials (Chaffin, 2004) and the omission of control groups (Gomby, 1999).

1.6.4   Summary

Collectively, the evidence on the contribution of current family support programmes suggests that for many programmes the initial promise of home visitation as a mechanism for improving the opportunities of children from families facing stress and disadvantage has not been fulfilled. While some programmes, most notably the Nurse Home Visitation Program (Olds et al, 1988; Olds et al, 1999; Olds, Kitzman et al, 2004), have reported moderate success, others have not. In summary, the current state of the evidence from home

visitation programmes shows that:

- Overall, the results of the impact of home visitation are mixed and inconclusive.
- Despite the failure of many programmes to produce any positive effects, other programmes have done so.
- There is an urgent need to identify the factors that make programmes work and to develop home visitation programmes that are well designed and well implemented to produce positive outcomes for at-risk children and families.

The inconsistent findings from the literature on home visitation are likely to reflect a variety of factors that conspire to both reduce programme efficacy and increase variability between programme findings. These factors include:

1. *Variations in programme content:* While all home visitation programmes share the common feature that the programme is delivered at home, there have been wide variations in programme content including: the theoretical assumptions of the programme; the goals of the programme; the methods by which the programme is delivered; the types of staff used to deliver the programme and the duration of the programme. Given these variations, the extent to which reviews of findings compare "like with like" is open to question and the inconsistent conclusion emerging from reviews may reflect the fact some programmes are better designed and more effective than others.

2. *Variations in client need:* A feature common to many home visitation

programmes is the attempt to address a wide array of client needs among families facing stress and difficulty. The diversity of client needs means that the objectives of programmes are broad and that often positive change may be confined to specific client sub-groups facing some problem (for example, changes in risks of maternal depression will be confined to those women who are at risk of depression by virtue of their personal characteristics and social environment). The implication of this is that between-group differences between experimental and control groups may often be small since not all clients are "at risk" of all outcomes. Given this, one might expect to find a pattern of small between-group differences between experimental and control groups. This requires the use of research designs that are sufficiently sensitive to detect small but consistent between-group differences.

3. *Programme involvement and compliance:* A further major threat to study validity comes from between-study variations in the extent of programme involvement and programme compliance. In particular, many home visitation programmes involve long-term involvement with families for periods up to five years. Inevitably, lengthy programme duration will mean that a proportion of families fail to receive the full programme as a result of a series of processes including: withdrawal from programmes; moving away from the area in which the programme is provided and

changes in family structure. The net result is that these factors in combination may lead to high attrition from programmes resulting in reduced programme efficacy (Gomby et al, 1999).

4. *Programme fidelity:* A further feature that may influence the success of family support programmes concerns the extent to which home visitors adequately deliver the programme of service. There is evidence from some studies that the practices of home visitors often depart substantially from programme principles and recommendations (Gomby et al, 1999). Such findings clearly suggest the importance of ensuring fidelity in the process of programme delivery.

5. *The experiences of the control group:* A further factor that has to be taken into account is the extent of exposure of members of the control group to alternative services. In particular, those in the control group are not strictly an "untreated" group but rather represent a group which may have access to services that may offer similar benefits to the programme under examination. Thus, the extent to which family support programmes show benefits may depend on the social context within which they are evaluated and in general these programmes are most likely to show benefits in communities with a limited array of childhood and family services and less likely to be found to be beneficial in communities with well-developed services.

The above considerations suggest that demonstrating the efficacy of family support programmes is by no means a simple task and the evaluation process needs to take into account: the nature and goals of the programme; the extent of client needs; the fidelity of programme delivery; the level of involvement in and compliance to the programme; and the social context within which the programme is delivered. Each of these factors is likely to combine to determine levels of programme success.

CHAPTER 2:
BACKGROUND TO THE
EARLY START PROGRAMME
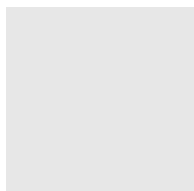
2.1 Introduction

This chapter sets the background to the research described in this report by presenting details on the Early Start service. The chapter provides a detailed description of the Early Start service, the rationale and processes of client recruitment, and the principles of service delivery. A detailed case history is provided to illustrate the process of service provision to a family enrolled in Early Start.

2.2 The Early Start Service

The Early Start service described in this report may be thought of as comprising two linked processes. The first is the client recruitment process by which clients are identified and enrolled in Early Start. The second is the service delivery process by which the clients enrolled in the programme are provided with services. These components of the programme are described below.

2.2.1 Client Recruitment: Rationale and Process

At the inception of the project, the Early Start consortium was aware of the need for the client recruitment strategy to address two goals that were, to some extent, in conflict. The first goal was to obtain population coverage of families facing stress and difficulty. The second goal was to ensure the client recruitment processes avoided labelling or stigmatisation as a result of population screening. After consideration of these issues, the strategy depicted in Figure 2.1 was developed. This strategy combines

methods of population screening and needs assessment to develop a client identification process in which the client has multiple opportunities to both learn about the Early Start service and to accept or decline this service.

The first stage of the client identification processes involved Plunket nurses throughout the Christchurch region administering a simple 11-point screening assessment based on the questionnaire used in the Hawaii Healthy Start Program for the identification of families facing stress and difficulty. This questionnaire contained items relating to maternal age, extent of family support, wantedness of pregnancy, substance use, family violence and child abuse risk. This questionnaire was administered to all new clients enrolled with the Plunket Society over the period from 1 January 2000 to 31 July 2001. To protect family confidentiality, nurses were asked to make this assessment informally and no questionnaires were completed. Nurses were asked to refer any family where two or more items were present. In addition, to ensure the screening process was comprehensive, nurses were asked to refer any family where they thought that the family could benefit from the Early Start programme.

This stage of the development of the programme involved several sessions in which Plunket nurses were briefed on the principles of the screening programme and ethical issues associated with the screening process were discussed. Plunket nurses were chosen as the source of client referral since at the time of the client recruitment process these nurses saw about 95% of women giving birth in Christchurch within three months of the birth. Thus, screening by Plunket nurses provided a close to universal coverage of the Christchurch population.

For those families who were identified as possible clients for the Early Start programme, the Plunket nurse provided information on both Early Start and the proposed outcome evaluation. For those mothers who agreed to participate in the programme, the Plunket nurse obtained a signed consent form indicating that the family was willing to enter the randomised

trial. The consent form was forwarded to Early Start who then contacted the family. The Plunket nurse provided a weekly return to Early Start indicating the number of families declining the trial.

To preserve client privacy, an information "firewall" was set up so that the only information the Plunket Society supplied to Early Start was consent forms for those families who consented to entering the trial and statistical returns of the number of families declining the offer to enter the trial. In this way, the client identification process provided an opportunity for the Plunket Society to refer clients to Early Start on the basis of signed and informed consent while at the same time protecting the privacy of those families who declined to enter the trial and those not eligible for the trial.

Figure 2.1 Client recruitment strategy

Initial assessment by Plunket nurses	▶	Not eligible
Eligible	▶	Declined
Consent sent to Early Start		
Family offered initial service	▶	Declined
Family needs assessment	▶	Not eligible
Offer of full service		Family offered three-monthly contact (Level 4)
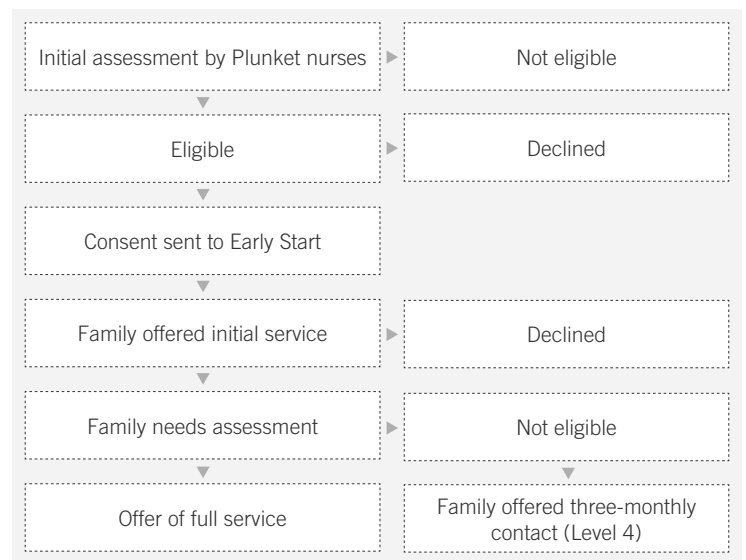
The second stage of the client enrolment process involved the provision of a one-month introductory period for each client referred to Early Start. It was recognised that the screening process employed by the Plunket nurses was likely to be only approximate and would involve "false positive" instances in which families not in need of Early Start had been identified. In addition, the Board of Early Start felt there was a need for a thorough process of client needs assessment to form the foundations of programme delivery. To address these issues, the second stage of the recruitment process

involved enrolling clients for a one-month introductory period. This offered clients a period of time in which they could learn about the Early Start programme without making a long-term commitment to the programme. This period also provided the Early Start programme with an opportunity to conduct an in-depth needs assessment to determine whether the family would benefit from Early Start.

The third stage of this process involved an in-depth evaluation of the client's need for the Early Start service. This evaluation was based on a modified version of the Kempe Family Stress Checklist, which is used in the Hawaii Healthy Start Program (Duggan et al, 2000; Kempe, 1976). Any family scoring 25 or more on this questionnaire was eligible for the Early Start programme. This cut-off point was set using the criterion developed in Healthy Start.

The final stage of the client enrolment process involved an offer of service to the eligible clients. Subject to client agreement and signed consent, eligible clients then entered the service. Families not meeting eligibility for the full Early Start service were not abandoned but were offered Level 4 service involving a telephone call every three months. The process of client recruitment thus aimed to overcome issues relating to screening and informed consent by using a multi-stage process in which the aim was to acquaint clients with the nature of the Early Start programme and to seek their informed consent to participate in this programme.

2.2.2 Principles of Service Delivery

Following entry into the Early Start programme, clients were offered an intensive family support programme lasting up to five years. The key features of programme design, content and delivery are described below.

1. Programme Philosophy

Over the history of Early Start a clear philosophy of programme delivery has evolved. This philosophy centres around building a collaborative partnership between the family support worker and client with the aim of this partnership being to maximise the health and well-being of the client child

and family. While Early Start treats the referred child as the primary client, it recognises that the well-being of the child can only be achieved through the more general health and well-being of the family. The Early Start programme does not use an exclusively strengths-based model of intervention because the organisation holds concerns about the limitations of strengths-based approaches in the identification and management of family deficits in such areas as child abuse and neglect. The underlying philosophy of the Early Start programme focuses on developing *problem solving partnerships* between family support workers and parents, with these partnerships working on building strengths and eliminating deficits to ensure the health and well-being of children and their families.

2. Methods of Service Delivery

a) Staffing: The service was delivered by home visitation provided by family support workers (FSWs) who had tertiary level qualifications in nursing, social work or a related discipline. All staff were recruited for the project by a panel of assessors convened by the Board of Early Start, with this panel always consisting of the manager of Early Start (Mrs H Grant) and at least one Māori member of the Board. Staff were provided with a five-week training programme to familiarise them with the principles of the programme.

b) Caseloads: The aim of Early Start was to provide FSWs with small manageable caseloads of families to ensure their ability to deliver intensive service. Typically, the caseloads for full-time staff contained between 10 and 20 families depending on the level of family need and progression through the programme. The caseload size was determined by a caseload point formula that weighted families according to the extent of family need, with each full-time worker being assigned 15 caseload points. This number of caseload points corresponds to 15 families of "average" levels of need.

c) Programme content: The Early Start programme aimed to produce positive changes in a wide range of outcomes spanning child and family health and well-being. These areas included:

- *Improvement of child health:* Here the focus of the programme was on achieving a series of specific goals that were likely to maximise child health outcomes. These goals included:
  - Timely medical visits for common childhood morbidity.
  - High levels of compliance to immunisation and well-child care checks.
  - Reductions in hospital visits for preventable childhood morbidity including childhood accidents and poisoning.
  - Improvements in home safety and home environment.
- *Reduction of child abuse:* Here the focus was to achieve a series of goals that were aimed at reducing child abuse and neglect. These goals included:
  - Reduced agency contact for child abuse and neglect.
  - Reduced use of physical punishment by parents.
  - Increased awareness by parents of child abuse and neglect issues.
  - Effective use of the Child, Youth and Family Service in cases where clear abuse and neglect issues were evident.
- *Improvements in parenting skills:* A theme that pervaded the delivery of Early Start was a focus on improving parenting skills in the areas of:
  - Parental sensitivity.
  - Positive parenting.
  - Non-punitive parenting.
- *Supporting parental physical and mental health:* Early Start recognises that the extent to which parents can operate effectively will depend on the extent to which issues relating to parental physical and mental health are addressed. Key goals of the programme were:
  - Reduction in the rate of unplanned and/or unwanted pregnancy.
  - Early detection and treatment of depression.

- Assistance with management of mental health and substance use disorders.
- Encouragement to use general practitioner services.
- *Encouraging family economic and material well-being:* Since many participants in the Early Start programme were single parents, they were often subject to a series of economic limitations that spanned welfare dependence, economic deprivation and financial hardship. These difficulties in turn were likely to exacerbate other issues facing families. To address this issue the goals of Early Start spanned the following areas:
  - Reducing levels of welfare dependence.
  - Encouraging the use of budgeting and budgeting services.
  - Encouraging workforce participation.
  - Encouraging forward economic planning and saving.
- *Encouraging stable positive partnerships:* Pilot research into families in Early Start suggested the presence of relatively high levels of inter-partner conflict, inter-partner violence, and partner separation. Since family change and violence were factors that were likely to have a pervasive influence on other aspects of childhood functioning, a further goal of Early Start was to reduce tensions relating to partner relationships including:
  - Reduction of partner violence and partner conflict.
  - Improvements in partner relationships.
- d) Programme delivery: The delivery of the Early Start programme centred around a process of home visitation aimed at achieving the goals described above. The key components of this process included:
  - *Client assessment:* For the purposes of programme delivery, clients were classified into four levels with each

level reflecting the needs of the client family. Each level was provided with home visitation services that reflected the level of need of the family. The levels were:

- Level 1. High need: One–two hours home visitation per week.
- Level 2. Moderate need: Up to one-hour home visitation per fortnight.
- Level 3. Low need: Up to one-hour home visitation per month.
- Level 4. Graduate: Up to one-hour contact (phone/home visitation) per three months.

All families entered the programme at Level 1 and, with the passage of time, changed levels depending on progress. Changes in levels were decided by discussion between the FSW and the relevant supervisor.

- *Forward and individual family plans:* Families entering Early Start faced a range of issues and difficulties. For this reason it was important for Early Start to develop a programme of service delivery that took into account the family's specific needs and circumstances. To address this issue, two family plans were devised for each client family. The first – the Family Support Plan (FSP) – was prepared by the FSW in conjunction with the relevant Early Start supervisor and set out a plan of work for the next three-month period. This plan addressed both issues that are common to all families (eg child health, parenting) and issues that were specific to a given family (eg child abuse and neglect, family violence). The FSP described the issues to be addressed and the steps needed to form a partnership with the family to address these issues. The second family plan – the Individual Family Plan (IFP) – was prepared by the family in conjunction with the FSW and set out the goals the family would like to achieve over the next three months. The IFP thus provided the client family with a voice in the

planning of service delivery.

- *The collaborative approach:* To address the issues covered in the family plans, FSWs used a collaborative problem-solving model based around the following general principles:
  - Understanding the client's individual and cultural perspectives.
  - Actively involving clients in the service by sharing ideas and experiences and involving clients in problem solving.
  - Assisting clients to seek and generate their own solutions.
  - Providing support and assistance for clients to implement their solutions.
  - Teaching, mentoring and providing the client with alternative strategies and solutions where these were needed.
  - Acting as an interpreter for the client in dealing with new material, ideas or suggestions.

  These principles were employed flexibly by FSWs to address the key issues raised in the two family plans.

In addition to working on the specific issues contained within the family plans, Early Start also provided a broad programme of family support in the areas of child health, parenting, child abuse and neglect, parental health, family violence, and family economic well-being. This role involved monitoring various areas of child and family functioning and providing advice, support and assistance as problems and issues arose.

A limitation of the original version of the Early Start programme was that it did not contain a systematic parenting programme that was delivered to all parents. Rather, programme delivery in this area relied on the skills and abilities of FSWs to teach and model parenting skills. This limitation has recently been addressed by including the Triple P: Positive Parenting Program (Sanders, 1999) as part of the service. This programme involves parents in

10 one-hour weekly sessions designed to introduce parents to the principles of effective parenting.

A final component of the services delivered by Early Start was the development of effective liaison with other organisations and services in the Christchurch region. Families enrolled in the programme presented with a broad range of health, social, emotional, financial and related issues. Much of the work of the FSWs involved identifying these issues and assisting the client family in the effective use of local services. This process was underwritten by FSWs acquiring a wide knowledge of local services and a good working liaison with these services.

e) Cultural issues: The aims of the Early Start programme were to provide a "mainstream" service that was available to all ethnic and cultural groups. Under article three of the Treaty of Waitangi, Early Start was obliged to make this service available to Māori, and the general requirements of the Treaty of Waitangi implied that the service was delivered to Māori in a culturally appropriate and respectful manner. These obligations under the Treaty of Waitangi were addressed in a number of ways including:

• Initial hui with local Māori seeking approval of the suitability of the Early Start programme for Māori.

• The appointment of Māori to the Board of Early Start, with the original Board containing three Māori directors.

• Recruitment of Māori staff to work with Māori clients, with three of the original team of 12 workers being Māori.

• Marae-based cultural training provided by Māori Board members and Kaumātua.

The aim of this process of consultation, recruitment and staff training was to produce an organisational climate in which the Early Start programme could be delivered to Māori in a culturally appropriate way while maintaining the underlying philosophy and principles of the programme. It has been the practice of Early Start that, wherever possible,

Māori families should be offered family support by Māori workers.

f) Ensuring programme fidelity: At the inception of Early Start it was recognised that a major challenge faced by the service was that of ensuring fidelity of programme delivery so that all staff were delivering a similar programme of support in a similar way. To encourage fidelity of programme delivery, a number of features were built into the organisation and administration of the Early Start programme. The most important of these was the employment of dedicated programme supervisors who had the full-time task of monitoring and supervising each worker's programme of work. This was achieved largely through the mechanism of weekly supervisory meetings with each FSW. During these meetings, the FSW's work from the preceding week would be reviewed and issues relating to the implementation of family plans would be examined. These meetings also provided supervisors with an opportunity to review the safety of each worker's practice and to monitor the health and well-being of the FSW.

A further mechanism by which fidelity to the programme was encouraged was through the development of organisational databases that recorded levels of client achievement in key areas relating to preventive health, home safety, child abuse and neglect, and participation in parenting programmes. These databases provided the organisation with continuous monitoring of key outputs. This monitoring role was reinforced by the need to provide the funder with regular three-monthly reports on client outcomes. Finally, to ensure the organisation developed a unified approach and culture, regular weekly meetings were held at which all staff could participate in discussion of issues facing the organisation.

2.3 Illustrative Case History

To illustrate the process of service provision, an illustrative case history is provided. This case history describes the services provided

to an actual family enrolled in Early Start. To protect family privacy, names have been changed. The family involved has seen the case history material and is agreeable to the publication of the details provided.

Emma, Stuart and Aroha

The family: At the time of enrolment, Emma was a 39-year-old Māori woman living with her partner Stuart, who was Pākehā, aged 40. Their child, Aroha, was six weeks old. Emma had three other children who were in the care of the Child, Youth and Family Service (CYFS) and both parents had a history of opiate addiction.

Family issues at the assessment period: To provide an assessment of the family needs, Emma and her family were enrolled in Early Start for a four-week assessment period. The purpose of this period was to acquaint the family with the Early Start programme and to allow the family support worker time to assess both family strengths and challenges. This initial assessment identified a series of challenges facing the family. These included:

- Difficulties in managing Aroha who was a restless, low-birthweight infant experiencing methadone withdrawal (Emma had been on methadone throughout her pregnancy).
- Poor housing conditions and a lack of furniture.
- Low income and family debt.
- Drug abuse (Stuart was using heroin and other drugs, while Emma was on methadone and using cannabis regularly).
- Maternal depression.
- Limited social support.
- Relationship difficulties.

In addition to these challenges, the family was suspicious, difficult to engage and evasive in their initial dealing with the FSW.

Nonetheless, there was evidence of some strengths and resiliency in the family situation:

- Emma was determined to discontinue her previous way of life and commit herself to providing Aroha with adequate care. She was also motivated to restore contact with her children who were in the care of CYFS.

- Emma was a good communicator who was able to articulate her goals and aspirations.
- Emma was compliant with the methadone programme and had ceased using heroin although she admitted to using cannabis.

The family plans: At the end of the assessment period, two family plans were developed. The first was the Individual Family Plan (IFP) developed by Emma and Stuart. This plan described the family's goals and aspirations for the future. The IFP developed by Emma and Stuart focussed on a series of goals which included establishing a warm and comfortable home environment, learning about being a good parent, having regular access with the children currently in care and remaining compliant to their methadone treatment plan.

The second family plan was the Family Support Plan (FSP) developed by the family support worker in conjunction with the Early Start supervisor. This plan focussed on a series of goals relating to child and family well-being and concentrated on addressing current care and protection concerns, addressing the poor housing conditions, teaching and establishing consistent and appropriate baby care routines, initiating child and home-safety measures, seeking assessment and treatment for possible depression, liaising with the methadone clinic regarding current drug treatment regimes, ensuring contraception advice and establishing regular Plunket nurse follow-up.

Following the development of the IFP and FSP these plans were discussed with Emma and Stuart and their signed consent to participate in the Early Start programme was obtained.

Service delivery during the first six months: During the first six months, the family was visited an average of three times per week by the FSW. Each visit lasted about 1.5 hours. This level of visiting was justified by the high level of need within the family and the concerns of CYFS about possible care and protection issues. Home visitation focussed upon a number of key themes which spanned issues in the IFP and FSP. These issues were:

- Child health, including health care, immunisation, home safety and infant feeding.
- Family physical and economic environment including housing, furnishing and budgeting.
- Providing Emma and Stuart with parenting advice and support.
- Child care and protection issues (ie ensuring that Aroha was being reared in an adequate, safe home environment).
- Assisting Emma to manage depression.
- Assisting the family with reliable contraception.

Progress at six months: At the end of the first six months the growing partnership between the family support worker and the family had achieved the following goals:

1. Child health: The family had become enrolled with a single general practitioner who addressed the family's health needs. Aroha had made regular visits to this GP and was up to date with her immunisations and well-child checks. Regular visits had been made to the Plunket nurse. With the FSW's support, Emma had instituted a number of home-safety practices which included ensuring Aroha lived in a smoke-free zone, installation of smoke detectors and a fire safety plan and ensuring Aroha slept in a non-prone position. Aroha's withdrawal from methadone was complete and she had begun to thrive, gain weight and sleep better.

2. Maternal health: Emma had been seen by her GP who monitored her depression. Although no medication was provided, Emma's symptoms of depression had alleviated. Emma had also arranged a reliable method of contraception. However, she continued to use cannabis on a regular basis.

3. Family economic circumstances: The family had been referred to the City Mission for budgeting advice and as a result Emma was able to manage on a very tight budget with some assistance from the food bank.

4. Crisis support: At the beginning of the six-month period, Emma called her FSW

three to four times a week regarding family crises. As the family circumstances began to improve, these calls reduced in frequency to about one per week.

Service delivery 6–12 months: With improvements in the family situation, home visitation was reduced to once per week. During this period, service delivery built on the foundations laid in the first six months and concentrated on the main issues addressed in the first six months (child health, parenting, maternal health, family economic circumstances and crisis support) but addressed some further issues that emerged in this period. These issues included:

- Family violence: Emma and Stuart's relationship had deteriorated, resulting in arguments and family violence that threatened both Emma and Aroha's safety and well-being.
- Drug abuse: While Emma was generally compliant with her methadone regime, she had continued to use cannabis on a regular basis. A further goal during this period was to eliminate Emma's cannabis use.

Progress made by 12 months:

1. Child health: As a result of a change of residence, the family had changed their GP. The family maintained regular contact with this GP and Aroha was up to date with all immunisations and well-child checks. She also continued to see the Plunket nurse. Emma had provided a safe home environment which included the use of safety plugs, safe storage of

household poisons, use of safety gates for the stairs, installation of smoke alarms and a fire safety plan.

2. Maternal health: Emma had no recurrence of her depression and had maintained her contraceptive regime. Her usage of cannabis had reduced and no further drugs were used.

3. Family material circumstances: Emma had moved to a state house and with the assistance of her FSW had created a warm and comfortable home for herself and Aroha. She was able to manage on her budget.

4. Parenting: Throughout this six-month period, the FSW had worked with Emma to develop a Child Profile Book that provided a record of Aroha's development. The FSW used this book as a method of introducing Emma to issues in child development and parenting.

5. Family circumstances: Emma and Stuart had separated and after a difficult period had reached an agreement for Stuart to have supervised access with Aroha. As a result of improvements in home circumstances and family functioning, the Department of Child, Youth and Family Services closed their file on the family with respect to Aroha. Emma's older children remained in the care of the department.

6. Crisis support: With the improvements in the family circumstances there was a progressive reduction in contact for family crises.

Service delivery 12–24 months: During this period service delivery concentrated on the same major themes as for the first 12 months (child health, maternal health, parenting, family material circumstances, family issues and crisis support) with a number of further themes developed by Emma for her IFP. These themes were:

• Emma wished to undertake major life changes and leave her past criminal and drug-addicted lifestyle behind.

• Emma wished to open a bank account.

• Emma wished to re-establish contact with her children.

In addition, the Early Start plan included

further goals to ensure Aroha was enrolled in early childhood education and with the preschool dental service.

Progress at 24 months:

1. Child health: The family continued their contact with the GP and Aroha was up to date with all immunisations and well-child checks. She continued her contact with Plunket and was enrolled with the preschool dental service. As a result of her exposure to methadone during pregnancy, Aroha required extensive dental treatment and specialist care.

2. Maternal health: Emma had no recurrence of depression, maintained her contraceptive regime and had markedly reduced cannabis intake from very frequent usage to only using cannabis very occasionally. Emma also engaged in anger management training so she could avoid confrontations with authority figures.

3. Parenting: The FSW continued her use of the Child Profile Book to introduce Emma to issues in child development and to keep a record of Aroha's development. She encouraged Emma to both play with and read to Aroha every day.

4. Material circumstances: With the assistance of the FSW and the Christchurch City Mission budget service, Emma was able to open a bank account. Her home environment was now tidy and well-ordered and Emma was able to live within her budget.

5. Early childhood education: With assistance from her FSW, Emma enrolled Aroha in preschool for nine hours per week and throughout the next six months Aroha did not miss a single session.

6. Family Issues: Emma had been able to re-establish contact with two of her older children who visited on a regular basis.

Service delivery 24–36 months: With the improvements in the family situation, home visitation was reduced to one visit per fortnight. Over the next year, the FSW continued to visit to work on issues in the family plans. These themes covered the areas described earlier but included some further issues relating to the family's changing circumstances. These issues included:

- Managing Aroha's dental problems. (As noted above Aroha faced severe dental problems that required she had a diet of soft and mashed foods.)

- Increasing Aroha's range of experiences.

- Managing behaviour problems.

- Initiating community participation.

Progress at 36 months:

1. Child health: At three years, Aroha is a healthy child who has had regular contact with her family doctor and Plunket nurse. All her immunisations and well-child checks are up to date. Although Aroha has faced severe dental problems, these are resolving and Emma has been able to provide her with an appropriate diet.

2. Parenting: Emma has continued to invest in developing her parenting skills. She reads to Aroha every day and has enrolled her in the local library. With the assistance of her FSW, she has enrolled Aroha in swimming lessons. Following concerns about behaviours Aroha learned at preschool, Emma has learned about non-punitive parenting methods and is enrolled (with her new partner) in the Triple P parenting course provided by Early Start. Aroha is now enrolled in kindergarten, attends every afternoon and Emma is actively participating as a parent helper.

3. Family material circumstances: Emma has formed a new and supportive partnership in which income and debt issues are being resolved. Her home is tidy and well run.

4. Maternal health: Emma has no recurrence of depression, is using contraception reliably and has remained fully compliant with the methadone treatment plan. She no longer uses cannabis.

5. Family issues: Emma has formed a supportive relationship with her new partner and also sees her older children regularly.

At 36 months the family was assigned to monthly home visitation.

The above case history serves to illustrate a number of the key features of the Early Start programme. These features are:

1. *Long-Term Commitment to Building Partnerships with Parents:* Over the 36 months that Emma and her family were enrolled in Early Start they were seen by their family support worker on over 100 visits. These visits laid the foundations for the development of a partnership between the FSW and the family with this relationship developing from an initial suspicion and evasion to later warmth and trust. This can be seen by tracing the number of times Emma kept to scheduled appointments and not evading contact by missing appointments or cancelling at the very last minute. During the first six months, she missed 12 appointments by not being at home when the FSW arrived for a scheduled visit, which dropped to two during the next six-month period and for the following two years all scheduled appointments were kept. If Emma was unable to keep an appointment it was always for a valid reason and she cancelled in plenty of time and rescheduled and then attended that appointment.

2. *Balance Between Needs and Strengths:* As may be seen from the example, the programme involved both recognition of family needs and family strengths with the programme working with the family to develop strengths to face the multiple and serious needs of the family.

3. *Balance Between Service and Family Goals:* A further area in which the service delivery attempted to strike a balance was between the specific goals of the service (eg timely use of medical services) and family goals (eg Emma's desire to be re-united with her other children). The aims of the service delivery process were to develop a programme that recognised the importance of both types of goals and to form a partnership that achieved these goals.

4. *Flexibility of the Programme:* Although the Early Start service focuses on positive outcomes in a number of areas of childhood and family functioning, the programme of service delivery provided is tailored to child and family circumstances as they change with time.

CHAPTER 3:
The Justification and Planning of the Randomised Trial

This chapter provides the justification of the randomised trial and an overview of the research design employed including 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 was of the view the most compelling evidence for programme effectiveness would come from a randomised trial in which a group of children and families receiving the Early Start programme were 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.

There have been a number of critiques of the use of randomised trials in the evaluation literature which have suggested that such trials have a number of limitations (see, for example, Donaldson & Christie, 2004). Arguments against randomised trials suggest that: randomised trials are not always the best methods of determining causality, particularly because causality is complex and there are multiple influences on outcomes; randomised trials can be ethically unacceptable in many situations; data sources can sometimes be insufficient to use randomised trials; and alternative methods may be equally valid and equally effective (Donaldson & Christie, 2004). Despite such claims, the fact remains that, in the mainstream literature on family support programmes, results from randomised trials have served as the "gold

standard" for assessing programme efficacy and effectiveness. The reason for an emphasis on randomised trials as opposed to other methods of evaluation has been clearly articulated by Chaffin (2004) who comments on the limitations of non-randomised evaluations:

> *"When it comes to establishing safety and effectiveness there is simply no substitute for randomized trials"* (page 589).

> *"Relying on … largely non-randomized designs … can lead to serious errors. Across a range of intervention and prevention topic areas, non-randomized designs are particularly vulnerable to over-estimating the size of intervention effects …"* (page 591).

This is also well reflected in the major reviews reported in *The Future of Children* (Gomby et al, 1999; Olds & Kitzman, 1993) which have assessed the efficacy and effectiveness of these programmes almost exclusively on the basis of results from randomised trials. Olds and Kitzman note that "randomized trials … when adequately designed and conducted, produce substantially better estimates of program effects than do estimates derived from other types of research". Furthermore, "by comparing home-visited and control groups … the investigator can determine with a degree of statistical confidence the extent to which the differences observed between program and control groups are due to chance" (pages 54–55).

For these reasons, the Board of Early Start elected to invest in the development of a randomised trial of the Early Start service. The overall research design employed was largely dictated by the funding provided to Early Start under the original Family Start 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. 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 ethical approval for each phase of the trial.

3.2 Overview of the Research Design

3.2.1 Recruitment of Clients and Randomisation to Groups

Clients for the trial were recruited over a 19-month period from 1 January 2000 to 31 July 2001. Over this period, a total of 443 families were referred to Early Start as being eligible for the randomised trial. At the point of client identification, clients were randomised to either the Early Start or control group using a computer-generated series of random numbers, with 220 clients being assigned to the Early Start group and 223 to the control group. The clients in the Early Start group were offered the services described previously, whereas the control group were not offered the Early Start service but had access to the full range of other health, education, welfare, and related services available in Christchurch. (In Chapter 4 we examine the extent to which this method of group assignment led to comparable groups.)

3.2.2 Assessments

1. *Family interviews:* Clients in both groups were assessed at the point of trial enrolment and at six, 12, 24 and 36 months post-enrolment using a home-based interview conducted by an interviewer (Mrs R Deighton) employed by the CHDS. These interviews were conducted with the parent with the greatest involvement with the child (usually the child's mother) and lasted between one to two hours. Table 3.1 gives a summary and overview of the issues covered in each interview. 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; Mrs E Ridder) 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 requested to provide signed consent for the research group to 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 obtain verification of the child's immunisation status and history of well-child care. In addition, case note information was obtained for all hospital attendances/admissions.

3. *Assessment of intelligence*: Following the 36-month assessment, families were visited by trained psychometric testing staff who administered the Wechsler Preschool and Primary Scale of Intelligence (WPPSI-III, Wechsler, 2002). The WPPSI is a test designed to assess the intelligence of preschool children and provides three scores: a verbal intelligence score; a performance intelligence score; and an overall intelligence score.

3.2.3 Variables for Analysis

From the assessments described above, a series of variables was constructed to assess the outcomes of the Early Start programme at six, 12, 24 and 36 months post-enrolment. These variables spanned measures of: child health; child abuse and neglect; parenting; parental health; family social and economic well-being; and family stability, family relationships and family violence. The Appendix provides an account of the methods by which these variables were constructed and, where information is available, on reliability and validity.

3.2.4 Statistical Issues

The analysis of the present trial poses some technical issues since, as described in Chapter 4, there was considerable variation in the amount of service received by those in the Early Start group ranging from those who received little or no service to those who received services for at least three years. This variation in service provision poses issues for the evaluation since not all individuals were

Table 3.1 Summary of topics covered in each interview

6 months	12 months	24 months	36 months
Family Composition	Family Composition	Family Composition	Family Composition
Infant Feeding	Child Development	Child Development	Child Development
	Child Behaviour	Child Behaviour	Child Behaviour
Child Health	Child Health	Child Health	Child Health
Safety and Protection	Safety and Protection	Safety and Protection	Safety and Protection
	Early Education and Child Care	Early Education and Child Care	Early Education and Child Care
Maternal Health	Maternal Health	Maternal Health	Maternal Health
Maternal Depression	Maternal Depression	Maternal Depression	Maternal Depression
Cigarette Smoking	Cigarette Smoking	Cigarette Smoking	Cigarette Smoking
Alcohol Use (maternal, paternal)	Alcohol Use (maternal, paternal)	Alcohol Use (maternal, paternal)	Alcohol Use (maternal, paternal)
Other Substance Use (maternal, paternal)	Other Substance Use (maternal, paternal)	Other Substance Use (maternal, paternal)	Other Substance Use (maternal, paternal)
Treatment Seeking	Treatment Seeking	Treatment Seeking	Treatment Seeking
Criminal Offending (maternal, paternal)	Criminal Offending (maternal, paternal)	Legal Problems (maternal, paternal)	Legal Problems (maternal, paternal)
Partner Relationships	Partner Relationships	Partner Relationships	Partner Relationships
Family Finances	Family Finances	Family Finances	Family Finances
Family Planning	Family Planning	Family Planning	Family Planning
Life Events	Life Events	Life Events	Life Events
Child Discipline	Child Discipline	Child Discipline	Child Discipline
Parenting	Parenting	Parenting	Parenting
Cultural Participation	Service Utilisation	Service Utilisation	Service Utilisation
			Satisfaction with Early Start

exposed to the same amount of "treatment". The recognised method for treating this difficulty is to employ an "intention to treat" paradigm in which contrasts are made between the control group and all of those enrolled in the treated group, irrespective of the extent of treatment received. The justification for the intention to treat method is that this method maintains the randomisation to groups and thus preserves experimental validity (see, for example, Ellenberg, 1996). What the intention to treat approach assesses is the extent to which the entire programme, including the method of client retention, had benefits when compared with the control group. In cases where there is high attrition from the programme, the intention to treat paradigm may give a conservative estimate of the benefits to those who receive the full programme. Alternative approaches to this problem are to use regression methods to estimate the effects of the treatment, taking into account amounts of treatment and the factors that lead to these differences.

An important feature of any experimental design is the statistical power of the study to detect a given difference between the experimental and control group. The power of the present design is summarised in Table 3.2, which describes the power of the study to detect a mean difference of a given size between the Early Start and control groups (Cohen, 1977). These power calculations are based on the sample studied at 36 months (207 controls; 184 Early Start) and thus give a slightly conservative estimate of study power. The table shows the power of the study to detect: a small effect of .2 of a standard deviation between means; a moderate effect of .5 standard deviations between means; and a large effect of .8 standard deviations between means. The table shows the design had very good power (> .99) to detect moderate (.5) and large (.8) effects. For a small effect (.2) power is only moderate ($1-\beta$ = .51). The design had 80% power to detect a difference of .28 standard deviations between groups suggesting adequate levels of power for small to moderate effect sizes in the range of 0.25–0.50 standard deviations.

3.3 The Ethics of Randomisation

An objection to the research design above is that some families were denied a potentially beneficial treatment on the basis of a random process. It could be suggested this process was unfair and that families should have received services on the basis of need rather than a chance process. There are two general responses to such criticisms. First, the claim that the services provided by Early Start were potentially beneficial is by no means self-evident. Indeed, as the literature review in Chapter 1 suggests, the majority of studies have found that family support programmes do not have detectable benefits (see reviews by Gomby et al, 1999; Olds & Kitzman, 1993). Under these circumstances there is a clear equipoise on the issue of whether or not Early Start was a beneficial programme, therefore the use of a randomised trial to resolve this uncertainty is justified. Second, the claim that services should be provided to those in need presupposes the groups who will benefit from the programme have already been identified. However, to do this would require a randomised trial to assess programme benefits.

Finally, it should be noted that the research design used did not result in any client family

Table 3.2 Statistical power to detect small, moderate and large effects (α = .05)

	Small (.2 SD)	Moderate (.5 SD)	Large (.8 SD)
Power	0.51	>.99	>.99

being denied services they would normally receive. Rather, some families were potentially benefited by being offered a service they would not have received had the trial not been conducted. All of these considerations suggest that the use of a randomised design to evaluate the efficacy and effectiveness of Early Start was ethically justifiable. This conclusion was affirmed by the ethical assessment conducted by the Canterbury Ethics Committee which gave approval for the trial to proceed.

3.4 Concluding Comment

The preceding two chapters provided an overview of the two arms of the research process – the Early Start service and the research design – which form the foundations of the outcome evaluation described in this report. The next chapter will present details on the implementation of both of these aspects, looking at such issues as: compliance and participation in the service and the research; the characteristics of families enrolled in the trial; and the factors associated with varying levels of participation in both the service and research arms of the trial.

CHAPTER 4:
Client Recruitment, Characteristics and Retention

4.1 Introduction

The purpose of this chapter is to provide a statistical background to the research results reported in subsequent chapters. This background focuses on providing descriptive and accounting information on the following aspects of the study design:

- A statistical account of the recruitment process describing the numbers screened, eligible for the randomised trial and entering the trial.

- A comparison of the characteristics of families enrolled in the Early Start and control groups.

- A statistical analysis of patterns of participation, non-participation and dropout from the service provision process.

- A statistical analysis of patterns of participation, non-participation and dropout from the data collection.

Each of these analyses serves a particular function in the overall evaluation process. Specifically:

- The analysis of the recruitment process provides an account of the extent to which it was possible to recruit those deemed to be eligible for the study.

- The analysis of the characteristics of families at baseline provides a description of the families enrolled in the trial and also provides a test of the extent to which randomisation to groups led to equivalent groups of families.

- The analysis of service retention describes the extent to which it was possible to deliver the programme to families who entered the trial.

- The analysis of research participation describes the extent to which it was possible to gather data on the outcomes of those who entered the trial.

4.2 The Client Referral Process

Recruitment for the trial took place over a 19-month period during which Plunket nurses throughout the Christchurch area screened all families with new infants using the screening method described in Chapter 2. Table 4.1 gives an account of the recruitment process. The table shows that over the recruitment period a total of 4,523 families were seen by Plunket nurses. Of these families, 588 (13%) were deemed to be eligible for the trial on the basis of the screening criteria. Of those eligible, 443 (75%) agreed to participate in the trial. This participation rate is slightly lower than that found in the pilot study of the screening process (80%, Fergusson et al, 1998). This may have been due to a number of factors including: the duration of the recruitment period; the number of nurses involved; and resistance among families to entering a controlled trial in which they were randomised to groups.

Under ideal circumstances it would have been of interest to compare the characteristics of those entering and those declining to enter the trial. Unfortunately, such a comparison was not possible owing to the information "firewall" that was set up between Early Start and the Plunket service. The nature of this firewall was that, to protect the privacy of Plunket clients, no information would be supplied to Early Start about clients who either declined or were not eligible for the service. 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, but will 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).

4.3 Characteristics of Children and Families Entering the Randomised Trial

Families who agreed to enter the Early Start trial were randomly assigned to the Early Start group and 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 (see Table 3.1 for an overview of interview content). Of the control group, 221 out of 223 families were interviewed whereas of the Early Start group, 206 out of 220 were interviewed. (The loss of 14 families from the Early Start group arose from those families who originally entered the trial but 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.

4.3.1 Social and Demographic Background

Table 4.2 compares 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; the mean age of fathers was around 27. The

Table 4.1 Client recruitment

Measure	N	%
Assessed by Plunket nurse	4,523	100
Screened positive	588	13.0
Agreed to enter trial	443	9.8

majority (>70%) of parents lacked formal educational qualifications. Between 25% to 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 that the randomisation had ensured these groups were equivalent with respect to social and demographic background.

4.3.2 Socio-Economic Background

Table 4.3 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 levels of $415 per week (assessed in 2000, 2001).

On average, families had debts (excluding mortgages) of over $1,500, and over a third 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.

Table 4.2 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 4.3 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)	$1662	$1515	.59 [1]
% Family income inadequate/very inadequate	32.9	40.4	.11 [2]

[1] t-test for independent samples.
[2] chi-squared test.

4.3.3 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 4.4, 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: over half reported being reared in a single-parent family or witnessing inter-parental family violence and over 40% reported being the victim of physical or sexual abuse in childhood.

Approximately 45% 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.

4.3.4 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.

Table 4.4 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
% Inter-parental 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.

Table 4.5 Comparison between the Early Start and control groups on parental adjustment

Measure	Controls (N = 221)	Early Start (N = 206)	p [1]
Maternal Adolescence			
% 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 the Youth Court	14.0	16.0	.56
% Became pregnant before age 16	12.7	14.6	.57
Maternal Psychological Adjustment			
% 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.

Partners included both those who were living with the mother and non-resident partners. Table 4.5 compares the Early Start and control groups on these measures. The table shows that 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 depression 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% were described as having been in trouble with the law; between 9% to 17% had current problems with alcohol or drugs; between a quarter to 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.

4.3.5 Pregnancy and Childbirth

Table 4.6 compares the Early Start and control groups on a series of measures relating to past and current pregnancy and childbirth. The table shows that the mean age at which women had first become pregnant was 19. Approximately one in seven women enrolled in the trial who had a previous child reported that 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 under 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 3.2kg and almost 85% of mothers reported breastfeeding their child. As with other comparisons, there were no significant differences between the Early Start and control groups.

4.3.6 Overall Conclusions

The series of comparisons shown in Tables 4.2 to 4.6 leads to two general conclusions. First, as a group, the families entering the trial were subject to disadvantages in a number of areas that spanned: socio-economic deprivation and problems; adverse

Table 4.6 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 (grammes)	3207	3258	.45 [1]
% Mother breast-fed child	83.7	85.4	.62 [2]

[1] t-test for independent samples.
[2] chi-squared test.

maternal childhood experiences; maternal and partner adjustment problems; and 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 three to four comparisons significant at the .10 level. The findings thus provide considerable reassurance that the assignment to the experimental and control groups produced equivalent groups of families.

4.4 Participation in Early Start

An important aspect of the research design concerns the extent of participation in Early Start among those enrolled in the service. Because the Early Start programme requires long-term enrolment in the service lasting for up to five years, it is clear that with the passage of time some families will "drop out" from the service for a variety of reasons. These reasons will include: withdrawal from the service, leaving the Christchurch region,

the child entering alternative care, and the death of children. Furthermore, among those remaining with the service the degree of participation may vary from those showing strong commitment and involvement with the service to those whose ties with the service are weaker. In turn, these variations in programme exposure and programme involvement will have a bearing on the extent to which it is possible to show positive programme outcomes. In general, as programme participation and commitment decline there will be a corresponding decline in the potential for the programme to show a positive benefit. Table 4.7 describes participation in the Early Start programme at: baseline, six, 12, 24 and 36 months following programme enrolment. At each time, the 220 clients enrolled in the trial were classified into one of three groups:

1. Currently Active: This group included all clients currently receiving services, including those who were graduated and on Level 4 (see Chapter 2 for a description of the level system).

2. Currently Inactive: This group included those clients who were currently enrolled with the service who were not receiving the service. These clients included those who had no contact with the service because of problems of client availability or elusiveness and those who were temporarily unavailable for other reasons (eg away from Christchurch for a protracted period of time).

3. Lost from the Service: This group included those clients who were no longer enrolled with the service. There were four reasons for service loss: the family had declined the service; the family had moved from the district; the family was referred to another service provider; or the child was dead (two children in the Early Start group died during the course of the trial).

Table 4.7 Rates of participation in Early Start at enrolment and six, 12, 24, and 36 months
post-enrolment

Service Duration (months)	Active[1]		Inactive[2]		Lost from Service[3]	
	N	% of sample	N	% of sample	N	% of sample
0	206	93.6	0	0.0	14	6.4
6	174	79.1	21	9.5	25	11.4
12	162	73.6	21	9.5	37	16.8
24	143	65.0	25	11.4	52	23.6
36	131	59.5	38	17.3	51	23.2

[1] Includes all those receiving service or graduated from service.
[2] Includes those currently enrolled but not in contact with the service.
[3] Includes those withdrawing from service, those leaving district, those referred to another service provider, deaths.

The table shows that:

- At 12 months, just under three-quarters of clients were actively receiving the service; just under one in 10 was inactive and 17% were lost from the service.

- At 24 months, just under two-thirds of the clients were actively receiving the service; just over 10% were inactive and nearly a quarter were lost from the service.

- At 36 months, just under 60% of the clients were actively receiving the service, one in six clients was inactive and just under a quarter were lost from the service.

These statistics make it clear that, while the Early Start service had relatively good client retention given the complexity of the service being delivered, with the passage of time there was a decline in the number of clients actively receiving services and an increase in those inactive and lost from the service. From the standpoint of the randomised trial, the important implications of these findings are that the amount of service received by clients varied from those who received no service (6.4%) to those who remained actively in the service at 36 months (59.5%). These variations in service delivery will have implications for the extent to which the programme can demonstrate positive outcomes. Specifically, as noted in the previous chapter, the recommended method for analysing randomised trials is to use an "intention to treat" paradigm which analyses the outcomes of all of those assigned to the programme irrespective of their degree of

programme participation. In situations where there is a substantial programme dropout, the intention to treat approach will be biased towards under-estimating the effects of the programme on those who receive the programme.

To examine the extent to which losses from the programme were systematic, a series of analyses was conducted to determine the extent to which participation in the programme varied with the characteristics assessed at the baseline. These analyses were conducted for two definitions of failure to deliver the service: those who were lost from the service were compared with those who were enrolled with the service; and those who were lost from the service or inactive were compared with those actively receiving the service. All analyses led to very similar conclusions and produced the finding that dropout from the service (however defined) was unrelated to any of the characteristics assessed at baseline. The comparisons made spanned over 40 variables and the findings of these analyses are described below:

1. Demographic factors: The extent of service participation did not vary significantly with parental age, parental education, family size, or family type.

2. Socio-economic factors: The extent of service participation did not vary significantly with degree of welfare dependence, family income, family debt levels, or adequacy of income.

3. Maternal childhood: The extent of service participation did not vary significantly with maternal exposure to single

parenthood, family violence, child abuse, or childhood happiness.

4. Parental adjustment: The extent of service participation did not vary significantly with maternal adjustment in adolescence (crime, pregnancy, substance use), current maternal adjustment (alcohol use, cannabis use, depression), or partner characteristics (crime, substance abuse, aggression, family violence).

5. Pregnancy and childbirth: The extent of service participation did not vary significantly with previous pregnancy history, previous children fostered, the planning of pregnancy, smoking during pregnancy, substance use in pregnancy, pregnancy history, or birth weight.

This uniform absence of association between the extent of service delivery and a wide range of social, family and personal factors tends to suggest that losses to the service occurred in a non-systematic way that was unrelated to the families' circumstances at the time of enrolment in the trial. This suggests it would not be unreasonable to treat losses from the service as occurring at random.

4.5 Participation in the Research Process

As explained in Chapters 2 and 3, the evaluation of Early Start involved a service-delivery arm and a research arm, with these processes being conducted in parallel but independently from each other. A feature of this approach was that participation in the research process was not determined by the individual's extent of participation in the Early Start service. Table 4.8 reports on the numbers in the Early Start and control groups who were assessed at: baseline, six, 12, 24 and 36 months. The table shows a high level of participation in the trial with approximately 88% of those enrolled being assessed at 36 months. The table shows a consistent and significant (p < .01) trend for a greater percentage in the control group to participate in interviews than those in the Early Start group. This trend is largely explained by higher rates of loss from the Early Start group in the first six months of the trial. At the six-month assessment, only seven control families were not assessed in comparison to 24 families in the Early Start group.

The higher rate of non-participation among the Early Start families can be attributed to the fact that a total of 14 families assigned to the Early Start group either declined to take up the service at baseline or withdrew from Early Start shortly after enrolment. These families also declined any further involvement in the research process. This situation did not apply to the control group, where only two families declined to participate in the trial.

A further factor that may have sustained participation in the control group is that members of this group were offered an honorarium of $50 per completed interview. This honorarium was offered in recognition of the fact that those in the Early Start group received a substantial indirect benefit from the services provided and it was unreasonable to expect the control group to participate in the randomised trial for no tangible recognition of their contribution.

Parallel to the analysis conducted on service participation, an analysis was conducted to examine the extent to which research

Table 4.8 Participation in the research process

Assessment Period	Controls (N=223)		Early Start (N=220)	
	N	%	N	%
Baseline	221	99.1	206	93.6
6 months	216	96.9	196	89.1
12 months	216	96.9	198	90.0
24 months	211	94.6	187	85.0
36 months	207	92.8	184	83.6

participation varied with social, demographic or related factors. This analysis showed rates of participation in the research assessments were unrelated to a wide range of factors including: demographic factors; socio-economic circumstances; maternal childhood; parental adjustment; and pregnancy history. These findings provide considerable reassurance that losses to the research process were not systematic and that it was reasonable to treat these losses as though they occurred at random. While the weight of the evidence suggests losses from the control and Early Start groups occurred at random but at different rates, differential sample losses must be considered to be a potential threat to study validity. This issue will be examined in greater depth in Chapter 7.

4.6 Summary and Conclusions

This chapter has examined a series of issues that relate to the background statistical context of the results reported in subsequent chapters. These issues included:

1. *Client recruitment*: The client-recruitment strategy using Plunket nurses as a referral source was moderately effective, with 75% of eligible clients being referred to Early Start. These referral statistics provide an indicator of the level of need for services within the community and suggest in the region of 13% of families may have been eligible for Early Start services with just under 10% being willing to enter the service. While the overall service acceptance of 75% was satisfactory, the level of refusal implied that one in every four families eligible for the trial declined to enter. Because of the need to protect client privacy, comparisons between those agreeing to referral and declining referral were not possible. Thus the extent to which the sample of families entering the trial was representative of all families eligible for the trial is unknown. However, it is important to recognise this limitation on study design will not affect the internal validity of the trial: the comparisons between outcomes of the Early Start and control groups will reflect the differences in outcomes of children and families in

these two groups. However, the extent to which such differences can be generalised to all families eligible for the trial cannot be ascertained.

2. *Characteristics of families in the randomised trial*: The results presented in this chapter provide a profile of the characteristics of the families enrolled in the trial at the point of enrolment. As would be expected from the use of the screening process, the families enrolled in the trial emerged as showing signs of disadvantage and adversity in a number of areas including: socio-economic disadvantage and poverty; adversity during maternal childhood; parental adjustment; and previous pregnancy history. The overall trends in the data are well summarised by a number of key statistics that summarise the levels of disadvantage and adversity faced by families entering the trial. These findings show that: 90% of families were welfare dependent; over 50% of mothers reported being reared in a family where there was domestic violence; nearly 50% described being brought up in impoverished circumstances; a third of mothers and over 50% of partners had been in trouble with the police; and over 80% of pregnancies were unplanned.

    These findings make it clear that the disadvantages faced by families entering Early Start were not simply confined to poverty and economic deprivation but also spanned a wider range of issues relating to personal adjustment, and partner and family relationships. All of these disadvantages share the common feature that they may lead to stresses and difficulties that impair family functioning and the child-rearing capacity of the family.

3. *Adequacy of randomisation*: The comparisons between the Early Start and control groups made it possible to test the adequacy of the randomisation process. These comparisons showed randomisation to groups produced two groups that were equivalent on a wide range of social, demographic, family, and personal factors.

4. *Participation in the Early Start service.* The Early Start programme involves a lengthy period of service provision lasting for up to five years. A vulnerability of this protracted period of service provision is that families will drop out of the service for a variety of reasons including refusal to participate, changes of residence, death of children and other factors. In addition, among those families enrolled in the service, the degree of commitment to and participation in the programme may vary. Both dropout from the programme and variations in degree of programme participation will limit the effectiveness of programme delivery and for that reason also impose limitations on the extent to which the evaluation can demonstrate positive outcomes. Given the duration of the programme and the complexities of delivering the programme, the level of participation in the Early Start programme was highly satisfactory with over three-quarters remaining enrolled in the programme at 36 months and 60% actively receiving the service.

Nonetheless, the failure of the programme to deliver an appropriate level of service to all clients over the three-year period raises issues about the extent to which limitations on service delivery may have reduced the effectiveness of the programme. A critical issue concerns the extent to which programme participation was determined by systematic factors. An examination of the relationship between levels of programme participation and factors observed at the baseline assessment revealed little or no association between the extent of programme participation and a wide range of social, family, and individual factors. This result suggests that what determined levels of programme participation may have been non-systematic chance-like factors that acted to influence the duration of family participation in the programme. In turn, this suggests it would not be unreasonable to assume that dropout from the Early Start service occurred at random.

5. *Research participation*: The assessments of Early Start were conducted using a research process that was independent from, yet run in parallel with, the process of service delivery. An implication of this was that levels of participation in the research process could differ from levels of participation in Early Start. Overall, the analysis suggested high rates of research participation with nearly 90% of those enrolled in the trial being studied for 36 months. There was no evidence to suggest dropout from the research assessment was related to a wide range of social, individual, family and related factors. This result suggests it is not unreasonable to assume that dropout from the research process occurred at random.

In general, the results in this chapter provide considerable reassurance about the overall adequacy of the processes by which client families were recruited, assigned to experimental and control groups, provided with services, and assessed. At the same time, the results show a number of departures from an idealised randomised trial in which: all eligible clients were enrolled; all clients received the full service throughout the trial; and all clients were assessed at all times. Each of these departures from the idealised randomised design poses some threat to the internal or external validity of the trial: failure to recruit all eligible clients may compromise the external validity of the trial; failure to provide all enrolled clients with a full service may bias trial results toward under-estimating treatment benefits; and failure to assess all families at all times may lead to bias in the assessment of outcomes. Subsequent chapters will examine the extent to which departures from the idealised randomised design may have posed threats to the validity of conclusions drawn from this trial.

Chapter 5:
Child-Related Outcomes

5.1 Introduction

The aims of this chapter are to examine the extent to which the provision of the Early Start programme produced beneficial outcomes for the child who was the primary client of the programme. Specifically, the analysis focuses on the extent to which children in the Early Start series showed beneficial outcomes in the following areas in which clear service goals and objectives had been set. These areas included:

1. Child health: In this domain, the aims of Early Start focussed on a series of goals that included: effective use of general practitioner services for morbidity and preventive health care; increased use of positive child health practices (breastfeeding, smoke-free home environment, home safety); reduction of hospital attendance for accidents, injuries or poisonings; and effective use of preschool dental services.

2. Early childhood education: During the second and third years of the Early Start programme a major emphasis of programme delivery was that of ensuring high rates of participation in early childhood education.

3. Encouragement of positive parenting: A major focus of the programme of home visitation employed by Early Start was on the encouragement of positive and non-punitive parenting practices.

4. Reduction of child abuse and neglect risks: Closely aligned with the emphasis on positive parenting, Early Start also focussed on approaches to reduce and manage child abuse and neglect.

5. Reductions in child problem behaviours: A further goal of the Early Start programme was to assist parents to develop skills to manage difficult child behaviours.

Against this background the present chapter examines the extent to which children in the Early Start group showed positive outcomes in the areas of health, early childhood education, parenting, child abuse and behavioural adjustment when compared with children in the control series.

5.2 Overview of Analysis and Statistical Methods

5.2.1 Assessments

The assessments of the outcomes described in this chapter have been described in detail in Chapter 3 and the Appendix. Data was gathered during the course of interviews conducted with parents at six, 12, 24 and 36 months following trial enrolment, and supplemented by information from general practitioner and hospital records. The following information was gathered as part of these assessments.

1. General Practitioner Visits

These included: the number of visits made to the family doctor by 36 months; whether the child was up to date with all immunisations at 36 months; and whether the child had received all well-child checks provided by the family doctor by 36 months. As noted in Chapter 3, information on the utilisation of immunisation and well-child checks was cross-validated by obtaining information from family doctor records. All information was gathered subject to the signed consent of the child's parents.

2. Hospital Attendance

Information on hospital attendance was gathered from hospital records up to the 36-month follow-up. Two measures of hospital attendance were analysed. These included: rates of hospital-treated accidents and poisonings; and hospital attendance for any other reason, including surgery and childhood illnesses.

3. Home-based Preventive Health Care Practices

Measures of home-based preventive health practices, including breastfeeding, smoke-free home environment and levels of home safety, were assessed on the basis of parental report. To assess overall levels of home safety, an index was constructed based on a count of the number of home-safety features present in each family's home at 12, 24 and 36 months (eg working smoke alarms; plug protectors on electric sockets; childproof latches on cupboards; locked medicine cabinet/safe storage of medicines; safe storage for poisons; hot water temperature set at 60˚C or lower; fireguards for open fires, log burners, or heaters; and an escape route planned in case of fire). This index ranged from zero to a maximum of nine.

4. Preschool Dental Care

To determine the extent to which families used free preschool dental services, parents were asked at the 36-month follow-up whether they had enrolled their child with preschool dental services.

5. Use of Early Childhood Education

At the 12, 24 and 36 month follow-up interviews information was gathered from parental reports about the use of early childhood education. This information was used to construct two indices of early childhood educational participation. These indices included: whether the child was attending early childhood education at the 12, 24 and 36 month assessments; and the estimated duration of involvement in early childhood education.

6. Maternal Reports of Parenting Behaviours

These reports were based on a series of 49 items derived from the Child Rearing Practices Report (CRPR, Block, 1981; Dekovic, Janssens, & Gerris, 1991), the Adult-Adolescent Parenting Inventory (AAPI, Bavolek & Keene, 1999; Hanson, 1990) and custom-written items assessed at 36 months. Factor analysis of the selected items suggested that these items measured two general dimensions of parenting attitudes and behaviours:

a) Positive parenting attitudes. Those scoring high on this factor tended to agree with statements suggesting they found parenting a rewarding task.

b) 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.

A total parenting score was constructed by summing the positive parenting and non-punitive parenting scales. All measures were scaled to a mean of 10 with a standard deviation of one.

7. Measures of Child Abuse and Neglect

To assess exposure to child abuse and neglect, a series of indices were constructed:

a) Parental reports of severe physical punishment: This was based on the severe/very severe assault subscales of the Parent-Child Conflict Tactics Scale (Straus, Hamby, Finkelhor, Moore, & Runyan, 1998) assessed at 12, 24 and 36 months. These subscales comprise eight items that measure severe punitive behaviours (eg "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 any incident of assault over the assessment period.

b) Parental reports of agency contact for issues relating to child abuse and neglect: At six, 12, 24, and 36 months parents were asked to describe their contacts with the Department of Child, Youth and Family Services (CYFS) over the period since the previous assessment. These accounts were analysed to identify those contacts that indicated actual or suspected incidents of child abuse and neglect.

c) Hospital admissions: Hospital records were also scrutinised to identify cases of hospital attendance for verified cases of abuse or neglect. An analysis of the records revealed seven cases in which children had been admitted to hospital with clear signs of abuse or neglect.

8. Measures of Behavioural Adjustment

Children's behavioural adjustment was assessed at 36 months using 50 items from the Infant Toddler Social and Emotional Assessment scale (ITSEA, Briggs-Gowan & Carter, 1998). These measures spanned a series of behavioural dimensions including (over)activity, aggression/defiance, peer aggression, emotional negativity, inhibition/separation problems and depression/withdrawal. These dimensions were categorised further into two overall scores of externalising behaviours (activity, aggression/defiance, peer aggression, and emotional negativity) and internalising behaviours (inhibition/separation problems and depression/withdrawal). Scoring followed the convention that a higher score implies greater behavioural problems. All measures were scaled to a mean of 10 with a standard deviation of one.

9. Measures of Intelligence

To assess intelligence at age three, the Wechsler Preschool and Primary Scale of Intelligence (WPPSI, Wechsler, 2002) was administered by trained examiners. The WPPSI was originally designed to be used with lower-class preschool children and consists of three subscales: Verbal IQ; Performance IQ; and Full Score IQ (a composite score of the verbal and performance subscales).

5.2.2 Analysis Approach

The analysis in this chapter uses an "intention to treat" design in which participants in the experimental series are contrasted with those in the control series on all available data irrespective of the extent to which those in the experimental series had participated in Early Start. This approach was used since it is widely accepted in the literature on controlled trials that the use of an intention to treat design preserves experimental validity (Ellenberg, 1996). In the comparisons below, sample sizes vary depending on the age at which comparisons are made (see footnotes to the tables) but the main comparisons made are between the 184 families receiving Early Start and the 207 control families who were assessed at 36 months.

5.2.3 Statistical Methods

The analysis uses the following statistical approaches. For analyses involving single outcomes assessed on a single occasion, standard tests of between-group differences were applied (eg chi-squared test of independence for comparison of proportions, t-test for independent groups for comparison of means) and a measure of effect size (see below) was obtained based on the standardised between-group difference. For analyses involving repeated measures on the same outcome over time, the test of between-group differences and the measure of effect size were based on an analysis of the accumulative or pooled group differences over time. (Where appropriate, preliminary analyses were conducted to test for group x time of measurement interactions. No significant interactions were detected.) The choice of statistical test for each outcome is indicated in each table.

5.2.4 Effect Size Estimates

To measure the size of effects, the analysis uses Cohen's "d" throughout. Cohen's d is defined as the standardised difference between means or proportions (Cohen, 1977). While the interpretation of d is somewhat arbitrary, Cohen suggests that an effect size of d = .20 can be described as "small", an effect size of d = .50 as "medium" and an effect size of d = .80 as "large". For consistency of presentation, effect sizes are always presented as positive in sign when the difference between groups is in favour of improved outcomes for the Early Start group.

5.2.5 Multivariate Testing

An issue raised by the use of multiple outcomes concerns the effects of conducting multiple tests on the interpretation of the significance test and the risk of type I statistical errors. In the final part of the results section of this chapter, we develop a multivariate regression modelling approach that permits simultaneous estimation of the probability of observed trial results under the null hypothesis of no differences.

5.2.6 Treatment of Missing Data

In this chapter we analyse the results using all data available on all participants at each time of observation. However, as noted in Chapter 4, there were significantly higher rates of sample loss in the Early Start series

Table 5.1 Comparison between the Early Start and control groups on general practitioner visits

Measure [1]	Controls	Early Start	p	Effect size (d)
Mean (SD) number of GP visits				
0–12 months	12.1 (5.6)	13.4 (6.6)		
12–24 months	5.2 (4.7)	6.1 (6.0)		
24–36 months	3.2 (4.1)	4.0 (4.8)		
Total visits (0–36 months)	20.7 (10.5)	23.5 (13.8)	<.05 [2]	.24
% Up to date with immunisations				
6 months	83.3	90.8		
12 months	94.0	92.4		
24 months	91.9	92.5		
36 months	91.9	92.5	.83 [3]	.02
% Up to date with well-child checks				
6 months	67.8	80.8		
12 months	30.1	41.9		
36 months	30.1	41.9	<.05 [3]	.25

[1] Sample sizes for different comparisons were: 12 months (Controls N = 216, Early Start N = 198); 24 months (Controls N = 211, Early Start N=187); 36 months (Controls N = 207, Early Start N = 184); 0–36 months (Controls N = 206, Early Start N = 182).

[2] t-test for independent samples.

[3] chi-squared test.

than in the control series. Although these losses did not appear to be systematic, they raise possible threats to study validity. In Chapter 7 we examine these threats in detail through the use of missing data estimation methods.

5.3 Results

5.3.1 Child-related Outcomes

1. General Practitioner Visits

Table 5.1 compares the Early Start and control groups on three measures of general practitioner attendance: first, the mean annual rate of attendance at 12, 24 and 36 months; second, the percentage of children who were up to date with immunisation at 12, 24 and 36 months; and third, the percentage of children who were up to date with well-child checks at six,12 and 36 months. Measures of immunisation and well-child checks were based on general practitioner records data, whereas the total number of attendances was based on parental reports data.

The table shows those in the Early Start group had a significantly higher rate of attendance at family doctors: at 36 months

the Early Start series had seen their family doctors an average of 23.5 occasions compared to an average of 20.7 visits for the control series (p < .05). This higher overall rate of medical contact was also reflected in higher rates of general practitioner contacts for well-child checks (p < .05). However, rates of immunisation were similar for both groups with in excess of 90% of both groups being up to date with immunisation at 36 months. For significant outcomes (medical attendance, well-child checks) effect sizes were small (d = .24; .25 respectively), and for immunisation the effect size was negligible (d = .02).

2. Hospital Attendance

Table 5.2 compares the Early Start and control groups on a series of measures of hospital attendance and admission. These measures were derived from hospital casenote material gathered on the Early Start and control groups. The table shows those in the Early Start group had significantly lower rates of hospital attendance for injury and poisoning (p < .05). The Early Start group also had lower rates of attendance for other reasons, although this difference was not statistically significant (p = .21). The effect size estimates for hospital attendance ranged from .13 to .22.

Table 5.2 Comparison between the Early Start and control groups on hospital attendance

Measure [1]	Controls	Early Start	p	Effect size (d)
% Attended hospital for accident/injury or accidental poisoning				
0–12 months	6.9	7.1		
12–24 months	12.8	5.9		
24–36 months	9.7	8.7		
Ever (0–36 months)	26.3	17.5	<.05 [2]	.22
% Attended hospital for any other reason				
0–12 months	50.9	46.5		
12–24 months	45.5	43.9		
24–36 months	37.2	29.4		
Ever (0–36 months)	74.2	68.4	.21 [2]	.13

[1] Sample sizes for different comparisons were: 12 months (Controls N = 216, Early Start N = 198); 24 months (Controls N= 211, Early Start N=187); 36 months (Controls N = 207, Early Start N = 184); Ever (Controls N = 205, Early Start N = 177).
[2] chi–squared test.

Table 5.3 Comparison between the Early Start and control groups on home-based health provisions

Measure [1]	Controls	Early Start	p	Effect size (d)
% Breastfed for six months or more	34.1	31.3	.54 [2]	-.05
% Smoke-free home/smoke-free area				
12 months	93.1	96.5		
24 months	91.5	87.7		
36 months	93.7	92.9		
Always smoke free (0–36 months)	81.6	80.8	.84 [2]	-.02
Mean (SD) number of home-safety features				
12 months	4.8 (1.7)	4.8 (1.8)		
24 months	4.9 (1.6)	5.4 (1.6)		
36 months	5.1 (1.4)	5.2 (1.4)		
Average (0–36 months)	4.9 (1.3)	5.2 (1.3)	.09 [3]	.17

[1] Sample sizes for different comparisons were: 12 months (Controls N = 216, Early Start N = 198); 24 months (Controls N = 211, Early Start N = 187); 36 months (Controls N = 207, Early Start N = 184); 0–36 months (Controls N = 206, Early Start N = 182).
[2] chi-squared test.
[3] t-test for independent samples.

3. Home-based Preventive Health Care Practices

Table 5.3 compares the Early Start and control groups on a series of measures of home-based preventive health practices including breastfeeding, smoke-free home environment and levels of home safety. The home-safety index reported in this table was based on a count of the number of home-safety features present in each family's home at 12, 24 and 36 months. This index ranged from zero to a maximum of nine. The table shows the two groups were not significantly different in rates of breastfeeding or the percentages of children living in a smoke-free environment (d = -.05, -.02 respectively). However, there was a marginally significant (p = .09) trend for the Early Start group to have more home-safety features than the control group (d = .17).

4. Preschool Dental Care

Table 5.4 shows the proportions of children in the Early Start and control groups who were enrolled with the preschool dental service by the 36-month follow up. The table shows rates of dental service enrolment for the Early Start group were significantly higher (p < .05) than for the control group. The effect size for this comparison was d = .20.

5. Enrolment in Early Childhood Education

Table 5.5 compares the Early Start and control groups on rates of attendance at early childhood education facilities and the total duration of attendance up to 36 months. The table shows over the three-year follow-up period there was a consistent trend for the children in the Early Start group to have higher rates of participation in early childhood education. Overall, 90.8% of the Early Start group had ever attended early childhood education by the 36-month follow-up, compared to 84.5% of the control group (p = .06). The higher overall rate of attendance among the Early Start group was also reflected in the duration of early

Table 5.4 Comparison between the Early Start and control groups on preschool dental care

Measure	Controls (N = 207)	Early Start (N = 184)	p	Effect size (d)
% Enrolled with dental service at 36 months	62.8	72.3	<.05 [1]	.20

[1] chi-squared test.

childhood education. By the 36-month follow-up, the Early Start group had attended early childhood education facilities for a mean of 16.4 months compared to a mean of 13.6 months for the control group (p < .05). The effect size estimates for these comparisons ranged from d = .19 to .22.

6. Parenting

Table 5.6 compares the Early Start and control groups on a series of measures of parenting attitudes assessed on the basis of maternal reports at the 36-month follow-up. These measures include scales of positive parenting attitudes, non-punitive parenting attitudes, and a total parenting score derived from a sum of the two attitude scales. For each group, the table reports the mean scores in each measurement domain. To simplify presentation, all measures have been scaled to a mean of 10 with a standard deviation of one. This convention makes it possible to translate between-group differences in means into standard deviation units and thus into direct estimates of effect size.

Inspection of the table shows that, on all comparisons, the Early Start group had significantly (p < .05) better overall mean scores than the control group. The effect size estimates for these comparisons ranged from .22 to .27, suggesting evidence of small but consistent tendencies for the Early Start group to report more positive and non-punitive attitudes to parenting.

7. Child Abuse and Neglect

Table 5.7 compares the Early Start and control groups on two measures of child abuse and neglect up to the 36-months follow-up: parental reports of the use of severe/very severe physical assault and parental reports of agency contact for abuse/ neglect.

The table shows there were marked and consistent differences in the rate of assault between the two groups. By 36 months, 11.7% of the control series had been exposed to severe physical assault compared to only 4.4% of the Early Start group (p < .01). The effect size for this comparison was d = .26.

Table 5.5 Comparison between the Early Start and control groups in attendance at early childhood education

Measure [1]	Controls	Early Start	p	Effect size (d)
% Attending early childhood education				
12 months	38.9	50.5		
24 months	56.9	65.2		
36 months	73.9	82.6		
Ever by 36 months	84.5	90.8	.06 [2]	.19
Mean (SD) duration of attendance (months)	13.6 (12.4)	16.4 (12.8)	<.05 [3]	.22

[1] Sample sizes for different comparisons were: 12 months (Controls N = 216, Early Start N = 198); 24 months (Controls N = 211, Early Start N = 187); 36 months (Controls N = 207, Early Start N = 184); 0–36 months (Controls N = 206, Early Start N = 182).
[2] chi-squared test.
[3] t-test for independent samples.

Table 5.6 Comparison between the Early Start and control groups on maternal parenting attitudes at 36 months [1]

Measure	Controls (N = 207)	Early Start (N = 184)	p [2]	Effect size (d)
Mean positive parenting attitudes 36 months	9.88	10.14	<.01	.26
Mean non-punitive attitudes 36 months	9.90	10.12	<.05	.22
Mean total parenting score 36 months	9.87	10.14	<.01	.27

[1] For ease of interpretation, all parenting scores have been standardised to a mean of 10 and a standard deviation of one.
[2] All p-values from t-tests for independent samples.

There were, however, no differences in rates of agency contact for child abuse and neglect (p = .39; d = .04).

Table 5.7 does not contain specific information on hospital admissions for child abuse and neglect. The reason for this was that this outcome was, fortunately, uncommon and only seven children were admitted to hospital with frank signs of abuse and neglect. Of these children, two were in the Early Start group and five were in the control group. While the number of cases involved is too small for this comparison to be tested for statistical significance, the general trends in hospital admission are, nonetheless, consistent with the results in Table 5.7 and, particularly, the lower rates of severe/very severe physical assault in the Early Start group. Table 5.8 provides a summary of the hospital admissions for child abuse and neglect in the two groups.

8. Child Behavioural Adjustment

Table 5.9 compares the Early Start and control groups on a series of measures of child behavioural adjustment assessed at the 36-months follow-up. As explained previously, these measures are based on the externalising and internalising scales of the ITSEA (see Appendix). The externalising scale assesses behavioural domains relating to activity, aggression/defiance, peer aggression and emotional negativity. The internalising scale assesses domains relating to inhibition/separation problems and depression/withdrawal. The table compares the mean scores for the two groups on each behavioural domain. The table also reports means for the aggregate externalising, internalising scales and a total behaviour score combining over all domains. As with the parenting measures above, all variables have been scaled to a mean of 10 and a standard deviation of one to permit comparisons between measures and to provide a direct translation of mean

Table 5.7 Comparison between the Early Start and control groups on measures of child abuse and neglect

Measure[1]	Controls	Early Start	p[2]	Effect size (d)
% Severe/very severe physical assault by any parent				
0–12 months	3.2	0.5		
12–24 months	5.7	2.1		
24–36 months	6.8	1.6		
Ever by 36 months	11.7	4.4	<.01	.26
% Parental report of agency contact for abuse/neglect (cumulative)				
0–12 months	14.8	10.1		
0–24 months	19.0	15.0		
0–36 months	21.3	19.6	.39	.04

[1] Sample sizes for different comparisons were: 12 months (Controls N = 216, Early Start N = 198); 24 months (Controls N = 211, Early Start N = 187); 36 months (Controls N = 207, Early Start N = 184); Ever/0–36 months (Controls N = 206, Early Start N = 182).
[2] chi-squared test.

Table 5.8 Hospital admissions for severe child abuse and neglect

Controls			Early Start		
Child	Reason	Perpetrator	Child	Reason	Perpetrator
1	Serious parental neglect	Both parents	1	Table thrown at child, head injury	Father
2	Shake injury	Mother	2	Concern over care of child	Mother
3	Serious bruising on bottom	Father			
4	Facial bruising	Father			
5	Head butted, concussed	Father			

differences into standard deviation/effect size estimates. All variables are scored using the convention that an increasing score implies increasing behavioural problems.

The table shows that, in all comparisons, the Early Start group scored consistently lower than the control group. For the externalising subscales, means for the Early Start group were between .07 to .24 standard deviations lower than in the control group, and for the internalising subscales the Early Start means were .21 to .25 standard deviations lower. These differences were mirrored in the aggregate scale scores, and suggest evidence of small but consistent tendencies for the Early Start series to have fewer behavioural problems than the control series. For the aggregate externalising scale the mean difference between the groups was marginally significant (p = .06), and for the internalising and total behaviour scales the differences were significant (p < .05). The effect size estimates for these comparisons ranged from d=.19 to d=.26.

9. Effects on Child Intelligence

At the 36-month follow-up, the intelligence of both groups was assessed using the WPPSI. Table 5.10 compares the Early Start and control groups on assessments of verbal IQ, performance IQ, and total IQ. The table shows that, while the Early Start group scored slightly higher on all tests, in no case did these differences approach statistical significance. The effect size estimates for these comparisons ranged from d = .04 to .08.

5.3.2 Multivariate Tests

A limitation of the analyses reported in Tables 5.1 to 5.10 is that the significance testing procedure does not take into account the fact that the analyses use multiple tests of significance. It is well known the use of multiple comparisons can give rise to misleading conclusions, since the use of multiple tests increases the probability of type I statistical errors for any given

Table 5.9 Comparison between the Early Start and control groups on child behavioural adjustment at 36 months[1]

Measure	Controls (N = 207)	Early Start (N = 184)	p[2]	Effect size (d)
Mean externalising scores				
Activity	10.10	9.86		
Aggression/defiance	10.05	9.94		
Peer aggression	10.07	9.92		
Emotional negativity	10.11	9.88		
Total externalising score	10.09	9.90	.06	.19
Mean internalising scores				
Depression/withdrawal	10.12	9.87		
Inhibition/separation	10.10	9.89		
Total internalising score	10.12	9.86	<.01	.26
Mean total behaviour score	10.11	9.87	<.05	.24

[1] For ease of interpretation, all child behaviour scores have been standardised to a mean of 10 and a standard deviation of one.
[2] All p-values from t-tests for independent samples.

Table 5.10 Comparison between the Early Start and control groups on measures of child cognitive ability

Measure	Controls (N = 196)	Early Start (N = 171)	p[1]	Effect Size (d)
Mean (SD) Verbal IQ	97.0 (11.8)	97.7 (12.5)	.58	.04
Mean (SD) Performance IQ	96.8 (13.2)	98.1 (13.8)	.37	.08
Mean (SD) Total IQ	96.5 (12.5)	97.7 (12.9)	.39	.07

[1] All p-values from t-tests for independent samples.

comparisons. To address this issue there is a need to employ multivariate methods to test the overall levels of significance. The approach used to assess overall significance levels is described below.

Figure 5.1 shows a plot of the values of Cohen's "d" for the 22 statistical tests reported in Tables 5.1 to 5.10. Under the null hypothesis of no between-group differences one would expect to find the values of Cohen's d approximately normally distributed around a value of zero. As the figure shows, this is far from the case and with two exceptions the values of d were positive and were distributed around a mean of .16. This plot clearly provides intuitive evidence suggesting it is unlikely the series of results reported in Tables 5.1 to 5.10 reflected chance variation.

Further inspection of the figure shows a bimodal distribution with peak values at 0 to .05 and .21 to .25. This distribution is consistent with the view that the overall distribution of d values reflected a mixture of two sets of results:

1. A series of results in which treatment effects were negligible with this series having a mean in the region of 0 to .05. These results largely reflect the absence of association between the intervention and measures of immunisation, home safety, agency contact and childhood intelligence.

2. A series of results in which there were

appreciable treatment effects with this series having a mean in the region of .21 to .25. These results reflect the general tendency for children in the Early Start group to fare better on child health, preschool education, parenting, child abuse and behaviour.

To provide a more formal test of the general null hypothesis that the two groups did not differ on the series of measures used in Tables 5.1 to 5.10, a multivariate regression model was fitted to the data. This model was:

$$Y = B\,X + U$$

where Y was the nx1 vector of outcome measures, X was the dichotomous treatment variable (Early Start/Control) and U the nx1 vector of disturbances. These disturbance terms were permitted to be correlated. The parameter matrix B was the nx1 vector of regression coefficients linking the treatment variable to the n outcomes. The null hypothesis test was Ho: B=0 implying that variations in treatment status had no effect on the outcome measures Y. A path diagram depicting this model is given in Figure 5.2.

A potential difficulty in fitting this model was that the outcome measures included a series of different types of measures involving dichotomies, count and continuous measures. Since the outcome measures Y did not have a multivariate normal distribution, methods of estimation using maximum likelihood were not justified. In this instance, model fitting was conducted using asymptotic distribution free methods of weighted least squares. These methods may be applied to fit structural models involving any multivariate distribution of continuous or categorical variables, and to produce estimates of test statistics that are corrected for departures from multivariate normality (Joreskog & Sorbom, 1993). From this model it was possible to construct a log likelihood ratio chi-square test of the null hypothesis B=0 by fitting a model in which all elements of B were constrained to zero. However, to avoid problems of multicollinearity, it was necessary to omit the total parenting and total behaviour scores from this analysis. The value of the resulting test statistic was $\chi^2 = 65.2$ (df = 20; p < .0001) suggesting the presence of a highly significant deviation

Figure 5.1 Frequency distribution of effect size (Cohen's d) estimates for child-related outcomes

observed mean = .16
expected mean (under Ho) = 0.0
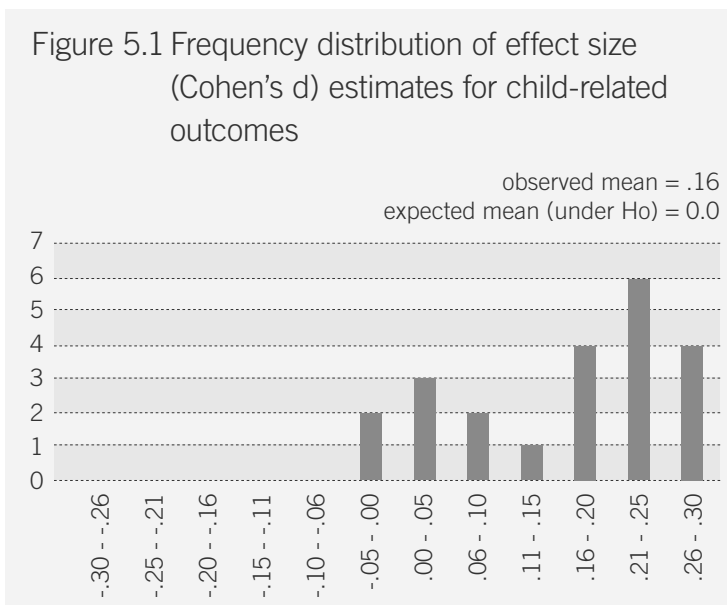
from the general null hypothesis. This result is consistent with the plot shown in Figure 5.1, which shows a systematic trend for the Early Start treated group to have better outcomes on measures of child health; preschool education; parenting; child abuse and neglect and early behavioural adjustment.

5.4 Summary and Conclusions

The analyses in this chapter suggest that the provision of Early Start services results in small but pervasive benefits for a series of outcomes relating to child health, preschool education, parenting, physical child abuse and early behavioural adjustment. A summary and review of the major findings is given below:

1. Child health: Children in the Early Start group received health benefits in a number of areas. These included: greater attendance at family doctors; higher uptake of well-child services; reduced rates of hospital admission for accidents and poisonings; and increased use of preschool dental services. In each case these outcomes represent areas targeted by the Early Start programme. However, there were some exceptions to the findings of health benefits for the Early Start group. First, the Early Start group had a similar rate of immunisation to the control group. The reason for this was

that both groups had a high uptake (> 90%) of immunisation. This high rate of immunisation appears to reflect circumstances prevailing during the course of the trial. In particular, in 2000 the Pegasus GP group introduced a home visitation service focussed on improving the uptake of immunisation. It would appear this service was effective in raising overall rates of immunisation in Christchurch, thus reducing the potential for the Early Start service to produce benefits in this area. Other areas in which no benefits were found were breastfeeding and the use of smoke-free zones. In addition, the programme had only small but marginally significant effects on levels of home safety.

2. Preschool education: An important goal of the Early Start service was to encourage high levels of participation in preschool education. In line with this focus, the evaluation shows children in Early Start spent significantly longer in early childhood education than their peers in the control group. Nonetheless, these differences were not large. This was mainly because of a high level of uptake of preschool educational services by the control group. By the 36-month follow-up, 84% of controls had enrolled in preschool education compared to 91% of the Early Start group.

3. Parenting: An ongoing focus of the Early Start programme was to increase rates of positive and non-punitive parenting. Findings showed that, at the 36-month assessment, parents in the Early Start programme had significantly higher rates of positive and non-punitive parenting attitudes. However, these effect sizes were relatively small, involving differences of approximately .25 standard deviations between the Early Start and control groups.

4. Child abuse and neglect: A goal that was closely related to increasing the rates of positive parenting was a reduction in risks of child abuse and neglect. On two measures there was evidence suggestive of programme benefits in this area. First, parents in the Early Start group reported a far lower use of severe methods of

Figure 5.2   Multivariate regression model
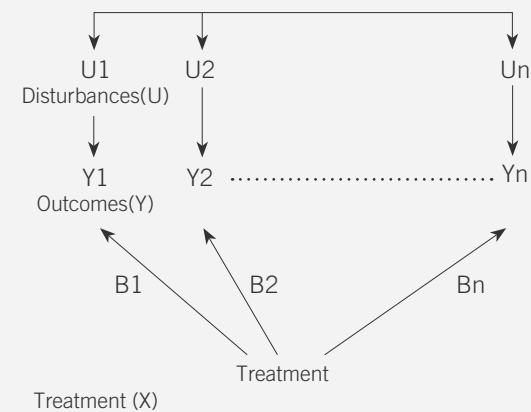
Model $Y = B X + U$
Null Hypothesis $B = 0$ ($B1=0$, $B2=0…Bn=0$)

punishment than control parents (4.4% vs 11.5%). Second, these differences were paralleled by differences in rates of hospital admission for child abuse and neglect, with five of the seven children admitted to hospital for these reasons coming from the control group. However, no differences were found on reported rates of contact with the Department of Child, Youth and Family Services for child abuse and neglect issues. At the same time, it should be noted that measures of official contact may have been biased by the fact families in the Early Start group were under greater surveillance as a result of their participation in Early Start. This feature is likely to have biased study results against finding a difference on this measure.

5. Child behaviour: One of the goals of Early Start was to reduce rates of problem behaviours in the preschool years. The results at 36 months show small but statistically significant reductions in problem behaviour scores for externalising, internalising and total behaviour problems.

6. Intelligence: Although children in the Early Start group had slightly higher IQ scores than the control group, in no case was this difference statistically significant, suggesting that enrolment in Early Start did not lead to improvements in child cognitive ability. At the same time, it is important to recognise that the focus of the Early Start programme was not on cognitive ability but rather on the areas of child health and psychosocial adjustment.

7. Overall effects: The preceding analyses suggested the Early Start programme produced a series of small benefits for child-related outcomes. However, it could be suggested these conclusions were biased by the use of multiple significance testing procedures. However, as shown in the preceding section, this conclusion is implausible. An "omnibus" test of significance using multivariate regression methods showed highly significant departures ($p < .0001$) between the observed data and the general null hypothesis of "no between-group differences". A plot of effect sizes using

Cohen's "d" supported this view. This plot showed a distribution of d values around a mean of approximately .16. Under the general null hypothesis of "no difference" one would expect to find the values of "d" distributed around a mean of zero.

The above findings need to be considered in the light of two major issues. The first issue concerns the reasons for the programme showing only small benefits in most areas. The second issue concerns the effects of various threats to the validity of study conclusions. These issues are examined below:

1. Explanations of Small Effects

There are a number of possible explanations of the relatively small effect sizes found in this analysis. These include:

a) Services provided to the control group: As we have noted previously, the control group were not an "untreated" group. Rather, this group had access to the wide range of health, education and social services available in Christchurch. These services account for relatively high levels of service utilisation seen within the control group. This point is particularly relevant with respect to immunisation, where the comparison between the Early Start group and the control group became a comparison of the effectiveness of home visitation provided by Early Start and home visitation provided by Pegasus Health to the control group. These considerations suggest one of the reasons for the relatively modest benefits found in this trial was that the control group was exposed to a relatively high level of service provision within the Christchurch region.

b) Programme retention: As noted in Chapter 3, the analysis used in this report uses an "intention to treat" design in which all those enrolled in Early Start are compared with the control group. While this method preserves experimental validity, it has the liability of producing conservative estimates of effect sizes in cases where those in the experimental group fail to receive treatment for various reasons. These issues are particularly important in the evaluation of family

support services owing to the lengthy duration of service delivery and the consequent possibility of programme dropout due to residential change, resistance to the programme, or other factors. The role of programme dropout is clearly evident in this study, with only just over 50% of those originally enrolled in Early Start being active participants in the programme at the 36-months follow-up. The effect of this level of dropout will be to bias study findings towards the null hypothesis of no treatment difference.

c) Fidelity of programme delivery: In addition to issues relating to the experience of the control group and programme retention, a further threat to validity arises from the extent to which family support workers delivered the intended programme. There is a substantial literature that suggests the programme of home visitation delivered by family support workers may differ substantially from that intended by the programme designers (eg Duggan, Fuddy et al, 2004; Gomby et al, 1999; Kitzman, Cole, Yoos, & Olds, 1997). In the development of Early Start, these problems of programme fidelity were recognised and addressed through the use of regular supervisory sessions that focussed on the achievement of specific programme goals. Nonetheless, even with such supervision it is likely there was some variation in the extent to which various aspects of the programme were delivered to families.

d) Family heterogeneity: Randomised trials of home visitation services such as Early Start differ from standard clinical trials in two important respects. First, those enrolled in the trial are not a homogeneous group of families facing a common set of issues. Rather, families are a heterogeneous group, with heterogeneous needs, facing a range of circumstances and issues. Second, the families enrolled in the service do not receive a fixed programme of service delivery but rather a programme of service delivery tailored to their needs. Both sets of factors are likely to result in a situation in which programme delivery

results in small pervasive benefits rather than large specific benefits. This point can be illustrated by considering the findings for severe punishment reported in Table 5.7. This table shows that parents in Early Start reported quite a substantial reduction in the use of severe punishment when compared with the control group (4% vs 11%). However, the potential for the Early Start programme to show benefits was limited by the fact the great majority of parents enrolled in the trial would not have used these methods of punishment, irrespective of whether or not they were enrolled in Early Start.

e) Family potential for change: Home visiting and family support programmes begin with the optimistic assumption that, by the formation of a supportive and equal partnership between the home visitor and the client family, it is possible to work towards positive family change (Webster-Stratton, 1998). While this assumption is an important foundation of family support provision, it is clear that in practice this ideal will not always be achieved. There may be difficulties of "fit" between the worker and the family, the family may be resistant to programme goals, or life stresses and problems may limit the extent to which the family is able to participate in the programme. In turn, each of these factors will limit the potential of a home visitation programme to produce positive change.

A common feature of the explanations above is that they are all likely to conspire to reduce effect sizes and bias trial results towards the null hypothesis. Specifically: the provision of alternative services within the community may dilute treatment effects; dropout from, or failure to participate in, the programme will reduce the size of treatment effects; failure of staff to adequately deliver the programme will reduce programme efficacy; heterogeneity in the client population may impose limitations on the extent of achievable benefit; and limitations on family ability to respond to the programme may reduce the overall effectiveness of the programme. These considerations suggest it would be unrealistic to expect trials of family

support programmes to show evidence of large changes in the outcomes of client families. Rather, what one might expect to find is that successful programmes show evidence of small but consistent gains across a wide range of family outcomes. Judged by that criterion, the present chapter suggests the Early Start programme was generally successful in its objectives of improving child-related outcomes.

2. Threats to Trial Validity

The preceding conclusions, however, need to be leavened by a recognition of potential threats to trial validity, which may have conspired to convey a false impression of the efficacy of Early Start.

These threats are reviewed below:

a)   Open nature of the trial: Of necessity, the evaluation of Early Start did not use a double-blind design in which trial participants and research interviewers were unaware of the intervention group to which the family belonged. This lack of blinding in the study design may have led to a bias in which those in the Early Start group provided more positive reports of parenting, child rearing and other outcomes than those in the control series. While there is no way of fully addressing this concern using the present design, there are a number of indications in the findings that suggest the absence of blinding was not a major threat to trial validity. In particular, measurements based on official data including general practitioner records and hospital records were unlikely to be subject to any bias as a result of the open nature of the trial. However, these measures showed benefits in the areas of well-child care, injury and poisoning, and hospital admission for child abuse and neglect that were consistent with the results for parent-reported outcomes. The robust nature of the findings across different measurement sources (parental report, general practitioner records, hospital records) argues against the view that the open nature of the trial posed a major threat to study validity.

b)   Sample attrition: A potential flaw in the study design was that losses to follow-up in the experimental group were greater than in the control group. This difference arose largely from higher rates of sample loss at the beginning of the trial and this appeared to be due to an initial inexperience of FSWs enrolling families in the service and the trial. The greater loss of families from the Early Start group is a potential threat to study validity since it could be proposed that those families remaining in the study were less at risk than those not remaining in the study. There are two possible ways of addressing this threat to validity. The first is to compare the Early Start and control groups that were studied on the basis of baseline measures to examine the extent to which these groups were equivalent at the baseline assessment. The alternative is to use methods of missing data estimation to estimate outcomes for all study participants. Chapter 7 uses these methods to examine study bias and finds no evidence of bias in study results as a result of losses to follow-up.

c)   Investigator bias: A further possibility that could be raised is that the results reflect an investigator bias in which study findings favourable to the trial have been presented and unfavourable outcomes have been suppressed. This criticism may be addressed in two ways. First, the outcomes chosen in this analysis have all followed directly from the stated goals of Early Start and we have tried to avoid including outcomes on the basis of "data dredging". For each outcome chosen there is a clear and self-evident justification for the inclusion of that measure in the analysis. Second, as we will show in the following chapter, while it is possible to demonstrate small but pervasive benefits for child outcomes, no such benefits can be shown for parent or family outcomes. The specificity of the evaluation results thus argues against the view that these results have been contaminated by an investigator bias.

CHAPTER 6:
Maternal and Family Outcomes

6.1 Introduction

The preceding chapter examined the benefits of Early Start for a number of child-related outcomes that span health, education, parenting, and behavioural adjustment. In this chapter the focus broadens to examine the extent to which the provision of Early Start services had benefits in the following areas:

- Maternal health and well-being.
- Family stability, family relationships and family violence.
- Family economic and material well-being.
- Family exposure to stress and adversity.

Each of these themes represents an area in which the Early Start programme had specific goals. These goals included: promoting maternal health and well-being; reducing and preventing family violence; increasing family material well-being; and reducing family susceptibility to stress and crisis.

6.2 Overview of Analysis and Statistical Methods

6.2.1 Assessments

Data on maternal health and well-being, family violence, material well-being and stress and crisis was gathered at interviews conducted with parents at six, 12, 24 and 36 months following trial enrolment.

1. Maternal Health and Well-Being

a) Contraceptive use: At each assessment, parents were asked whether they were currently using any form of contraception.

b) 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.

c) 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.

d) Maternal substance use (tobacco, alcohol and other drugs): At each assessment, parents were questioned concerning 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 substance use problems if they reported any abuse or dependence item over each 12-month follow-up period.

2. Family Stability, Family Relationships and Family Violence

a) 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.

b) Family violence: At each assessment, parents were questioned about partner violence using the revised Conflict Tactics Scale (CTS2) (Straus, Hamby, Boney-McCoy, & Sugarman, 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.

3. Family Economic and Material Well-Being

A number of measures were used to describe the economic and material well-being of the families in the trial. These measures included:

a) 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 were currently reliant on a social welfare benefit.

b) 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.

c) Adequacy of family income: At each assessment, parents were asked to rate the adequacy of their income to meet their basic family needs on a four-point scale from "income more than adequate" to "income very inadequate".

d) Debt: At each assessment, parents were asked to report the amount of money owed in debt, excluding hire purchase or mortgage costs.

e) 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.

f) Adequacy of accommodation: At each assessment, parents were asked to rate the adequacy of their accommodation to meet their family's needs on a four-point scale from "accommodation more than adequate" to "accommodation very inadequate".

g) Economic hardship: Economic hardship factors were recorded at each assessment from parental reports of those factors

experienced by the family since the last assessment. These economic hardship factors included, for example: "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.

4. Family Exposure to Stress and Adversity

At each point of assessment, respondents were questioned on a 45-item questionnaire regarding their exposure to stressful and adverse life events. These events were categorised into five dimensions of stress and adversity: illness and death; economic and financial problems and crises; family or social relationship problems; victimisation; and any other life events. 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.

6.2.2 Analysis Approach and Statistical Methods

The analysis in this chapter uses the same array of assumptions and methods to those used in Chapter 5. Specifically:

- The analysis uses an "intention to treat" design.
- The statistical methods used follow those used in Chapter 5 (eg chi-squared test of independence for comparison of proportions, t-test for independent groups for comparison of means).
- Effect sizes are measured by Cohen's d.
- The set of results is tested for overall significance using multivariate regression.
- The results are analysed using all data available on all participants at each time of observation (sample sizes are shown in the footnotes of each table).

Table 6.1 Comparison between the Early Start and control groups on maternal health and well-being

Measure [1]	Controls	Early Start	p^2	Effect size d
% Currently using contraception				
12 months	47.4	51.0		
24 months	56.4	53.2		
36 months	54.1	50.0		
Pooled average (0–36 months)	52.6	51.4	.80	-.02
% Ever pregnant to 36 months	47.6	42.9	.35	.09
% Major depression				
0–12 months	26.5	24.0		
12–24 months	21.3	17.7		
24–36 months	15.9	16.9		
Ever (0–36 months)	37.1	36.0	.82	.02
% Mother smoked cigarettes				
12 months	64.4	61.1		
24 months	65.9	66.3		
36 months	62.3	62.0		
Ever (0–36 months)	68.9	73.1	.37	-.09
% Substance use problems (past 12 months)				
12 months	19.9	27.3		
24 months	16.1	19.8		
36 months	13.5	19.0		
Ever (0–36 months)	33.0	38.5	.26	-.11

[1] Sample sizes for different comparisons were: 12 months (Controls N = 216, Early Start N = 198); 24 months (Controls N = 211, Early Start N = 187); 36 months (Controls N = 207, Early Start N = 184); Ever/0–36 months (Controls N = 206, Early Start N = 182).
[2] chi-squared test.

6.3 Results

6.3.1 Maternal and Family-Related Outcomes

1. Maternal Health and Well-Being

As part of the research process, data was gathered on a series of maternal outcomes at the 12, 24 and 36-month follow-up assessments. Table 6.1 compares the Early Start and control groups on a series of measures describing maternal health and well-being during the three-year follow-up. These measures span: contraception; further pregnancy; maternal depression; maternal substance use; and maternal cigarette smoking.

The table shows a consistent lack of association between maternal outcomes and group status. 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, major depression) women in the Early Start group fared better than women in the control group, whereas in others (use of contraceptives, substance use problems, cigarette smoking), women in the control group fared better than women in the Early Start group. These findings are consistent with the view that the provision of the Early Start service offered mothers no benefits in the areas described in Table 6.1. The effect size estimates for these comparisons ranged from -.11 to .09.

2. Family Stability, Family Relationships and Family Violence

Table 6.2 compares the Early Start and control groups on a series of measures of family stability, family relationships and family violence assessed at 12, 24 and 36 months. These measures include: the fraction of single-parent families; the number of

Table 6.2 Comparison between the Early Start and control groups on family stability, family relationships and family violence

Measure[1]	Controls	Early Start	p	Effect size d
% Single-parent family				
12 months	59.3	64.1		
24 months	60.2	64.2		
36 months	57.0	62.0		
Ever (0–36 months)	82.0	85.9	.30 [3]	-.11
% Separated from partner				
0–12 months	13.9	11.6		
12–24 months	9.5	10.2		
24–36 months	6.8	10.4		
Ever (0–36 months)	26.3	31.1	.31 [3]	-.10
Mean number of family changes (0–36 months)	1.1	1.0	.64 [2]	.05
% Mother assaulted by any partner				
0–12 months	12.5	9.1		
12–24 months	10.0	11.8		
24–36 months	7.3	8.7		
Ever (0–36 months)	22.3	26.4	.35 [3]	-.10
Mean number of residential changes (0–36 months)	2.5	2.5	.98 [2]	.00

[1] Sample sizes for different comparisons (excluding partner measures) were: 12 months (Controls N = 216, Early Start N = 198); 24 months (Controls N = 211, Early Start N = 187); 36 months (Controls N = 207, Early Start N = 184); Ever/0–36 months (Controls N = 206, Early Start N = 182).
[2] t-test for independent samples.
[3] chi-squared test.

families experiencing a parental separation; the mean number of family changes; the fraction of mothers who experienced any incident of assault by a partner; and the mean number of residential changes.

Inspection of Table 6.2 suggests two conclusions. First, families in both the Early Start and control groups were subject to what appears to be high levels of family change, instability and conflict, which

Table 6.3 Comparison between the Early Start and control groups on family economic and material well-being

Measure [1]	Controls	Early Start	p	Effect size d
% Welfare dependent				
12 months	78.7	84.3		
24 months	76.3	78.6		
36 months	71.5	70.1		
Ever (0–36 months)	86.4	89.6	.34 [3]	-.10
Mean family income per week				
12 months	$400	$407		
24 months	$443	$445		
36 months	$492	$499		
Average (0–36 months)	$443	$454	.47 [2]	.08
% Income inadequate/very inadequate				
12 months	21.9	22.3		
24 months	19.5	23.1		
36 months	15.5	17.4		
Ever (0–36 months)	37.4	41.8	.38 [3]	-.09
Mean amount of debt				
12 months	$2,541	$2,860		
24 months	$3,509	$3,362		
36 months	$4,619	$4,218		
Average (0–36 months)	$3,380	$3,582	.71 [2]	-.04
% Mother in paid employment				
12 months	20.8	18.7		
24 months	22.3	23.5		
36 months	26.6	31.5		
Ever (0–36 months)	38.8	42.9	.42 [3]	.08
% Partner in paid employment				
12 months	22.2	20.7		
24 months	25.1	22.5		
36 months	30.4	27.2		
Ever (0–36 months)	40.3	37.4	.56 [3]	-.06
% Accommodation inadequate/very inadequate				
12 months	8.3	4.0		
24 months	7.1	8.0		
36 months	8.2	12.0		
Ever (0–36 months)	20.9	19.2	.69 [3]	.04
Mean number of hardship factors				
12 months	5.7	5.5		
24 months	4.5	4.6		
36 months	4.2	4.5		
Average (0–36 months)	4.7	4.9	.60 [2]	-.05

[1] Sample sizes for different comparisons (excluding partner measures) were: 12 months (Controls N = 216, Early Start N = 198); 24 months (Controls N = 211, Early Start N = 187); 36 months (Controls N = 207, Early Start N = 184); Ever/0–36 months (Controls N = 206, Early Start N = 182).
[2] t-test for independent samples.
[3] chi-squared test.

extended throughout the 36-month follow-up period. Second, there was no consistent, significant tendency for families in the Early Start group to fare any better than those in the control group. These results are consistent with the conclusion that the provision of Early Start services failed to have any detectable effect on levels of family stability, inter-partner violence, or residential mobility. The effect sizes for these comparisons were d = -.11 to .05.

3. Family Economic and Material Well-Being

Table 6.3 compares the Early Start and control groups on a series of measures of economic and material well-being. These measures include: welfare dependence; family income; family debt; parental workforce participation; and adequacy of accommodation. Family income was assessed net of tax and included all sources of family income from welfare benefits and paid employment.

The table also includes the mean number of economic hardship factors (eg borrowed money from family or friends, unable to pay the bills, postponed visits to the doctor or dentist) the family had experienced in the past year (see Appendix).

The table shows that in both groups levels of material deprivation were high. The majority of families were welfare dependent, had low incomes, high debt and restricted material circumstances. Nonetheless, there was no evidence those in the Early Start group fared any better on these measures than those in the control group. The effect size estimates for these comparisons ranged from d = -.10 to .08.

4. Family Life Events and Stresses

At each point of assessment, respondents were questioned on a 45-item questionnaire regarding their exposure to stressful and adverse life events. Table 6.4 compares the Early Start and control groups on a series of measures relating to exposure to adverse life events. These events included: illness and death; economic and financial problems and crises; family or social relationship problems; victimisation; and other stressful life events. In addition, the table reports an overall life-

events score. The table shows that rates of adverse life events were similar for the Early Start and control groups. This finding suggests that the provision of Early Start services had no detectable effects on family susceptibility to stress, crisis, or adversity. The effect size estimates for these comparisons ranged from -.19 to .12.

6.3.2 Multivariate Tests

As with the analysis in Chapter 5, this chapter reports the results of multiple tests of statistical significance. This practice, in turn, raises the issue of conducting an omnibus test of the overall significance of the results. To address these issues the following approaches were used:

1. Plot of effect size: Figure 6.1 shows a plot of the Cohen's d values for the comparisons reported in Tables 6.1 to 6.4. The plot shows the values of Cohen's d were approximately normally distributed around a mean of -.02. This result is consistent with the conclusion that the Early Start programme had no detectable effects in the areas of maternal health and well-being, family stability, family violence, family material circumstances and family exposure to stress and crisis.

2. Multivariate regression: To take account of the effects of multiple significance testing, a multivariate regression model was fitted to the data (see Chapter 5 for details of this model). The log likelihood ratio chi-square test for the null hypothesis that the intervention was unrelated to the set of 24 outcomes analysed in this chapter was $\chi^2 (24) = 21.2$; p = .63. This test statistic confirmed the conclusion that there was no evidence of significant associations between group membership and maternal and family outcomes.

Table 6.4 Comparison between the Early Start and control groups on stressful life events and adversity

Measure [1]	Controls	Early Start	p	Effect size d
% Any illness/death in past 12 months				
12 months	44.4	51.0		
24 months	41.4	37.4		
36 months	41.6	39.7		
Ever (0–36 months)	77.1	72.0	.25 [3]	.12
% Any economic problem in past 12 months				
12 months	74.1	71.7		
24 months	68.7	71.1		
36 months	80.2	74.5		
Ever (0–36 months)	92.7	91.8	.72 [3]	.03
% Any family/social relationship problem in past 12 months				
12 months	65.3	61.1		
24 months	52.9	50.3		
36 months	47.3	54.9		
Ever (0–36 months)	82.9	81.3	.68 [3]	.04
% Any victimisation in past 12 months				
12 months	12.0	10.6		
24 months	7.1	11.2		
36 months	5.3	10.3		
Ever (0–36 months)	20.5	28.6	.06 [3]	-.19
% Any other stressful/adverse life event in past 12 months				
12 months	71.8	68.2		
24 months	69.7	72.7		
36 months	68.1	67.9		
Ever (0–36 months)	93.7	93.4	.91 [3]	.01
Mean number of life events in past 12 months				
12 months	4.2	3.9		
24 months	3.6	3.6		
36 months	3.5	3.6		
Total (0–36 months)	11.3	11.2	.91 [2]	.01

[1] Sample sizes for different comparisons were: 12 months (Controls N = 216, Early Start N = 198); 24 months (Controls N = 211, Early Start N = 187); 36 months (Controls N = 207, Early Start N = 184); Ever/0–36 months (Controls N = 206, Early Start N = 182).
[2] t-test for independent samples.
[3] chi-squared test.

Figure 6.1 Frequency distribution of effect size (Cohen's d) estimates for maternal and family outcomes

Observed mean = -.02
Expected mean (under Ho) = 0.0
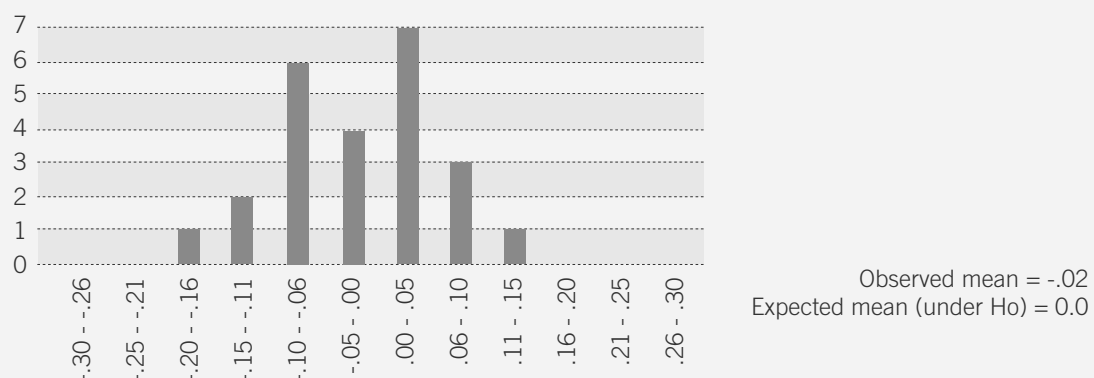
6.4 Summary and Conclusions

This chapter has examined the extent to which families enrolled in the Early Start service had outcomes that differed from the control group on a number of measures including: maternal health and well-being; family stability; family violence; family material circumstances; and family exposure to stress and crisis. In all of these areas no differences emerged. Specifically:

- Mothers in the Early Start group had similar outcomes to the control group in terms of contraception, pregnancy, depression and substance use.

- Mothers in the Early Start group had similar outcomes to the control group in the areas of single parenthood, partnership stability, exposure to domestic violence and residential mobility.

- Families enrolled in Early Start had similar outcomes to the control group in the areas of welfare dependence, family income, family debt, family living standards, parental workforce participation and family economic hardship.

- Families enrolled in Early Start had similar outcomes to the control group in the areas of exposure to stress and adversity.

These findings are generally consistent with the findings of the pilot study. The pilot study suggested that while the Early Start programme led to positive changes in the areas of child health, parenting, child abuse, early childhood education and related outcomes, the programme showed little apparent benefits in the areas of maternal and family outcomes (Fergusson et al, 1998). The pilot evaluation concluded that:

> *"Family support services are most effective in assisting mothers to acquire new skills in rearing their children but are less effective in addressing life style issues relating to substance abuse, partner relationships and family material circumstances"* (page 77).

The above conclusion is clearly consistent with the findings of this evaluation. This conclusion, in turn, raises important questions about the reasons for programme efficacy to vary in different areas. There are two possible explanations for this result. The first relates to the specific emphasis of the Early Start programme, which treats the child as the primary client and addresses issues of family context in relationship to that client. It may be that the child-centred emphasis of the Early Start programme favours the promotion of positive child-related outcomes rather than family-related outcomes. The alternative explanation relates to the potential for family change in different areas. The results of the present study could suggest the potential for family change is greater in areas of "new learning" involving child rearing and parenting practices than in areas involving long-standing personal, social and economic difficulties.

Irrespective of the reasons for a consistent lack of programme benefit in maternal or family functioning, the results of this evaluation clearly suggest home visitation programmes such as Early Start do not, and perhaps cannot, provide a complete solution to family social and economic problems. This suggests that such programmes need to be seen as one component of an integrated approach to assisting families facing stress and difficulty.

The present evaluation suggests a major problem facing families is that of restricted material and economic circumstances. As shown in Table 6.3, families in both groups had high welfare dependence, low incomes, high levels of debt, and limited material circumstances. In this population there are three factors that probably conspire to limit material circumstances. First, at any time approximately three-quarters of families were welfare dependent and their living standards were thus constrained by the levels of material well-being provided by benefit support. Second, the majority (over 80%) of families were single-parent families at some time during the 36-month study period. The absence of two parents prevented these families from achieving the economic and material benefits of two-income families. Finally, parents frequently lacked formal educational qualifications thus restricting their earning capacity. These factors (welfare dependence, single parenthood, limited education) conspired to create a situation in which such families were subject to a

"poverty trap" where welfare benefits were likely to provide a very similar level of income to that which could be obtained from full-time workforce participation. A clear challenge facing current social policy is to supplement home visitation methods such as Early Start with policies that both reduce welfare dependence and increase the earning capacity of these families.

Finally, it is important to consider these results in the context of possible threats to study validity. The previous chapter noted there were three major potential threats to trial validity. First, the open nature of the trial raised possibilities of response bias in which reports of family outcomes were biased by the family's knowledge of the intervention group to which they had been assigned. Second, it was noted that higher rates of sample loss in the assessments of the Early Start series could have biased trial results. Finally, it was suggested the results could be influenced by an investigator bias in favour of finding positive results for Early Start. A key feature of all of these threats is that they were likely to bias the results of the trial towards finding benefits for the Early Start programme. Specifically, the open nature of the trial could encourage control families to under-report difficulties, sample losses may have biased the Early Start results towards reducing rates of problems, and

investigator biases may have led to a search for findings supportive of Early Start. The results of this chapter provide further evidence to support the conclusion that these potential study biases did not exert any appreciable influence on the study findings. In particular, both this chapter and the previous chapter have used similar methods of data collection, data analysis, and data reporting but the chapters have led to quite different conclusions. Chapter 5 suggested the presence of small but pervasive benefits in the areas of child health, parenting, child abuse and neglect and child behaviour. The present chapter suggests a clear absence of association between the provision of Early Start and a range of parental and family outcomes. This specificity of association between Early Start and different types of outcomes clearly argues against the view that the results of the trial were biased as a result of the open nature of the trial, sample losses or investigator biases. This follows, since it would be expected that any bias arising from these sources would have affected the results for all outcomes and not simply those relating to child health, parenting and child behaviour.

CHAPTER 7:
Further Analysis

7.1 Introduction

The aims of this chapter are to elaborate on the analyses reported in earlier chapters. The chapter focuses on two key issues: effect modification and treatment of missing data.

7.1.1 Effect Modification

Thus far the analysis has treated those in the Early Start and control groups as homogeneous groups of families. However, it could be proposed that the benefits of the programme may vary with the type of family to which it is supplied. An important example of this issue concerns the extent to which the programme benefits were similar for Māori and non-Māori families. In particular, as noted in Chapter 1, Early Start was set up as a "mainstream" programme offered to both Māori and non-Māori in a culturally appropriate way. However, it could be suggested that the benefits of the programme may vary between ethnic groups so, for example, Māori gained less benefit from the programme than non-Māori. Similarly, previous research has suggested the benefits of family support programmes tend to be greater among high-risk families facing multiple sources of stress and difficulty than among lower risk families (Olds et al, 1994).

For these reasons it is important to explore the extent to which programme benefits may vary between different types of families including: Māori and non-Māori families; families having their first child; families with teen parents; and families facing multiple difficulties. To address each of these questions involves extending the research design to

stratify the sample by a further classificatory factor to examine the homogeneity of the treatment effects in different strata. This may be achieved by fitting statistical models that include a treatment x classificatory factor interaction term. As Gomby has discussed (Gomby, 1999), because such analysis is conducted on an ex post facto basis the conclusions drawn should be seen as being more tentative and exploratory than the overall results of the trial.

7.1.2 Treatment of Missing Data

As noted in Chapter 4, the research findings were subject to a relatively small amount of missing data due to the non-participation of respondents in the research interviews. As we showed, this missing data was not associated with respondent characteristics at the baseline assessment. For this reason we have treated the missing data as occurring at random. However, this assumption could be questioned. To address issues relating to missing data, we report further analyses in which the results are corrected for missing data using methods of missing data estimation and imputation.

7.2 Family Features and Programme Benefits

7.2.1 Ethnicity

Table 7.1 compares the Early Start and control groups on a series of outcome measures found to be significant in Chapter 5. 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; whereas 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 x ethnicity interactions, indicating that for most outcomes the results were similar for both Māori and non-Māori. Nonetheless, in two

Table 7.1   Māori ethnicity and observed outcomes to 36 months

Measure	Māori		Non-Māori		Treatment x Ethnicity
	Controls	Early Start	Controls	Early Start	
	(N = 75)	(N = 76)	(N = 131)	(N = 105)	p
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
Preschool 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

comparisons there was a significant treatment x ethnicity interaction. First, for general practitioner visits the differences for Māori were smaller than for non-Māori, suggesting in this instance programme benefits were weaker for Māori. Second, for internalising behaviour the differences for Māori were greater than for non-Māori suggesting in this instance programme benefits were stronger for Māori.

Closer inspection of 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 the 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, while for non-Māori the range of d values was .08 to .40, with a median of .21.

7.2.2 Parity

Table 7.2 compares the Early Start and control groups on the same series of outcomes that were analysed in Table 7.1. In Table 7.2, the sample is stratified by maternal parity and compares the outcomes for women having their first child with women who had more than one child.

Examination of the table shows that in all but one comparison there were no significant treatment x parity interactions. The results show that, among women having their first child, those enrolled in Early Start had lower rates of immunisation (89% vs 95%) whereas among those with two or more children those enrolled in Early Start had higher rates of immunisation (96% vs 91%). However, leaving aside this result, the table suggests the benefits for families having their first child were similar to those for families having their second or further children. This conclusion was confirmed by comparing Cohen's d for all outcomes for both groups. Effect sizes varied from .11 to .41 with a median of .24 for first child families and .13 to .32 with a median of .21 for families with more than one child.

Table 7.2 Parity and observed outcomes to 36 months

Measure	Primiparous		Multiparous		Treatment x Parity
	Controls	Early Start	Controls	Early Start	
	(N = 109)	(N = 99)	(N = 97)	(N = 82)	p
Child Health					
% Up to date with well-child checks by 36 months	25.7	45.5	35.1	41.5	.16
Mean GP visits by 36 months	20.3	23.1	21.2	23.6	.91
Hospital Attendance					
% Attended hospital for accident/injury by 36 months	26.6	17.0	26.0	18.8	.77
Preschool Dental Care					
% Enrolled with dental service at 36 months	63.3	68.7	62.9	76.8	.33
Early Childhood Education					
Mean duration of early childhood education (months)	13.9	16.9	13.4	15.9	.82
Parenting					
Mean positive parenting attitudes (36 months)	9.90	10.08	9.86	10.20	.45
Mean non-punitive attitudes (36 months)	9.87	10.10	9.92	10.13	.90
Mean total parenting score (36 months)	9.88	10.10	9.88	10.18	.67
Child Abuse and Neglect					
% Severe/very severe physical assault	11.9	4.0	11.3	4.9	.76
Child Behavioural Adjustment					
Mean externalising score (36 months)	10.09	9.92	10.09	9.86	.79
Mean internalising score (36 months)	10.05	9.73	10.20	10.01	.52
Mean total overall score (36 months)	10.09	9.84	10.14	9.89	.99

7.2.3 Age

Table 7.3 shows the data in Tables 7.1 and 7.2 re-analysed by stratifying the sample on the basis of age into those under 20 years and those 20 years and over. Although there were no significant treatment x age interactions, inspection of the table suggested that programme benefits tended to be smaller for those families with mothers aged under 20 years. This conclusion was confirmed by comparing Cohen's d for all outcomes for both groups. The values of d varied from -.04 to .38 with a median of .16 for women under 20 years and .16 to .30 with a median of .29 for women 20 years and over.

7.2.4 Disadvantage

To examine the extent to which programme benefits were modified by the extent of family disadvantage, a family disadvantage score was constructed by summing up the number of family disadvantages present at the baseline assessment. Fourteen disadvantage factors were included in the analysis. These disadvantages included measures of welfare dependence, maternal childhood stress and difficulty, exposure to child abuse, and exposure to partner violence.

Table 7.4 shows the sample classified into two groups by dichotomising the family disadvantage score at the median value. The table shows that in all comparisons there were no significant treatment x disadvantage interactions, indicating that for most outcomes the results were similar for all levels of disadvantage.

Closer inspection of the table shows the presence of a small but pervasive trend for benefits to be larger for those in the high disadvantage group. For the high disadvantage group Cohen's d ranged from .07 to .37 with a median of .31, whereas for the low disadvantage group Cohen's d ranged from .03 to .33 with a median of .22.

7.3 Treatment of Missing Data

Previous chapters of this report have compared the Early Start and control groups on all available data at each point of observation. This method of analysis assumes implicitly that missing observations

Table 7.3 Maternal age and observed outcomes to 36 months

Measure	Age <20		Age 20+		Treatment x Age
	Controls	Early Start	Controls	Early Start	
	(N = 61)	(N = 51)	(N= 145)	(N= 130)	p
Child Health					
% Up to date with well-child checks by 36 months	23.0	37.3	33.1	46.2	.77
Mean GP visits by 36 months	18.0	21.5	21.9	24.0	.60
Hospital Attendance					
% Attended hospital for accident/injury by 36 months	19.7	18.4	29.2	17.6	.32
Preschool Dental Care					
% Enrolled with dental service at 36 months	59.0	56.9	64.8	78.5	.10
Early Childhood Education					
Mean duration of early childhood education (months)	15.6	15.8	12.8	16.6	.20
Parenting					
Mean positive parenting attitudes (36 months)	10.03	10.13	9.82	10.13	.36
Mean non-punitive attitudes (36 months)	9.88	10.0	9.90	10.16	.55
Mean total parenting score (36 months)	9.94	10.07	9.85	10.16	.41
Child Abuse and Neglect					
% Severe/very severe physical assault	6.6	2.0	13.8	5.4	.85
Child Behavioural Adjustment					
Mean externalising score (36 months)	10.23	9.96	10.03	9.87	.62
Mean internalising score (36 months)	10.0	9.83	10.18	9.88	.55
Mean total overall score (36 months)	10.19	9.91	10.08	9.85	.83

on the outcome variables occurred "at random". While this assumption was justified on the grounds missing data was not related to the characteristics of the sample at baseline, it is clearly worth considering the extent to which missing data may have threatened experimental validity. One way of addressing this issue is to use missing data estimation to estimate the study results that would have been observed had all study participants been observed at all time periods. In this instance, the estimation of missing data was conducted in the following way:

1. For the 12 participants (11 Early Start; 1 Control) who were not assessed at baseline and who had missing data on 36-month outcomes, scores on each of the outcome variables were estimated by setting the outcomes for these participants equal to the mean score on the relevant variable for the control group.

2. For the remaining 43 participants (27 Early Start, 16 Controls) who completed the baseline interview but did not have complete outcome data, outcomes up to 36 months were estimated using missing data imputation methods. These methods involved a two-stage process. In the first stage, the data for participants with non-missing data at baseline and 36 months was used to derive regression models to predict each outcome at 36 months on the basis of baseline characteristics. These baseline factors included the broad range of family demographic, socio-economic, maternal childhood, pregnancy and related characteristics of the sample described in Chapter 4. In the second stage, the coefficients from the fitted regression models were then used to impute the value of each outcome for those with missing data on 36-month outcomes on the basis of their known baseline characteristics. Regression imputation of missing values was conducted using the missing data imputation procedures in the Stata 8 statistical package (StataCorp, 2003).

The missing data estimation methods described above gave, for each participant

Table 7.4 Extent of family disadvantage and observed outcomes to 36 months

Measure	High disadvantage		Low disadvantage		Treatment x Disadvantage
	Controls	Early Start	Controls	Early Start	
	(N = 89)	(N = 87)	(N = 117)	(N = 94)	p
Child Health					
% Up to date with well-child checks by 36 months	20.2	34.5	37.6	52.1	.76
Mean GP visits by 36 months	22.7	23.6	19.2	23.0	.24
Hospital Attendance					
% Attended hospital for accident/injury by 36 months	36.0	17.7	19.0	18.0	.07
Preschool Dental Care					
% Enrolled with dental service at 36 months	58.4	67.8	66.7	76.6	.84
Early Childhood Education					
Mean duration of early childhood education (months)	13.3	17.8	14.0	15.1	.19
Parenting					
Mean positive parenting attitudes (36 months)	9.71	10.01	10.01	10.24	.71
Mean non-punitive attitudes (36 months)	9.90	10.11	9.89	10.12	.94
Mean total parenting score (36 months)	9.77	10.07	9.95	10.20	.82
Child Abuse and Neglect					
% Severe/very severe physical assault	14.6	5.8	9.4	3.2	.89
Child Behavioural Adjustment					
Mean externalising score (36 months)	10.33	9.99	9.91	9.80	.25
Mean internalising score (36 months)	10.17	9.82	10.09	9.91	.39
Mean total overall score (36 months)	10.32	9.93	9.96	9.81	.22

enrolled in the trial, a complete set of data up to the 36-month follow-up. Using this information it was possible to compare the Early Start and control groups using a strict intention to treat design that included all respondents enrolled in the trial.

Table 7.5 compares the results obtained from analysing: the observed data; and the observed data where available supplemented by missing data estimates as described above. For each outcome measure and each data set the table reports the effect size as measured by Cohen's d, and the results of the test of significance.

Inspection of the results obtained from the observed data and the observed data supplemented by missing data estimates suggests both analyses led to very similar conclusions. For the observed data, the values of Cohen's d ranged from .19 to .27 with a median value of .24 whereas for the data supplemented by missing data estimates the values of Cohen's d ranged from .16 to .27 with a median value of .23.

Closer inspection of the results suggests that adjustment for missing data tended to increase some associations slightly and

reduce others slightly. These changes resulted in changes of the significance levels associated with two comparisons. First, on the basis of the observed data, enrolment with a dental nurse/dentist was significant but became non-significant (p = .10) after correction for missing data. Second, on the basis of observed data, the mean externalising score was marginally significant (p = .06) but became significant (p < .03) after correction for missing data. In general, correction for missing data led to similar conclusions to those found for observed data. This suggests it is unlikely that missing data was a threat to study validity.

Table 7.5 Supplementary analysis of missing data

Measure	Observed data		Adjusted for missing data	
	p	d	p	d
Child Health				
% Up to date with well-child checks by 36 months	<.05	.25	<.01	.27
Mean number of GP visits by 36 months	<.05	.24	<.05	.19
Hospital Attendance				
% Attended hospital for accident/injury by 36 months	<.05	.22	<.05	.21
Preschool Dental Care				
% Enrolled with dental nurse/dentist at 36 months	<.05	.20	.10	.16
Early Childhood Education				
Mean duration of early childhood education (months)	<.05	.22	<.03	.21
Parenting				
Mean positive parenting attitudes (36 months)	<.01	.26	<.02	.24
Mean non-punitive attitudes (36 months)	<.05	.22	<.04	.20
Mean total parenting score (36 months)	<.01	.27	<.02	.24
Child Abuse and Neglect				
% Severe/very severe physical assault (0–36 months)	<.01	.26	<.02	.24
Child Behavioural Adjustment				
Mean externalising score (36 months)	.06	.19	<.03	.21
Mean internalising score (36 months)	<.01	.26	<.02	.24
Mean total behaviour score (36 months)	<.05	.24	<.01	.25

7.4 Summary and Conclusions

The first topic examined in this chapter concerned the extent to which programme benefits varied with family characteristics including: ethnicity, parity, maternal age and family disadvantage. A commentary on the key findings and their implications is given below.

7.4.1 Ethnicity

The findings of the study suggested that, overall, there were few detectable differences in the outcomes for Māori and non-Māori families. Nonetheless, there was a small but pervasive tendency for Māori families to receive greater benefits than non-Māori families. This was evident in the 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 of relevance in the context of ongoing debates about the relative contributions of mainstream and Māori programmes. In particular, there have been strong claims made that 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, McGregor, Coggan, Bennett & McKenzie, 2000; Ministry of Health, 1998). The results of this evaluation of Early Start are not consistent with this view and suggest that mainstream programmes may in fact deliver similar if not greater benefits to Māori clients. However, it should be noted that the programme has been evaluated using mainstream measures of health and well-being. It may be proposed that programmes run by Māori for Māori offer additional cultural and spiritual benefits (Durie, 1998).

The benefits of Early Start for Māori may reflect a number of features of the design and implementation of the Early Start programme. In particular, in setting up the Early Start programme 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 ongoing process that involved: initial consultation with Māori about the programme design and directions; the establishment of a Board on which approximately 50% of members were Māori; investment in cultural training for all workers; and the employment of Māori staff. It would appear this combination of processes resulted in an organisational environment that produced outcomes for Māori families 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.

7.4.2 Parity, Age and Family Disadvantage

There have been suggestions in the literature on home visitation programmes that the benefits of the programme may vary with a number of client characteristics including the age of parents, family size, and the extent of family disadvantage (Gomby et al, 1999; Olds et al, 1994). In general, it has been suggested that programmes may be more beneficial for younger parents, parents having their first child and parents facing multiple disadvantages. The present study failed to find conclusive evidence linking any of these factors to improvements in programme outcomes. In general, outcomes were similar for primiparous and multiparous parents. There were, however, some suggestions the programme had greater benefits for older parents and those families facing multiple disadvantage. However, tests of interaction failed to show these differences to be significant. These findings suggest that, in general, the Early Start programme had similar benefits for client families regardless of parental age, family size or level of family disadvantage.

7.4.3 Missing Data

A potential threat to study validity arose from the fact there was a certain amount of missing data arising from: clients declining to enter the trial following randomisation; and failure to interview all participants at all times. Although the trial had complete data for approximately 90% of families enrolled in the trial, the presence of missing data poses a potential threat to trial validity. In the preceding chapters we have analysed the observed data on the assumption that data was missing at random. However, that assumption could be questioned. To address this issue, this chapter reported a re-analysis of the key findings using methods of missing data estimation. This analysis led to very similar conclusions to the main analysis suggesting that missing data did not pose a threat to study validity.

CHAPTER 8:
Summary, Conclusions and Recommendations

8.1 Introduction

The aims of this chapter are to provide an overview of the findings of this randomised trial of Early Start by examining a series of issues. These issues include:

• A summary of the overall findings of the trial.

• Consideration of threats to trial validity.

• Comparison of findings with results from both international and New Zealand studies of home visitation.

• An examination of issues for the future development and evaluation of Early Start.

8.2 Summary of the Overall Findings of the Randomised Trial

The results of the randomised trial fall neatly into two sets of results; with one set of results suggesting positive benefits for some outcomes and the other set of results suggesting no benefit for other outcomes. The outcomes for which positive findings were noted included: child health; early childhood education; parenting behaviours; child abuse and neglect; and child behaviour. There are two features that distinguish these outcomes. First, all concern high priority areas targeted by Early Start and second, all concern child-related rather than family-related outcomes. The weight of the evidence suggests the Early Start programme delivered small but consistent benefits in a number of areas relating to child health, education, child abuse, parenting and behavioural outcomes.

In contrast, an absence of benefit was noted in other areas including: maternal health; family economic conditions; family violence; and family stress. These were also areas targeted by Early Start but which had a lower priority than the child-related outcomes. What the results of the trial suggest is that, while the Early Start programme offered benefits for children, the impact of the programme on a wide range of family level factors was negligible.

There are two possible explanations of the differences in the child-related and family-related outcomes of this trial. The first is that these differences may reflect differences in the priority family support workers assigned to child-related and family-related outcomes. It may be that because the Early Start programme was targeted at children rather than families the effects on child-related outcomes were more marked than family-related outcomes. Alternatively, it may be that the differences reflect differences in the extent to which change is possible. This explanation was proposed in a discussion of the pilot study preceding this trial (Fergusson et al, 1998). In that discussion it was suggested programmes such as Early Start may be better at promoting "new learning" in areas relating to child health and development than in addressing long-standing personal, financial and related problems (Fergusson et al, 1998). Which of these explanations (if either) is correct is unknown, but the results of this trial clearly suggest a need for further programme development to examine the extent to which home visitation methods may lead to family level change.

A pervasive feature of the results was that the benefits of the Early Start programme tended to be small, with the mean effect size assessed by Cohen's d being .16. This falls into the range described as a small effect size (Cohen, 1977). This tendency for home visitation programmes to produce small effects has been discussed in a number of reviews of evaluations of these programmes (Gomby et al, 1999; Olds & Kitzman, 1993). An explanation of the small size of effect is that this may reflect the nature of both the client population and programme evaluation. Randomised trials of home visitation

programmes such as Early Start differ from standard clinical trials in two important ways. First, those enrolled in such trials are not a homogeneous population experiencing a common set of issues and problems. Rather, they represent a heterogeneous set of families characterised by a broad spectrum of factors that expose them to stress and difficulty. Second, the programme supplied does not provide each family the same treatment and set of experiences; rather the programme is tailored to fit the family circumstances. The net result of these factors is the evaluations of home visitation programmes describe the effects of a heterogeneous treatment method applied to a heterogeneous population. Under these circumstances one would not expect to find a large effect size for a specific outcome. Rather, it would be expected that programme benefits would be evident in a pattern of small pervasive benefits such as those found in this study.

This analysis was then extended to examine the extent to which programme benefits varied with family characteristics including ethnicity, parental age, family size and family disadvantage. This analysis showed the programme benefits did not vary to any great extent with these factors. However, there was some suggestion the programme offered greater benefits to Māori, older mothers, and families facing high levels of disadvantage. As we have noted in the previous chapter, the findings for Māori have implications for debates about the role of mainstream and Māori programmes in the delivery of services to Māori. The findings of the Early Start programme provide a clear illustration of the fact that appropriately designed mainstream programmes can offer similar benefits to Māori and non-Māori. At the same time it remains possible that programmes designed by Māori for Māori may offer cultural and spiritual benefits not provided by mainstream programmes (Durie, 1998).

8.2.1 Threats to Trial Validity

While the weight of the evidence from this study suggests Early Start had beneficial consequences for the children enrolled in the programme, this evidence needs to be

weighed in the light of a number of potential threats to trial validity. These threats are discussed below.

1. Client recruitment: A potential limitation of this trial was that not all clients eligible for the trial agreed to enter the trial. As shown in Chapter 4, approximately three-quarters of those eligible for the trial agreed to enter and a quarter declined. Furthermore, because of privacy issues it was not possible to ascertain the characteristics of those who declined to enter the trial. For this reason the extent to which those agreeing to enter the trial were representative of all families eligible to enter the trial is unknown. This feature of the study design limits the extent to which the findings can be generalised to families other than those in the trial but does not call the internal validity of the trial into question.

2. Study design: Of necessity, the Early Start trial used an open design in which participants were aware of the treatment group to which they belonged. Furthermore, many of the outcomes were assessed using participant reports. This feature of the evaluation poses a potential threat to validity to the extent that it could be suggested differences between the Early Start and control groups may have reflected differences in parental reporting behaviours rather than differences in child and family outcomes. However, this conclusion is not consistent with the fact that differences between groups were also found for general practitioner and hospital record data (see Chapter 5). Such findings argue against the view that differences in outcomes are solely a reflection of the effects of Early Start on the ways in which parents reported child and family outcomes.

3. The experiences of the control group: Community-based randomised trials differ from clinical trials in that the control group are not an "untreated" group but rather those who have access to the mix of health, education and social services available in the community at the time of evaluation. Such designs thus assess the extent to which a particular programme improves outcomes over and above any benefits provided by existing services. It follows from this that the extent to which the control group is exposed to alternative services will influence the outcomes of a randomised trial. In the present context, the Early Start service was evaluated in a community that has a generally high level of health, education and welfare services devoted to the well-being of children and families. Under these circumstances, the research design is biased against finding programme benefits. The fact Early Start showed evidence of benefits in a number of areas in such an environment would tend to reinforce the view that the programme is effective.

4. Variation in duration of service: As shown in Chapter 4, there was wide variation in the amount of service received by those enrolled in Early Start, ranging from those who declined to enter the service to those who remained in the service for 36 months, with the median duration of service participation being 24 months. To address these issues the present trial has analysed the results using an "intention to treat" paradigm in which results were analysed for all participants irrespective of the duration of service provision. This approach preserves trial validity but may lead to conservative estimates of the effect sizes that would have been observed had all families remained in the trial throughout the 36-month period.

5. Missing data: There was a relatively high level of sample retention in the research follow-up of the trial participants: at 36 months just under 90% of trial participants were interviewed. However, this response rate implies that data was incomplete for 10% of trial participants. It could be suggested missing data may have biased the trial results towards (or against) the research hypotheses. To address this issue, missing data estimation methods were used to estimate outcome measures for all trial participants. This re-analysis produced conclusions that were similar to the results for the observed trial data, suggesting that missing data was not a major threat to trial validity.

8.3 Comparisons with International and New Zealand Studies of Home Visitation

To place the findings of the Early Start trial in a more general context, it is useful to compare these findings with other evaluations of home visitation programmes conducted throughout the world. To achieve this, the findings of the Early Start trial were contrasted with the findings of a series of randomised trials of home visitation programmes, as described in the literature review in Chapter 1. This comparison is made in Table 8.1, which is an extension of Table 1.1 with the inclusion of additional findings from Australasian programmes including Early Start, a programme of home visitation in Queensland, Australia (Fraser et al, 2000), and the New Zealand Parents as First Teachers pilot trials (Livingstone, 1998).

As described in the literature review, outcomes are described by a simple scoring system in which + indicates the programme reported at least one statistically significant benefit for the outcome domain, - indicates no benefits were found, and * indicates the outcome was not assessed.

Inspection of Table 8.1 indicates that, for the great majority of programmes, benefits were either not found or data was not reported. The exception to this trend is the Nurse Family Partnership Program which has found a similar array of benefits to those found for Early Start. This comparison raises the interesting feature of the factors common to both programmes that may have encouraged programme success. Comparison of Early Start with the Nurse Family Partnership Program shows these programmes differ in a large number of

Table 8.1 Findings from randomised trials of home visitation

Programme	Outcomes					
	Child Abuse and Neglect	Child Health	Parenting	Utilisation of Preschool Education	Child Behaviour	Maternal Life Course
Early Start	+	+	+	+	+	-
International Research						
Nurse Family Partnership						
Elmira [1]	+	+	+	*	*	+
Memphis [2]	+	-	+	+	-	+
Hawaii Healthy Start [3]	-	+	-	*	*	-
Healthy Families America						
Overall findings [4]	-	*	+	*	*	+
Parents As Teachers [5]						
Northern California	*	-	-	*	*	-
Southern California	+	+	-	*	*	-
Comprehensive Child Development Program [6]						
Overall findings	*	-	-	*	-	-
Australasian Research						
Queensland trial [7]	+	-	-	*	*	-
Parents as First Teachers[8]						
Dunedin	*	-	-	-	-	-
Gisborne/East Coast	*	-	-	+	-	-
South Auckland	*	-	-	-	-	-
Whangarei	*	-	+	+	+	-

[1] (Olds et al, 1988; Olds et al, 1994; Olds et al, 1999).
[2] (Kitzman, Olds et al, 1997; Kitzman et al, 2000; Olds et al, 1999; Olds, Kitzman et al, 2004).
[3] (Duggan, Fuddy et al, 2004; Duggan et al, 1999; Duggan et al, 2000).
[4] (Daro & Harding, 1999).
[5] (Wagner et al, 1996; Wagner & Clayton, 1999; Wagner et al, 1997).
[6] (St Pierre & Layzer, 1999).
[7] (Fraser et al, 2000).
[8] (Livingstone, 1998).

respects regarding programme philosophy, client recruitment and service delivery methods. However, there are three points of similarity that may be relevant to the apparent success of these programmes.

First, both programmes evolved in a research context in which the development of the programme and the design of the research were conducted by the same group of researchers and service providers (Olds et al, 1999; Olds, Kitzman, Cole, & Robinson, 1997). This aligning of the process of programme development with the process of research evaluation may have led to improved programme efficacy. This view is supported by a considerable literature which suggests programmes and treatments that are developed and evaluated within a research context tend to produce improved outcomes (Daro & Harding, 1999; Gomby, 1999; Olds, 1992; Olds et al, 1999; Olds & Kitzman, 1993). This may be due to the fact research involvement leads to more theoretically-driven and evidence-informed programmes that maximise the chances of finding positive programme benefits.

Second, both programmes employ professionally trained staff rather than paraprofessional or lay staff. The Nurse Family Partnership Program uses qualified nurses whereas Early Start uses either nurses or social workers. There is now evidence from controlled experiments that suggests programmes using professionally trained workers produce better outcomes than programmes using paraprofessionals (Olds & Kitzman, 1990; Olds et al, 2002).

Finally, both programmes have invested substantially in mechanisms to ensure fidelity of programme delivery. In Early Start, programme fidelity was maximised by a series of procedures that include: the development of clearly stated and operationalised programme goals; regular weekly supervision of staff to assess goals, directions and practice; and the development of databases to monitor key outputs. In the Nurse Family Partnership Program, fidelity of programme delivery has been maximised by a series of approaches including: carefully constructed and detailed programme protocols and programme objectives; and

regular clinical supervision by highly qualified and experienced clinical workers.

It is our view that it is probably this combination of factors (research involvement, professional staff and high commitment to programme fidelity) that has led the Early Start and the Nurse Family Partnership models to produce positive client outcomes.

8.4 Future Development of Early Start

The development of the Early Start programme is clearly work in progress. Throughout this randomised trial there was successive refinement and improvement of the programme as a result of growing experience in engaging with and delivering services to an often difficult to reach population. The net result of this process of evolution is that the present day version of Early Start probably is more effective and provides greater benefits than the programme developed in the 1990s. Ideally, it would be desirable to conduct a further trial to examine the extent to which various programme enhancements have improved the outcomes of the service. Furthermore, there are a number of areas in which further information or development is required. These include:

1. Longer-term benefits of Early Start: While the present trial has shown Early Start has short-term benefits in a number of areas, the extent to which these early benefits will translate into longer-term benefits is unknown. However, evidence from the Elmira trial of the Nurse Family Partnership Program suggests that early programme benefits have been sustained for up to 15 years with positive outcomes being evident in a number of areas including: behavioural adjustment (which was not measured in the early evaluations of the trial) and delinquency; child abuse and neglect; and maternal life course. At the present time, further evaluations of Early Start are planned up to the age of six years. An important extension of the study data collection will involve information from school teachers on school readiness, school performance, and behaviour at school. Since teachers

are unlikely to be aware of the participation of families in the Early Start group, this assessment will provide a measure of outcomes that is "blind" to programme participation.

2. Programme engagement and dropout: Although programme participants spent a median duration of 24 months in the Early Start programme, over time there was a progressive loss from the programme due to moving from Christchurch, programme withdrawal, and reluctance to continue with the programme. As noted earlier, these losses impose limitations on the potential success of the programme. Currently, work is being undertaken to examine the extent to which client engagement can be increased and rates of programme participation increased. Programme extensions in this area include an initial visit by supervisory staff to introduce the Early Start programme, and consideration of the employment of specialist staff to work with difficult to reach families.

3. Child and family outcomes: What emerges very clearly from this evaluation is the different effects of the programme on child and family outcomes, with the programme showing benefits for child-related but not family-related outcomes. The reasons for this difference remain unclear but one possibility is that the effectiveness of programme delivery differed for child-related and family-related outcomes. To explore this issue further, work is currently being conducted on the development of targeted approaches aimed at reducing rates of parental depression and unplanned pregnancy. This work may clarify the extent to which interventions targeted at parent-related outcomes may have beneficial effects.

8.5 Concluding Comments

The overall findings from this randomised trial clearly support the view that this programme was effective in producing positive outcomes in the areas of child health, early childhood education, service utilisation, parenting, child abuse and neglect, and behavioural adjustment. These programme benefits did not vary with client characteristics and the results suggest the outcomes for Māori and non-Māori participants were similar. The results of the trial compare favourably with both the international literature and local literature on home visitation programmes.

Despite the apparent success of the Early Start programme, a number of issues remain to be resolved. These include: evaluation of the longer-term outcomes of children and families enrolled in the programme; improvements in levels of programme engagement; and examination of the extent to which programme benefits can be extended to parental and family-related outcomes.

Finally, we would conclude with a methodological comment. While it has generally been recognised randomised trials provide the most powerful evidence of programme efficacy (Chaffin, 2004; Gomby, 1999), a growing number of evaluation theorists has argued this methodology has limitations that reduce the effectiveness of the randomised trial for evaluating community-based interventions (for example, Donaldson & Christie, 2004; Farquhar, 2003; Livingstone, 1998; Swerissen, 1999). These limitations include: ethical issues; problems of client compliance; the nature of the control population; and programme attrition – which may conspire to limit the extent to which such trials may be informative. From the standpoint of the evaluation group, one of the more important aims of our work on this evaluation was to provide an illustration of the way a well-designed and conducted randomised trial could provide informative results on programme effectiveness. We will leave the reader to judge how successful our efforts have been.

APPENDIX:
Selection And Measurement Of Trial Outcomes

Introduction

As noted in the text of this report, the Early Start and control groups were assessed at baseline and six, 12, 24 and 36 months following trial enrolment. At each point extensive data was gathered on child and family health, well-being and related outcomes. An important stage of the development of the evaluation described in this report was that of abstracting relevant measures of programme outcomes from the database of the study. One approach to this problem could have been to adopt an exploratory strategy by examining the relationships between a wide array of measures of child health and parental and family functioning to identify areas in which the Early Start programme had benefits. The risks of such "data dredging" methods in the analysis of randomised trials have been well documented (see, for example, Chan, Hrobjartsson, Haahr, Gotzsche & Altman, 2004; Lord, Gebski & Keech, 2004; Schulz & Grimes, 2005), and research methodologists have long advocated the specification of predetermined trial outcome measures (see, for example, Begg et al, 1996; Moher, Schulz & Altman, 2001). However, this approach also has limitations when applied to randomised trials of community-based interventions, since these interventions are often targeted at producing changes in multiple domains rather than for a specific outcome. These demands require that evaluations focus on a range of outcomes rather than a single well-specified outcome measure.

To address these competing demands to avoid thoughtless data dredging while at the same time assessing outcomes relevant to the trial, we adopted a strategy of assessing

multiple outcome measures in a number of domains of child and family functioning targeted by Early Start. These domains and the corresponding measures are listed below.

Child Outcomes

Child Health

The Early Start programme had developed a number of specific goals in this area. These included: effective and timely use of general practitioner services; increased use of immunisation and well-child services; improved levels of home safety; reduction in child accidents and poisonings; and increased use of preschool dental services. To examine the extent to which these goals were achieved, a series of measures was constructed. These measures included:

1. Frequency of visits to the family doctor assessed at six, 12, 24 and 36 months: This measure was based on parental reports of the number of visits made to the family doctor since the preceding interview at each point of assessment. Parental reports were used rather than general practitioner records, as initial investigations suggested it would be difficult to obtain a comprehensive history of medical attendance from family doctor records because of the complexities of checking records from multiple practitioners.

2. Immunisation: General practitioner records were used to assess whether or not children had received each of the following immunisations: six-week Hepatitis, DTPH, Polio; three-month Hepatitis, DTPH, Polio; five-month Hepatitis, DTPH, Polio; and 15-month DTPH, MMR. It was found possible to do this since the child's history of immunisation could be obtained from their current medical practitioner.

3. Well-child checks: Parallel to data collection on immunisation, data was also gathered from family doctor records on whether the child had received the following well-child checks: six-week check; three-month check; and nine-month check.

4. Hospital admissions and attendance: Signed consent was obtained from families for the research group to have access to hospital record data. This permission was provided by 98% of families. Using these consents, hospital record data was gathered on hospital attendances up to the 36-month follow-up. Hospital visit data was then coded to record the total number of medical attendances and the number of medical attendances for accidents and poisoning. All records were also scrutinised to identify cases of hospital attendance for verified cases of abuse or neglect.

5. Home safety: Throughout the follow-up, parents were questioned about their use of various home-safety features. These features included: working smoke alarms; plug protectors on electric sockets; childproof latches on cupboards; locked medicine cabinet/safe storage of medicines; safe storage for poisons; hot water temperature set at 60˚C or lower; fireguards for open fires, log burners, or heaters; and an escape route planned in case of fire. For each family, a home-safety score was constructed at each point of assessment by summing up the number of safety features present in the home.

6. Preschool dental care: To determine the extent to which families used free preschool dental services, parents were asked at the 36-month follow-up whether they had enrolled their child with preschool dental services.

Early Childhood Education

An important goal of the Early Start programme was to encourage early and consistent uptake of early childhood education services. These services included playgroup, playcentre, kohanga reo, kindergarten, child care centres/creche, and early intervention centres. To assess the utilisation of early childhood education, parental reports of attendance at such services were gathered from the 12-month follow-up onwards. This information was used to construct a series of indices of early childhood educational participation. These

included: the age at which the child was first enrolled in early childhood education; whether the child was attending early childhood education at the 12, 24 and 36 month assessments; and the estimated duration of involvement in early childhood education.

Parenting

A major goal of the Early Start programme was to increase rates of positive parenting practices and to reduce rates of punitive parenting. To assess parenting practices, parents were questioned on a 49-item measure that combined items from the Child Rearing Practices Report (CRPR, Block, 1981; Dekovic et al, 1991), the Adult-Adolescent Parenting Inventory (AAPI, Bavolek & Keene, 1999; Hanson, 1990) and custom-written items. Factor analysis of these items suggested they measured two correlated dimensions of parenting attitudes. The first dimension was defined as positive parenting attitudes and included such items as "my child and I have warm close times together", and "I enjoy being with my child for long periods". The second dimension was defined as punitive parenting attitudes and included such items as "smacking teaches children right from wrong", and "with some children, smacking is the only thing that will work". For ease of comparison, all measures were scaled to a mean of 10 with a standard deviation of one. For consistency, items on the punitive attitudes dimension were reversed to follow the convention that a high score implies declining problems and a low

score reflects increasing problems. The reliabilities of these scales over the three time periods, assessed using coefficient alpha, ranged from .73 to .89.

Child Abuse and Neglect

A key goal of Early Start was to reduce child abuse and neglect through: reduced use of severe/very severe physical assault by parents; increased awareness by parents of child abuse and neglect issues; and reduced agency contact for child abuse and neglect. A series of five indices was constructed to assess the impact of Early Start on child abuse and neglect outcomes:

1. Parental reports of severe/very severe physical assault: This was 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. The items included: "hit him/her with a fist or kicked him/her hard"; "grabbed him/her around the neck and choked him/her"; "hit him/her over and over as hard as you could"; "burned or scalded him/her on purpose"; "slapped him/her on the face, head or ears"; "threw or knocked him/her down"; "hit him/her on some other part of the body besides the bottom with something like a belt, hairbrush, a stick or some other hard object"; and "shook him/her". Parents were classified as engaging in severe/very severe physical assault if they reported at least one incident of assault for any parent over the previous 12-month period.

2. Parental reports of agency contact for issues relating to child abuse and neglect: At six, 12, 24, and 36 months parents were asked to describe their contacts with the Department of Child, Youth and Family Services (CYFS) over the period since the previous assessment. These accounts were then analysed to identify those contacts that indicated actual or suspected incidents of child abuse and neglect.

3. Hospital admissions: For each child, copies of hospital records were obtained

from all hospitals the child was known to have attended. Hospital record data provided an account of each admission and visit made by the child to that hospital. An analysis of the records revealed seven cases in which children had been admitted to hospital with clear signs of abuse or neglect.

Child Behaviour

Children's behavioural adjustment was assessed at 36 months using 50 items from the Infant Toddler Social and Emotional Assessment scale (ITSEA, Briggs-Gowan & Carter, 1998). These measures spanned a series of behavioural dimensions including:

- Activity (eg "s/he is restless and can't sit still"; "s/he goes from toy to toy faster than other children").

- Aggression/defiance (eg "s/he acts aggressive when frustrated"; "s/he throws or pushes away things s/he doesn't want").

- Peer aggression (eg "s/he is mean to other children on purpose"; "s/he takes toys away from other children").

- Emotional negativity (eg "s/he cries a lot"; "s/he is irritable or grouchy").

- Inhibition/separation problems (eg "s/he is very clingy"; "s/he cries or hangs on to you when you try to leave").

- Depression/withdrawal (eg "s/he has less fun than other children"; "s/he seems withdrawn").

These dimensions were further categorised into two overall scores of child behavioural adjustment. The first, "externalising behaviours", included the dimensions of activity, aggression/defiance, peer aggression and emotional negativity. The second, "internalising behaviours", included the dimensions of inhibition/separation and depression/withdrawal. Scale scores corresponding to these dimensions were constructed by summing the scale items using the scoring rule suggested by Briggs et al (1998). For ease of interpretation, all measures have been scaled to a mean of 10 with a standard deviation of one. Scoring followed the convention that a higher score implies greater behavioural problems. The

reliabilities of the resulting scales, assessed using coefficient alpha, ranged from .47 to .85, with a median of .75.

Cognitive Outcomes

A further goal of the Early Start programme was to provide positive childhood outcomes in the area of cognitive ability. To assess intelligence at age three, the Wechsler Preschool and Primary Scale of Intelligence (WPPSI, Wechsler, 2002) was administered by trained examiners. The WPPSI was originally designed to be used with lower-class preschool children and consists of three subscales; Verbal IQ, Performance IQ, and Full Score IQ (a composite score of the verbal and performance subscales). Spearman Brown split-half reliabilities of these scales were .86 for the performance scale, .93 for the verbal scale, and .94 for the total scale.

Family-Related Outcomes

The Early Start programme had specific goals regarding maternal and family outcomes. These goals correspond to goals adopted in other home visitation programmes such as the Nurse Family Partnership Program, and included: promoting maternal health and well-being; reducing and preventing family violence; increasing family material well-being; and reducing family susceptibility to stress and crisis.

To assess these outcomes, the following measures were used:

Maternal Health and Well-Being

1. Contraceptive use: At each assessment, parents were asked if they were currently sexually active and, if so, whether they were currently using any form of contraception.

2. 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.

3. 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.

4. Maternal substance use (alcohol, tobacco and other drugs): At each assessment, parents were questioned concerning 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 (World Health Organization, 1993) relating to alcohol and drug abuse and dependence. On the basis of this information, parents were described as having alcohol and/or substance use problems if they said yes to any abuse or dependence item.

Family Stability, Family Relationships and Family Violence

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 asked if they had separated from a resident partner and how many changes of address they had had in the period since the previous assessment.

At each assessment parents were also 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 psychological or physical assault by any partner in the period since the previous assessment.

Family Economic and Material Well-Being

A number of measures were used to describe the economic and material well-being of the families in the trial. These measures included:

1. Welfare dependence: At each assessment parents were asked whether either parent was currently in receipt of a social welfare benefit (including the unemployment benefit, domestic purposes benefit, sickness benefit or disability benefit). Families were defined as welfare dependent if they were reliant on a social welfare benefit.

2. Family income: Family income was assessed net of tax and included all sources of family income from welfare benefits and paid employment.

3. Family debt: At each assessment parents were asked to report the amount of money owed in debt, excluding hire purchase or mortgage costs.

4. Adequacy of income: At each assessment parents were asked to rate the adequacy of their income to meet their basic family needs on a four-point scale from "income more than adequate" to "income very inadequate". Table 6.3 reports the percentage of parents who stated their income was inadequate or very inadequate.

5. Parental workforce participation: At the 12, 24 and 36 month assessments parents were asked if they currently work in paid employment, including any part-time work. This was recorded separately for mothers and resident male partners.

6. Adequacy of accommodation: At each assessment parents were asked to rate the adequacy of their accommodation to meet their family's needs on a four-point scale from "accommodation more than adequate" to "accommodation very inadequate". Table 6.3 reports the percentage of parents who stated their accommodation was inadequate or very inadequate.

7. Economic hardship: Economic hardship factors were recorded at each assessment from parental report 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"; "postponed visits to the doctor or dentist"; "gone without meals on some days"; "bought

second-hand clothing"; "visited budget advisory service"; "been declared bankrupt"; "received a summons regarding unpaid bills"; "had to sell or pawn belongings to get money"; "needed to seek help from the food bank or social agency"; "needed to seek assistance from WINZ to pay bills"; and "moved to cheaper accommodation". A family hardship score was calculated by summing the number of hardship factors reported at each assessment period.

Family Exposure to Stress and Adversity

At each point of assessment, respondents were questioned on a 45-item questionnaire regarding their exposure to stressful and adverse life events. These events were categorised into five dimensions of stress and adversity: illness and death; economic and financial problems and crises; family or social relationship problems; victimisation; and other stressful life events. In addition, an overall score of stressful and adverse life events was calculated by summing the number of events reported by parents at each assessment.

References

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 Renormed*. Park City, UT: Family Development Resources, Inc.

Begg, C., Cho, M., Eastwood, S., Horton, R., Moher, D., Olkin, I., Pitkin, R., Rennie, D., Schulz, K. F., Simel, D., & Stroup, D. F. (1996). Improving the quality of reporting of randomized controlled trials. The CONSORT statement. *JAMA, 276*(8), 637–639.

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.

Boyd, A. (1997a). *Parents as First Teachers pilot Project Evaluation (PAFT): Report on South Auckland area*. Auckland: Auckland Uniservices Ltd.

Boyd, A. (1997b). *Parents as First Teachers Pilot Project Evaluation (PAFT): Report on South Whangarei region*. Auckland: Auckland Uniservices Ltd.

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.

Bugental, D. B., Ellerson, P. C., Lin, E. K., Rainey, B., Kokotovic, A., & O'Hara, N. (2002). A cognitive approach to child abuse prevention. *Journal of Family Psychology, 16*(3), 243–258.

Campbell, K., & Silva, P. A. (1997). *Parents as First Teachers Pilot Project Evaluation: Age Three Assessments. Final Report to the Ministry of Education on the Dunedin and Gisborne/East Coast Areas*. Dunedin: Dunedin Multidisciplinary Health and Development Research Unit.

Centre for Child and Family Policy Research. (2005). *Outcome/Impact Evaluation of Family Start: Final Report*. Auckland: Auckland Uniservices Ltd.

Chaffin, M. (2004). Is it time to rethink Healthy Start/Healthy Families? *Child Abuse & Neglect, 28*, 589–595.

Chan, A. W., Hrobjartsson, A., Haahr, M. T., Gotzsche, P. C., & Altman, D. G. (2004). Empirical evidence for selective reporting of outcomes in randomized trials: Comparison of protocols to published articles. *JAMA, 291*(20), 2457–2465.

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.

Cohen, J. (1977). *Statistical Power Analysis for the Social Sciences*. New York: Academic Press.

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. *The Future of Children, 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.

Donaldson, S. I., & Christie, C. A. (2004). *The 2004 Claremont Debate: Lipsey vs Scriven. Determining Causality in Program Evaluation & Applied Research: Should Experimental Evidence be the Gold Standard?* Retrieved February, 2005, from the World Wide Web: http://www.cgu.edu/include/SBOS_2004_Debate.pdf

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., & Sia, C. (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.

Duggan, A., McFarlane, E., Fuddy, L., Burrell, L., Higman, S. M., Windham, A., & Sia, C. (2004). Randomized trial of a statewide home visiting program: Impact in preventing child abuse and neglect. *Child Abuse & Neglect, 28*(6), 597–622.

Duggan, A., McFarlane, E. C., Windham, A. M., Rohde, C. A., Salkever, D. S., Fuddy, L., Rosenberg, L. A., Buchbinder, S. B., & Sia, C. (1999). Evaluation of Hawaii's Healthy Start Program. *The Future of Children, 9*(1), 66–90.

Duggan, A., Windham, A., McFarlane, E., Fuddy, L., Rohde, C., Buchbinder, S., & Sia, C. (2000). Hawaii's healthy start program of home visiting for at-risk families: Evaluation of family identification, family engagement, and service delivery. *Pediatrics, 105*(1), 250–259.

Durie, M. H. (1998). *Whaiora. Māori Health Development (2nd ed).* Auckland: Oxford University Press.

Ellenberg, J. H. (1996). Intention to treat analysis versus as-treated analysis. *Drug Information Journal, 30*, 535–544.

Evaluation Management Group. (2003). *Family Start Process Evaluation Final Report: A Summary and Integration of Components of the Process Evaluation Phase.* Wellington: Evaluation Management Group.

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 No. 52.* Auckland: Injury Prevention Research Centre, University of Auckland.

Farquhar, S. E. (2003). *Parents as First Teachers: A Study of the New Zealand PAFT Programme.* Wellington: Early Childhood Development.

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., & Lynskey, M. T. (1994). The childhoods of multiple problem adolescents: A 15-year longitudinal study. *Journal of Child Psychology & Psychiatry & Allied Disciplines, 35*(6), 1123–1140.

Fraser, J. A., Armstrong, K. L., Morris, J. P., & Dadds, M. R. (2000). Home visiting intervention for vulnerable families with newborns: Follow-up results of a randomized controlled trial. *Child Abuse & Neglect, 24*(11), 1399–1429.

Gomby, D. S. (1999). Understanding evaluations of home visitation programs. *The Future of Children, 9*(1), 27–43.

Gomby, D. S., Culross, P. L., & Behrman, R. E. (1999). Home visiting: Recent program evaluations – analysis and recommendations. *The Future of Children, 9*(1), 4–26.

Gray, A. (2001). *Family Support Programmes: A Literature Review. Report to the Evaluation Management Group (representatives from Ministry of Education, Ministry of Health, and Ministry of Social Policy).* Unpublished.

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.

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.

Joreskog, K. G., & Sorbom, D. (1993). *LISREL 8: Structural Equation Modelling with the SIMPLIS Command Language.* Chicago, IL: Scientific Software International.

Kempe, H. (1976). *Child Abuse and Neglect: The Family and the Community.* Cambridge, MA: Ballinger Publishing Company.

Kitzman, H., Cole, R., Yoos, H. L., & Olds, D. (1997). Challenges experienced by home visitors: A qualitative study of program implementation. *Journal of Community Psychology, 25*(1), 95–109.

Kitzman, H., Olds, D. L., Henderson, C. R. Jr., Hanks, C., Cole, R., Tatelbaum, R., McConnochie, K. M., Sidora, K., Luckey, D. W., Shaver, D., Engelhardt, K., James, D., & Barnard, K. (1997). Effect of prenatal and infancy home visitation by nurses on pregnancy outcomes, childhood injuries, and repeated childbearing. A randomized controlled trial. *JAMA, 278*(8), 644–652.

Kitzman, H., Olds, D. L., Sidora, K., Henderson, C. R. Jr., Hanks, C., Cole, R., Luckey, D. W., Bondy, J., Cole, K., & Glazner, J. (2000). Enduring effects of nurse home visitation on maternal life course: A 3-year follow-up of a randomized trial. *Journal of the American Medical Association, 283*(15), 1983–1989.

Livingstone, I. D. (1998). *Parents as First Teachers Summary Report – Evaluation of Pilot Project.* Wellington: Chartwell Consultants.

Lord, S. J., Gebski, V. J., & Keech, A. C. (2004). Multiple analyses in clinical trials: Sound science or data dredging? *Medical Journal of Australia, 181*(8), 452–454.

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. (1998). *Child Health Programme Review.* Wellington, New Zealand: Ministry of Health.

Moher, D., Schulz, K. F., & Altman, D. G. (2001). The CONSORT statement: Revised recommendations for improving the quality of reports of parallel-group randomised trials. *Lancet, 357*(9263), 1191–1194.

Olds, D. L. (1992). Home visitation for pregnant women and parents of young children. *American Journal of Diseases of Children, 146*(6), 704–708.

Olds, D. L. (2002). Prenatal and infancy home visiting by nurses: From randomized trials to community replication. *Prevention Science, 3*(3), 153–172.

Olds, D. L., Eckenrode, J., Henderson, C. R. Jr., Kitzman, H., Powers, J., Cole, R., Sidora, K., Morris, P., Pettitt, L. M., & Luckey, D. (1997). Long-term effects of home visitation on maternal life course and child abuse and neglect. Fifteen-year follow-up of a randomized trial. *JAMA, 278*(8), 637–643.

Olds, D. L., Henderson, C. R. Jr. , Cole, R., Eckenrode, J., Kitzman, H., Luckey, D., Pettitt, L., Sidora, K., Morris, P., & Powers, J. (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. *The Future of Children, 9*(1), 44–65.

Olds, D. L., Henderson, C. R. Jr., Tatelbaum, R., & Chamberlin, R. (1988). Improving the life-course development of socially disadvantaged mothers: A randomized trial of nurse home visitation. *American Journal of Public Health, 78*(11), 1436–1445.

Olds, D. L., & Kitzman, H. (1990). Can home visitation improve the health of women and children at environmental risk? *Pediatrics, 86*(1), 108–116.

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.

Olds, D. L., Kitzman, H., Cole, R., & Robinson, J. (1997). Theoretical foundations of a program of home visitation for pregnant women and parents of young children. *Journal of Community Psychology, 25*(1), 9–25.

Olds, D. L., Kitzman, H., Cole, R., Robinson, J., Sidora, K., Luckey, D. W., Henderson, C. R. Jr, Hanks, C., Bondy, J., & Holmberg, J. (2004). Effects of nurse home-visiting on maternal life course and child development: Age 6 follow-up results of a randomized trial. *Pediatrics, 114*(6), 1550–1559.

Olds, D. L., Robinson, J., O'Brien, R., Luckey, D. W., Pettitt, L. M., Henderson, C. R. Jr., Ng, R. K., Sheff, K. L., Korfmacher, J., Hiatt, S., & Talmi, A. (2002). Home visiting by paraprofessionals and by nurses: A randomized, controlled trial. *Pediatrics, 110*(3), 486–496.

Olds, D. L., Robinson, J., Pettitt, L., Luckey, D. W., Holmberg, J., Ng, R. K., Isacks, K., Sheff, K., & Henderson, C. R. Jr. (2004). Effects of home visits by paraprofessionals and by nurses: Age 4 follow-up results of a randomized trial. *Pediatrics, 114*(6), 1560–1568.

Public Health Group. (1996). *Cannabis: The Public Health Group's Advice to the Minister of Health 1995–1996.* Wellington: Public Health Group.

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.

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 & Family Psychology Review, 2*(2), 71–90.

Schulz, K. F., & Grimes, D. A. (2005). Multiplicity in randomised trials I: Endpoints and treatments. *Lancet, 365*, 1591–1595.

St Pierre, R. G., & Layzer, J. I. (1999). Using home visits for multiple purposes: The Comprehensive Child Development Program. *The Future of Children, 9*(1), 15–34.

StataCorp. (2003). Stata Statistical Software: Release 8.0. 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 & Neglect, 22*(4), 249–270.

Swerissen, H. (1999). Health promotion evaluation. Guest editorial. *Australian Journal of Primary Health – Interchange, 5*(4), 6–8.

Wagner, M. M., Cameto, R., & Gerlach-Downie, S. (1996). *Intervention in Support of Adolescent Parents and Their Children: A Final Report on the Teen Parents as Teachers Demonstration.* Menlo Park, CA: SRI International.

Wagner, M. M., & Clayton, S. L. (1999). The Parents as Teachers program: Results from two demonstrations. *The Future of Children, 9*(1), 91–115.

Wagner, M. M., Clayton, S. L., Gerlach-Downie, S., & McElroy, M. (1997). *An Evaluation of the Northern California Parents as Teachers Demonstration.* Menlo Park, CA: SRI International.

Webster-Stratton, C. (1998). Parent training with low-income families: Promoting parental engagement through a collaborative approach. In J. R. Lutzker (Ed), *Handbook of child abuse research and treatment* (pp 183–210). New York: Plenum Press.

Wechsler, D. (2002). *WPPSI-III Administration and Scoring Manual.* San Antonio: The Psychological Corporation.

World Health Organization. (1993). *Composite International Diagnostic Interview (CIDI).* Geneva, Switzerland: World Health Organization.