The Development of a Screening Tool to Evaluate Infants who are HIV Positive

N C Hilburn

A thesis submitted to the Faculty of Health Sciences of the University of the Witwatersrand, for the degree of Doctor of Philosophy.

Johannesburg **2010**

ABSTRACT

HIV/AIDS continues to be one of the greatest health challenges which South Africa faces. The epidemic in children is closely linked to that in women, the prevalence of which continues to grow according to antenatal statistics from the South African Department of Health (DOH). HIV is known to invade the central nervous system at the time of infection, and causes widespead damage. In children, this leads to a well-researched condition known as HIV encephalopathy, which affects all areas of neurodevelopment. The effects of timely initiation of antiretroviral therapy on alleviating the impact of encephalopathy have been well described.

Neurodevelopmental delay is a stage four disease indicator according to the World Health Organisation (WHO), and therefore is a criterion for initiation of Highly Active Antiretroviral Therapy (HAART). HAART is often only administered according to the virologic and immunologic status of a child, as standardised neurodevelopmental assessment tools are not widely available in South African clinics. When HAART initation is dependent on immunologic status, it is often too late to prevent encephalopahy. To date, the only means of prevention of this condition is early initation of HAART, which has not been widely available in South Africa. Stringent guidelines for the commencement of this therapy according to the WHO, and the South African Department of Health (DOH) have had to be followed, leading to late initiation of HAART, and widespread central nervous system encephalopathy. Studies which have been carried out in South African clinics have demonstrated the high prevalence of this condition. Once there is evidence of encephalopathy, children should be referred for assessments in all facets of development, and where necessary, for rehabilitation. A standardised developmental screening tool which is suitable for use in a developing country is therefore necessary in order to screen for neurodevelopmental delays to allow for further assessment and referral to rehabilitation services, as well as providing an additional assessment criterion for initiation of HAART.

Paediatric HIV clinics in developing countries are understaffed, and children may be seen by junior staff or screened by nurses due to the high numbers of clinic attendees. This often results in neurodevelopment being inadequately assessed and children are therefore not referred for intervention services. A standardised screening tool, which

could be administered by clinic staff in order to ensure correct and timely referral of children for further assessment and intervention is therefore necessary. This is of importance both locally and internationally where a screening tool, which has been developed specifically for this purpose, does not exist.

The aim of this study was therefore to evaluate the agreement between the Bayley-III Screening Test and the Bayley Scales of Infant Development (3rd version) in a population of HIV positive infants in order to evaluate its appropriateness for use in South Africa. The Bayley Scales of Infant Development have long been considered the 'gold standard' in infant developmental assessment, which is why this tool was chosen to evaluate the Bayley-III Screening Test against. The developmental scores in each facet (cognitive, motor or language) were evaluated to determine which should be included in an assessment tool for this population. Further objectives for the study were to adapt the screening tool to the needs of the population, or to develop a new screening tool should the Bayley-III Screening Test not prove suitable for use in this population.

In order to meet the aims and objectives, a cross-sectional study was conducted where 112 HIV positive infants between the ages of six and eighteen months were assessed using the Bayley-III Screening Test and the Bayley Scales of Infant Development (3rd version) (BSID III). The infants were stratified into four age groups namely 6-8 months, 9-12 months, 13-16 months, and 17-18 months. Children were recruited from Harriet Shezi Children's Clinic at Chris Hani Baragwanath Hospital in Soweto.

The agreement between the Bayley-III Screening Test and the Bayley Scales of Infant Development (3rd version) was analysed using Kappa, for the overall group, and for each age group. Overall agreement between the tools was as follows: K=0.58 for the Cognitive facet, K=0.82 for the Expressive Communication facet, K=0.76 for the Receptive Communication facet, K=0.44 for the Fine Motor facet and K=0.57 for the Gross Motor facet. These values indicate that the Bayley-III Screening Test is therefore not acceptable for clinical use, as excellent agreement ($k \geq 0.75$) in all facets would be necessary for this purpose.

A new screening tool therefore had to be developed. The infant's developmental scores from the BSID III were analysed to determine which facets of development were most

severely affected, and therefore which facets should be included in a new screening tool. Gross motor function was demonstrated to be the area which was most severely affected, followed by cognitive function. A gross motor screening tool would therefore be suitable for use in this population, as no equipment would be necessary. Gross motor development is the most universally similar aspect of development, which is not completely dependent on cultural or socioeconomic factors which often have an influence on language and cognitive development.

Item selection from the BSID III was undertaken to determine which items should be included in a brief screening tool. In each of the four age groups, item selection occurred as follows: Two items which discriminated the At-Risk, from Emerging and Competent groups (less than 20% in the At-Risk group, and 100% in the other groups) were selected. Two items, which discriminated between children in the 'Emerging' and 'Competent' categories on the BSID III were selected (0-5% of children who were At-Risk obtained credit, 30-50% of the Emerging group obtained credit, and 100% of the Competent group obtained credit). Lastly, two items were selected which discriminated the Competent group from the other two groups (100% or as high as possible in the Competent group, and 0% in the other groups).

The new gross motor screening tool was assembled using the selected items, scoring was allocated, and it was tested against the scores obtained on the Gross Motor facet of the BSID III for the initial 112 infants. Agreement between the tools was analysed using Kappa, and refinements were made according to the discrepancies. This was done three times, until the Kappa value revealed excellent agreement between the tools ($k = 0.87$). A panel of experts was then invited to examine the new gross motor screening tool, and to comment on it, and further adjustments were made accordingly.

Preliminary concurrent validity testing of the new gross motor screening tool was then carried out against the Gross Motor facet of the BSID III on 60 children, who were recruited from the Harriet Shezi Children's Clinic at Chris Hani Baragwanath Hospital in Soweto. Statistical analysis revealed that the agreement between the BSID III and the new screening tool was excellent ($k=0.85$). The diagnostic properties of the new gross motor screening tool were as follows: sensitivity 97.4%, specificity 85.7%, positive predictive value 92.7%, and negative predictive value 94.7%. These values indicate that

the statistical properties of the tool are excellent, and the tool will not be predisposed to underreferrals or over-referrals. Preliminary reliability testing was carried out on 15 children for test-retest/intrarater reliability and 15 children for interrater reliability. Interrater, test-retest and intrarater reliability were excellent (r=1, r=0.98, r=0.98 respectively). Further testing of reliablity and validity should be undertaken in order to establish these properties, and standardisation should also be carried out on healthy children. Given the need for an assessment tool of this nature in South Africa and other developing countries, and the statistical properties thus far, the tool may be used clinically for the purposes for which it was developed.

DECLARATION

I, Nicole Clare Hilburn, declare that this is my own unaided work except for the help given by the persons listed under the acknowledgements.

Signed this day in Johannesburg

Signature

Date

ACKNOWLEDGEMENTS

I wish to acknowledge the following people for the roles they have played in the completion of this study:

Dr JL Potterton and Prof AV Stewart for supervision, mentorship and friendship.

Prof M Westaway for help and guidance

Mrs L Goldberg for assistance with data collection.

Mrs L Rawstorne for assistance with data collection

Ms N Whitehead for assistance with data collection

Prof P Becker for statistical help (Medical Research Council of South Africa).

Dr T Meyers for permission to conduct this study at the Harriet Shezi Clinic.

Dr H Moultrie for support and advice

All staff at Harriet Shezi Clinic for their help.

Caregivers and children for their participation

Carnegie Corporation for funding

My family for all their support

PUBLICATIONS AND PRESENTATIONS IN THE FIELD OF PAEDIATRIC HIV

Publications

Baillieu N (Hilburn), Potterton J 2008 The Extent of Delay of Language, Motor and Cognitive Development in HIV Positive Infants. Journal of Neurologic Physical Therapy September 32 (3) 118 - 121

Potterton J, Stewart A, Cooper P, Baillieu N, Gajdosik C 2009 Neurodevelopmental Delay in Children Infected with HIV in Soweto, South Africa. Vulnerable Children and Youth Studies March 4 (1) 48 - 57

Presentations

SASP Congress 2005 (28 May 2005) Baillieu N, Potterton J and Becker P "The Extent of Delay of Language, Motor and Cognitive Devlopment in HIV Positive Infants"

Wits Paediatric Clinics Workshop (2006) Baillieu N "Neurodevelopment in Paediatric HIV"

University of the Witwatersrand Faculty of Health Sciences Research day: August 2006 Baillieu N, Potterton J and Becker P "The Extent of Delay of Language, Motor and Cognitive Devlopment in HIV Positive Infants"

Platform presentation at WCPT Vancouver, 2007 Baillieu N, Potterton J and Becker P "The Extent of Delay of Language, Motor and Cognitive Devlopment in HIV Positive Infants"

Platform presentation at University of Pretoria Physiotherapy Congress, October 2008 Baillieu N, Potterton J and Becker P "The Extent of Delay of Language, Motor and Cognitive Devlopment in HIV Positive Infants"

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

LIST OF ABBREVIATIONS

AAP – American Association of Pediatrics
AIDS – acquired immunodeficiency syndrome
AIMS – Alberta Infant Motor Scale
ART – antiretroviral therapy
ASQ – Ages and Stages Questionnaire
BINS – Bayley Infant Neurodevelopmental Screener
BMI – body mass index
BSID III – Bayley Scales of Infant and Toddler Development III
CDC – Centre for Disease Control
CNS – central nervous system
CSF – cerebrospinal fluid
CT – computed tomography
DDST – Denver Developmental Screening Test
DOH – Department of Health
GH – growth hormone
HAART – highly active antiretroviral therapy
HESSI – Household Economic and Social Status Index
HINT – Harris Infant Neuromotor Test
HIV – human immunodeficiency virus
ICC – intraclass correlation coefficient
IL-6 – interleuken-6
IQ – intelligence quotient
LMIC's – low and middle income countries
MAI – Movement Assessment of Infants
MDI – Mental developmental index
MPI-1a – macrophage inflammatory protein-1a
MRI – magnetic resonance imaging
NGT – Nominal Group Technique
NNRTI - non-nucleoside reverse transcriptase inhibitor
NPV – negative predective value
NRTI - nucleoside reverse-transcriptase inhibitor

NSP – National Strategic Plan

PDGMS – Peabody Developmental Gross Motor Scale

PDI – Psychomotor Developmental Index

PI – protease inhibitor

PMTCT – prevention of mother to child transmission

PPV – positive predictive value

RNA – ribonucleic acid

SD – standard deviation

SES – socieconomic status

SIV – Simian Immunodeficiency Virus

TB - Tuberculosis

TLC – total lymphocyte count

TNF – tumour necrosis factor

US - United States

WHO – World Health Organisation

DEFINITION OF TERMS

CD4 count

The CD4 count is used to quantify the immune status of a patient. The CD4 count is the absolute number of CD4+ T lymphocytes (McFarland, 2005). Children who attend Harriet Shezi clinic have their CD4 count taken at baseline, and then at three monthly intervals. CD4 counts are used to determine eligibility for highly active antiretroviral therapy (HAART) and to monitor treatment response. Interpretation of CD4 numbers in children requires the recognition that the normal number of cells declines with age over the first six years of life. Normal values for age have therefore been calculated.

CD4 percentage

CD4 percentage is the fraction of total lymphocytes that express CD3 and CD4 cell surface markers. CD4+ T-cell percentage fluctuates less than absolute counts, as variations in total white cell count do not affect the percentage as much (McFarland, 2005). The CD4 percentage threshold for the commencement of HAART is as follows: less than 25% for infants less than 12 months and less than 20% for children aged 12 to 36 months. Children older than three years are eligible for HAART when the CD4 count is less than 15% (Department of Health, 2008).

International paediatric guidelines recommend monitoring both percentage and absolute CD4 cell count, and determining disease staging based on the lower of the two values. A sustained decrease of 50% in the absolute count or percentage is accepted as significant evidence of disease progression (McFarland, 2005).

HIV-1

There are two major types of HIV: HIV-1 and HIV-2. Although similar to HIV-1, HIV-2 has a different sequence of nucleotides in its genome. Studies suggest that both HIV-1 and HIV-2 evolved from Simian Immunodeficiency Virus (SIV), but from different simian species. HIV-1 diverged from chimpanzee SIV and HIV-2 from sooty mangabey SIV. HIV-2 is less transmissible than HIV-1 both vertically and between partners, and progression to end-stage AIDS is slower. HIV-2 is more commonly found in countries in West Africa (Pratt, 2003).

Viral load

Viral load is an indication of the amount of viral RNA in the plasma, and is a direct representation of the amount of circulating virus. The measure is used to monitor the child's response to HAART, and no change or an increase in viral load, despite having commenced highly active antiretroviral therapy (HAART) is an indication of poor response or adherence. Harriet Shezi Clinic uses the nucleic acid sequence based amplification (NASBA), with the optimal result on this test being a viral load of less than 25.

Antiretroviral regimen

Children who are managed at the Harriet Shezi Clinic need to meet the South African Department of Health criteria for the commencement of antiretroviral therapy. The clinic aims to start every child who meets the Department of Health Criteria on HAART using the following regimen: Stavudine (d4T) (NRTI), Lamivudine (3TC) (NRTI), Lopinavir/ritonavir (Kaletra®) (PI). HAART exposure is defined as the concomitant use of at least three drugs from at least two classes of HIV drugs (NRTI, NNRTI and PI) (WHO, 2008), and therefore the combination of drugs used at the clinic qualifies as HAART.

Composite Score (For the Bayley Scales of Infant Development III)

Composite scores are derived from sums of subtest scaled scores. They are scaled to a metric with a mean of 100 and a standard deviation (SD) of 15, and range from 40-160 (Bayley, 2006a). The interpretation of the scores is shown in Table 1 below:

Table 1.1 Interpretation of BSID III Composite Scores

Composite	Classification
130 and above	Very Superior
120 – 129	Superior
110 – 119	High Average
90 – 109	Average
80 – 89	Low Average
70 – 79	Borderline
69 and below	Extremely Low

(Bayley, 2006b)

Scaled Score (For the Bayley Scales of Infant Development III)

Scaled scores are derived from the subtest total raw score, and range from one to 19, with a mean of 10, and a SD of three (Bayley, 2006a). The conversion of the scaled score to the categories on the Bayley-III Screening Test for comparison is shown in Table 1.2 below:

Table 1.2 Conversion from BSID III to Bayley-III Screening Test Category

Scaled Score on BSID III	Category on the Bayley-III Screening Test
1 – 4	At-Risk
5 – 7	Emerging
8 +	Competent

(Bayley, 2006c)

Kappa

In this study, Kappa has been used to evaluate agreement between two assessment tools. It is an index that compares the agreement against that which might be expected by chance (Portney and Watkins, 2000). The interpretation that will be used in this study is shown in the table below:

Table 1.3 Interpretation of Kappa scores

Kappa Score	Interpretation
≤ 0.4	Poor Agreement
> 0.4 - < 0.75	Moderate agreement
≥ 0.75	Excellent agreement

(Landis and Koch, 1977)

Z-score

$$\text{Z-score (or SD-score)} = \frac{\text{observed value - median value of the reference population}}{\text{standard deviation value of reference population}}$$

(WHO, 1997)

The z-score system expresses the anthropometric value as a number of standard deviations or z-scores below or above the reference mean or median value. The advantages of using z-scores are as follows: z-scores have the same statistical relation to the distribution of the reference around the mean at all ages, which makes results comparable across ages groups and indicators, and z-scores are sex-independent. These characteristics allow further computation of means, standard deviations, and standard error to classify a population's growth status (WHO, 1997).

Chapter 1

INTRODUCTION

It is estimated that 33.4 million people worldwide were living with the Human Immunodeficiency Virus (HIV) in 2008, over half of whom were women (UNAIDS, 2009). In 2008, an estimated 390 000 children in sub-Saharan Africa were infected (UNAIDS, 2009). The HIV epidemic amongst children is closely linked to that of women, since the majority of paediatric infections are the result of vertical transmission from mother to child (UNAIDS, 2006; Thorne and Newell, 2000).

More than twenty years have elapsed since the original description of the effects of HIV on the central nervous system (CNS) (Epstein et al, 1985), and the nervous system has been described as one of the most frequent and serious targets of HIV (Raskino et al, 1999; Simpson and Berger, 1996). Involvement of the nervous system in patients infected with HIV is a well-described condition known as HIV encephalopathy, and may come about as a result of direct HIV action, opportunistic infection caused by the immunodeficiency, or both (Fragoso et al, 1999).

Significant delays in mental and motor development of HIV infected infants have been found over the first two years of life (Potterton et al, 2009a; Potterton et al, 2009b; Baillieu and Potterton, 2008; Van Rie et al, 2008; McGrath et al, 2006a; Drotar et al, 1997; Msellati et al, 1993). In a study conducted in South Africa on children under three and a half years of age, it was found that 72% had severe motor delay, and 52% had severe cognitive delay (Potterton et al, 2009b). Such impairments are likely to become more severe as the children get older (Blanchette et al, 2001; Belman, 1992).

Neurological and developmental signs are often markers of HIV infection in infants, which may precede other signs of disease progression (Bisiacchi et al, 2000; Belman, 1992). HIV infection of the developing CNS of infants and children is characterised by either a progressive or static loss of previously acquired developmental milestones with cognitive, behavioural and motor manifestations (Udgirkar et al, 2003; Belman et al, 1988; Epstein et al, 1986; Sharer et al, 1986; Belman et al, 1985).

Development of encephalopathy is one of the most severe complications of HIV but its frequency at different ages has long been disputed (Tardieu et al, 1992). HIV encephalopathy may occur very early in the course of HIV infection and 88.1% of children who develop it do so within the first two years of life (Newell et al, 1998). This is more common in children infected with HIV who present in early infancy and show rapid progression, as they may have been infected in-utero during the last weeks of pregnancy, which is the period of fastest brain growth (Newell et al, 1998).

The most frequent manifestations of HIV-associated progressive encephalopathy are: cognitive impairment, developmental delays, corticospinal tract lesions, acquired microcephaly, movement disorders, and ataxia. Progressive motor dysfunction is common and often results in a loss of milestones (Belman, 1992). Risk factors for paediatric CNS involvement include the timing of infection, advanced maternal disease at delivery, rapid progression with early advanced immune suppression, high plasma viral load in infancy, and factors such as home environment and socioeconomic status (Robertson et al, 2008).

HIV encephalopathy is one of the criteria for initiation of Highly Active Antiretroviral Therapy (HAART) (ANNECCA, 2005). Encephalopathy is one of the most common CNS disorders amongst HIV infected children worldwide, and while this is a HAART eligibility criterion, neurodevelopmental assessment is rarely performed in developing countries due to a shortage of skilled staff and a lack of validated assessment tools for these settings (Van Rie et al, 2007). This results in missed opportunities for timely initiation of HAART as well as for the prevention and management of HIV-associated neurodevelopmental impairment. Validation and standardisation of neurodevelopmental assessment tools for developing-world contexts are vital in order to identify early warning signs of HIV-associated CNS disorders and to allow the integration of neurodevelopmental assessment into paediatric HIV care and treatment programmes worldwide (Van Rie et al, 2007).

The Bayley Scales of Infant Development II (Bayley, 1996), have been extensively used in HIV/AIDS research worldwide, and therefore have been used in South Africa in order to be able to compare results with worldwide studies (Potterton et al, 2009a; Potterton et al, 2009b; Baillieu and Potterton, 2008; McGrath et al, 2006a; Aina and

Morakinyo, 2005; Chase et al, 2000; Drotar et al, 1997; Gay et al, 1995). The Bayley Scales have been found to be sensitive to developmental changes in the first two years of life of infants who are medically fragile (Niccols and Latchman, 2002), and these findings support the clinical validity of the scale, which can be applied to those with HIV. Like its predecessors (Bayley Scales of Infant Development I and II), the Bayley Scales of Infant Development III (BSID III) is an individually administered instrument that assesses developmental functioning of infants and young children 1-42 months of age. The primary purpose of the BSID III is to identify children with developmental delay and to provide information for intervention planning (Bayley, 2006a). The BSID III is a revision of the BSID II, in order to improve the quality and enhance the utility of the instrument. The BSID III was standardised on a national sample of 1,700 children in the United States of America who were stratified by age, sex, race/ethnicity, parent education level and region (Bayley, 2006b). The BSID III is reliable and valid, and is predictive of later developmental outcome (Bayley, 2006b), however, this is not an appropriate tool to use in busy HIV clinics in South Africa, as the administrator needs to be trained and the assessment is lengthy.

The BSID I (Bayley, 1993) was normed on over 700 South African children from both urban and rural areas, and was found to be suitable for use on South African infants (Richter and Griesel, 1992; Richter et al, 1988). The South African infants scored above the test norms up to 10 months of age on both the Mental and Motor Scales, and from 10 months up to the middle of the second year the groups obtained very similar scores. The trend across the whole age range in the South African sample was for urban children to score higher than rural children on both the Mental and Motor scales, although this was not statistically significant across the age range (Richter and Griesel, 1988). Based on these findings, the authors recommended that the BSID be used without adaptation on South African children. A study was also done in Uganda, examining normative data for the BSID II on infants between eight weeks and 30 months (Aina and Morakinyo, 2005). The infants scored above the normal values in both the Mental and Motor scale of the BSID II, and furthermore, their scores in the motor scale were significantly higher, especially in the lower age groups of eight, 16 and 32 weeks when compared with data from the Western World (Aina and Morakinyo, 2005). The BSID I, II and III are well correlated (Bayley, 2006b; Bayley,

1996), and therefore the results from the above can be applied to the BSID III and its applicability for use on South African infants.

The Bayley-III Screening Test (Bayley, 2006c) was developed using items from the BSID III. It is designed to assess briefly the cognitive, language and motor functioning of infants and young children between 1-42 months of age. The primary purpose is to determine whether a child shows competence in age-appropriate tasks, shows evidence of emerging age-appropriate skills, or shows evidence of being at-risk for developmental delay. It was standardised on a sample of 1,675 children and included representative proportions of infants according to selected demographic variables. It is also reliable and valid (Bayley, 2006c).

Screening is the process by which conditions are identified by tests that can be applied rapidly on a large scale (Beaglehole et al, 1993). The screening test should be cost-effective, reliable and valid, and should be sensitive and specific (Beaglehole et al, 1993). Even though there is no gold standard (an ideal test covering all areas of development, being equally applicable to all ages, with construct validity, and an accuracy approaching values of 100% for sensitivity and specificity), many screening tests are well standardised, with acceptable levels of sensitivity and specificity in the context of documented developmental delays. The test should be acceptable to the population and should be non-invasive and culturally sensitive (Rydz et al, 2006).

There is therefore a need for an assessment tool which can be administered by doctors or nurses at the HIV clinics, which is less lengthy, yet is able to identify those children who need referral to services for encephalopathy and developmental delay. At present, South African children are not being screened, and therefore are not adequately referred for services (Potterton et al, 2009a). The development of a screening tool which is suitable for use in this setting has been incorporated into the National Strategic Plan (NSP) for HIV for 2007-2011, and the South African government has made a commitment to ensuring that all HIV positive children under three years of age undergo developmental screening (NSP, 2007 – 2011).

1.2 Problem statement

HIV is known to affect the CNS, and consequently neurodevelopment in infants and children. Paediatric HIV clinics in developing countries are understaffed, and doctors often see over 100 children a morning; children may be seen by junior staff or screened by nurses due to the high numbers of clinic attendees. This often results in inadequate assessment of neurodevelopment and children are therefore not referred for intervention services. A standardised screening tool is needed, which is able to be administered by clinic staff in order to ensure correct and timely referral of children for further assessment and intervention. Children who are clinically eligible for HAART due to the presence of encephalopathy are not being detected and therefore a screening tool would assist with this. This is of importance both locally and internationally where a screening tool, which has been developed specifically for this purpose, does not exist.

1.3 Aim of the study

The aim of this study was to evaluate the agreement between the Bayley-III Screening Test and the Bayley Scales of Infant Development (3rd version) in a population of HIV positive infants in order to determine its appropriateness for use in South Africa. The separate areas of the tool were evaluated to determine which facet (cognitive, motor or language) is most valuable, and could be used alone. The tool was to be adapted accordingly to make it feasible for use in terms of time, equipment and technical expertise. Should the tool not be suitable for use, a more appropriate screening tool was to be developed.

1.4 Objectives

- To determine how the Bayley-III Screening Test needed to be adapted for use in developing countries such as South Africa by shortening the assessment time and decreasing the need for equipment;
- To determine which areas of the screening tool were most useful for use on this population.
- To test and compare results from the BSID III and the Bayley-III Screening Test on a sample of HIV positive infants;
- To determine the socioeconomic status of the sample, and whether there was a relationship between health status and developmental outcomes;

- To evaluate the suitability of the Bayley-III Screening Test for use in this population;
- Should the Bayley-III Screening Test not prove to be suitable, a new tool was to be developed, and preliminary testing of the new tool against the BSID III was to be undertaken.

1.5 Organisation of the Thesis

Figure 1.1 below outlines the way in which the study was conducted.

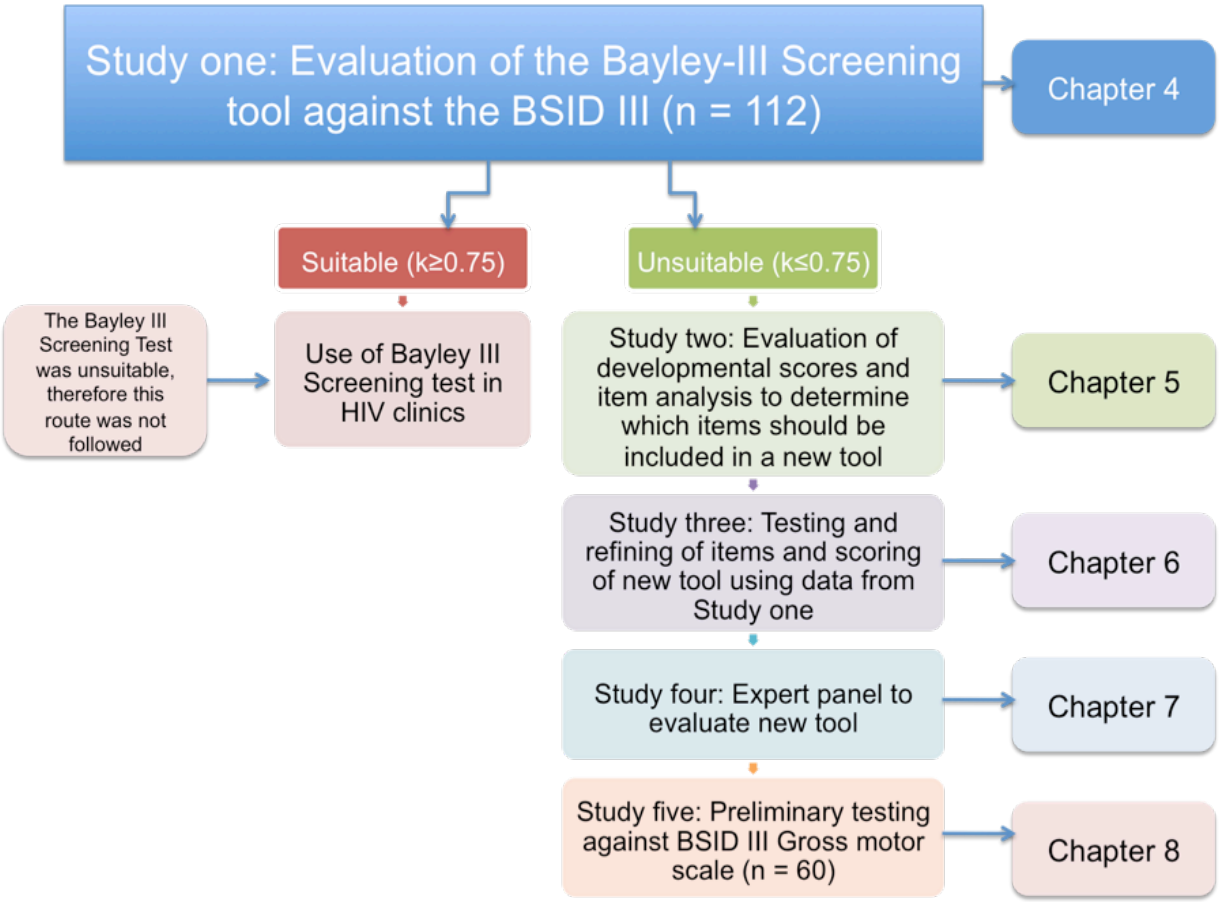


Figure 1.1 Diagram depicting the organisation of the thesis

As can be seen above, the sections of the thesis are divided into a number of different studies, which are each presented in a separate chapter. **Study one** consisted of the testing of The Bayley-III Screening Test against the Bayley Scales of Infant and Toddler development III (BSID III) in order to determine its suitability for use as a screening tool in South Africa. It was found to be unsuitable (K<0.75), and therefore

the final objective of the study was undertaken, and a new screening tool had to be developed. This is described in chapter four.

Study two was then undertaken: this involved analysis of developmental scores in order to determine which facet of development was most severely affected, and which areas of development should be included in a new screening tool. Item analysis of the Bayley Scales of Infant and Toddler Development III was undertaken in order to determine which items should be included in a new tool. Study two is discussed in chapter five.

Study three: Once the items for the new tool had been selected, and scoring had been allocated, testing of the first version against the BSID III was undertaken. Following analysis of the results, the relevant sections were refined, and the second version was tested against the BSID III. Again, the results were analysed and the relevant sections were adjusted. Study three is discussed in chapter six.

Study four consisted of a meeting with a panel of experts who were asked to comment on the new screening tool. Suggestions were put forward by the panel members and discussed within this meeting, and relevant changes were made to the tool following this discussion. This is described in chapter seven.

Study five consisted of the testing of the final version of the screening tool against the BSID III on 60 infants between the ages of 6 and 18 months. This is described in chapter eight.

All results are discussed in detail in chapter nine, and conclusions are drawn in chapter ten.

Chapter 2

LITERATURE REVIEW

2.1 Introduction

In this chapter an overview of paediatric HIV will be given with a more detailed discussion on the neurological complications of this disease. The impact of paediatric HIV on growth will be presented, and the socioeconomic implications on development and growth will be discussed. Developmental screening and surveillance in paediatrics will be presented, and necessary properties of standardised assessment tools are discussed. Six commonly used standardised paediatric developmental screening tools are compared to illustrate this.

Articles were sourced for this review using Pubmed, CINAHL, PSYCHInfo, and Cochrane Collaboration searches. A hand search was also conducted in the Health Sciences Library of the University of the Witwatersrand. Key words used in searches included; HIV, children, encephalopathy, developmental delay, neurodevelopment, socioeconomics.

2.2 Epidemiology of the HIV/AIDS Pandemic

Globally, an estimated 33.4 million (31.1-35.8 million) people worldwide were living with HIV in 2008, over half of whom were women (UNAIDS, 2009). Whilst the prevalence of people living with HIV has stabilised since 2000, the overall number of people living with HIV has increased as new infections occur each year and outnumber deaths due to HIV/AIDS.

Sub-Saharan Africa remains the region most heavily affected by HIV. In 2008, sub-Saharan Africa accounted for 67% of HIV infections worldwide, 68% of new HIV infections among adults, and 91% of new HIV infections among children. The region also accounted for 72% of the world's AIDS-related deaths in 2008 (UNAIDS, 2009). Mother-to-child transmission continues to account for a substantial portion of new HIV infections in many African countries. In 2008, an estimated 390 000 [210 000–570 000] children were infected in sub-Saharan Africa (UNAIDS, 2009).

In 2008, an estimated 1.9 million [1.6 million–2.2 million] people living in sub-Saharan Africa became newly infected with HIV. This brought the total number of people living with HIV to 22.4 million [20.8 million– 24.1 million]. The rate of new HIV infections in sub-Saharan Africa has slowly declined, with the number of new infections in 2008 approximately 25% lower than at the epidemic's peak in the region in 1995, but the number of people living with HIV in sub-Saharan Africa slightly increased in 2008, which is partly due to improved access to HIV treatment, with the number of new infections now exceeding the number of deaths. Adult (15–49 years) HIV prevalence declined from 5.8% [5.5–6.0%] in 2001 to 5.2% [4.9–5.4%] in 2008. In 2008, an estimated 1.4 million [1.1 million–1.7 million] AIDS-related deaths occurred in sub-Saharan Africa, which represents an 18% decline in annual HIV-related mortality in the region since 2004. South Africa is home to the world's largest population of people living with HIV (5.7 million) (UNAIDS, 2009).

Globally, new HIV infections in children peaked in 2000-2002, and are now decreasing due to the stabilisation of HIV prevalence among women overall, as well as increasing coverage of programmes for preventing mother-to-child transmission of HIV (UNAIDS, 2009). This is not the case in most countries in Sub-Saharan Africa, where rates are not decreasing as quickly.

In South Africa, Botswana and Zimbabwe, HIV is the underlying cause of more than one third of all deaths among children under the age of five (Mason, 2006). Without antiretroviral treatment, the progression of HIV infection in children is particularly aggressive, and many children die at a young age (Brahmbhatt et al, 2006; Newell et al, 2004; Taha et al, 2000). HIV data from antenatal clinics in South Africa suggest that the country's epidemic may be stabilising (Department of Health, South Africa, 2007), but there is no evidence of major changes in HIV-related behaviour. Meanwhile, the estimated number of maternal, paternal, and double orphans due to HIV/AIDS in South Africa in 2008 stood at 1.7 million (UNAIDS, 2009).

2.3 Mother-to-Child Transmission

The HIV epidemic amongst children is closely linked to that amongst women, since the majority of paediatric infections are the result of vertical transmission from mother to child (UNAIDS, 2009). Over 95% of HIV positive children in sub-Saharan Africa are

infected as a result of vertical infection (Dabis and Ekpini, 2002). Vertically transmitted HIV occurs in an immature and developing organism by transplacental or intrapartum transmission from mother to child (Rausch and Stover, 2001; Belman, 1992). HIV infection rates among pregnant women attending South African antenatal clinics have steadily increased from less than one percent in 1990 to 29.1% in 2006 (UNAIDS, 2008). Encouragingly, this represents a slight drop from its peak of 30.2% in 2005, but the pandemic is still not under control (UNAIDS, 2008). The cumulative rates of HIV transmission from mother-to-child are 25-40% with the breakdown as follows: under 10% are infected during pregnancy, 10% to 20% at birth, and 10% to 20% during breastfeeding (Luzuriaga and Sullivan, 2002). Maternal-, obstetric- and infant-related factors have been associated with the risk of vertical transmission, with the most important maternal factor being viral load (European Collaborative Study, 1996). Overall, in the absence of interventions to prevent transmission, two-thirds of mother-to-child transmission takes place in the peri-partum period, which can either occur in utero (around a third of cases), mostly late in the third trimester or intra-partum (around two-thirds of cases) (Chouquet et al, 1999).

Obstetric risk factors associated with an increased risk of mother-to-child transmission include vaginal delivery, ascending infection after prolonged rupture of membranes, chorioamnionitis, materno-foetal micro-transfusions during uterine contractions and absorption of the virus through the infant's digestive tract (Newell, 1998), as well as invasive obstetric procedures and co-infection with another sexually transmitted infection (European Collaborative Study 1999; The European Mode of Delivery Collaboration, 1999; Van Dyke et al, 1999; Mandelbrot et al, 1998; Landesman et al, 1996; Tovo et al, 1996; Ryder et al, 1989). The onset of perinatally acquired HIV infection is early with most infected infants showing signs of disease by six months of age. Children who are infected through vertical transmission are more likely to suffer from neurodevelopmental problems compared with those infected later on, as the most significant period of brain development is affected (Willen, 2006).

In a randomised trial of breastfeeding versus formula feeding which was undertaken in Kenya, 44% of infections in the breast-feeding group were attributable to breast-milk transmission (Coutsoudis et al, 2004; Nduati et al, 2000). The first few weeks of life are a particularly high-risk period, since colostrum and early milk have a higher viral

load than milk that is produced later (Nduati et al, 2001). However, transmission can take place at any time during breastfeeding; a meta-analysis of nine trials showed a cumulative risk that remains constant from one month to 18 months of age (Coutsoudis et al, 2004). Acquisition of infection through breastfeeding is often attributable to mixed feeding, where children who experience mixed feeding as opposed to exclusive breastfeeding during the first six months of life are at a higher risk of vertical transmission (Coovadia and Bland 2007; Coovadia and Coutsoudis 2007).

2.4 Entry of HIV into the Central Nervous System

Over twenty years have elapsed since the original description of the assault of HIV on the central nervous system (CNS) (Epstein et al, 1985) and the nervous system is among the most frequent and serious targets of HIV (Simpson and Berger, 1996). The incidence of CNS involvement is not known, although it is thought to occur in most HIV infected children (Fragoso et al, 1999). Early encephalopathy may be related to the occurrence of pathological events during late foetal life (Rausch and Stover, 2001; Tardieu et al, 2000). Studies have shown that significantly more children under three years of age show evidence of CNS disease, compared with children who have survived to over six years of age (Blanche et al, 1990). Involvement of the nervous system in HIV infected patients may come about as a result of direct HIV action, opportunistic infection caused by the immunodeficiency, or both (Zink et al, 2006; Fragoso et al, 1999). In children, CNS effects usually result directly from HIV-1 infection and not from opportunistic infections (Belman, 1992). The pathogenesis of brain injury in HIV-1 infected patients remains incompletely understood and is confounded by factors such as timing of infection, variable treatments, and lack of access to brain tissue for evaluation (Fuller et al, 2004).

The CNS is a viral reservoir for HIV-1 (Kolson, 2002). It has been well established that HIV-1 invades the CNS early in infection, and HIV has been found in the cerebrospinal fluid (CSF) at or near the time of seroconversion (Spector et al, 1993; Davis et al, 1992; Epstein et al, 1987; Goudsmit et al, 1986; Ho et al, 1985; Resnick et al, 1985), as well as in the aborted foetuses of HIV infected mothers as early as fifteen weeks of gestation (Lyman et al, 1990). Once inside the CNS, viral replication occurs (Trujillo et al, 2007; Lipton and Gendelman, 1995; Davis et al, 1992). Although there is evidence

of primary HIV CNS injury, the virus does not directly infect neurons, and instead causes neuronal damage indirectly by infecting CD4+ lymphocytes, macrophages and other cells in the CNS (Ellis et al, 2009; Davis et al, 1992; Peluso et al, 1985). Mature viral particles feature a glycoprotein coat, or envelope, that allows them to identify, attach to, and enter these specific cell types (Ellis et al, 2009). The neurodegenerative process is unleashed by HIV itself, and once inside the CNS, HIV produces distinctive pathological changes (Trujillo et al, 2007; Masliah et al, 2000).

There is little evidence for direct invasion of the CNS by cell-free viruses (Dunfee et al, 2009; Brouwers et al, 1995; Epstein and Gendelman, 1993). The primary immune cell targets of HIV infection within the CNS are microglia/macrophage cells derived from monocytes (Dunfee et al, 2009; Ellis et al, 2009; Zink et al, 2006; Williams et al, 2001; Takahashi et al, 1996). Outside the CNS, HIV infects T-Lymphocytes (T-cells) and monocytes through virus interaction with the CD4+ receptor and co-receptors on the surface of these cells. HIV infection of T-cells results in cell lysis and death, yet HIV infection of monocytes, while productive, does not kill these cells (Ellis et al, 2009). Actively infected monocytes in the blood traffic into the brain, whereupon they differentiate into a host of CNS cell types, most notably perivascular macrophages, perivascular microglia and resident microglia (Dunfee et al, 2009; Zink et al, 2006; Hickey and Williams, 1999). Once inside the brain, HIV particles and proteins can be localised within monocyte derived cells (Zink et al, 2006; Gabuzda et al, 1986; Wiley et al, 1986).

Primary HIV infection of the CNS results in neural tissue damage (Ellis et al, 2009; Brouwers et al, 1995; Davis et al, 1992; Navia et al, 1986; Sharer et al, 1986; Wiley et al, 1986) and neuron dysfunction and death are the indirect consequences of HIV infection of microglia and macrophages (Zink et al, 2006; Epstein and Gendelman, 1993; Pulliam et al, 1991). These infected cells secrete a battery of pro-inflammatory cytokines and other soluble factors including HIV proteins, which, over a sustained period of time, and in high concentrations, are toxic to nearby neurons (Ellis et al, 2009; Zink et al, 2006). Pro-inflammatory cytokines, neurotoxic metabolites and viral gene products, which may cause damage to cells, are expressed in the brain parenchyma shortly after infection and throughout the course of infection (Banks et al, 2006). Therefore neuronal cell damage may begin long before neurological symptoms

appear (Kolson, 2002). Infected or activated T-lymphocytes trafficking into the CNS may also contribute to the cytotoxic cascade (Ellis et al, 2009; Pulliam et al, 1991). HIV-1 infection of brain macrophages may limit the proliferation of nearby astrocytes (Tardieu et al, 1992). While HIV can infect astrocytes and endothelial cells (the cell types that form the blood brain barrier) to a limited extent, HIV does not productively infect neurons (Ellis et al, 2009; Takahashi et al, 1996; Wiley et al, 1986). Neurons and astrocytes can be destroyed by HIV-1 infected monocytic cells after adhesion of these cells to their membranes (Zink et al, 2006; Tardieu et al, 1992).

The pathological hallmarks of HIV include the multi-nucleated giant cell, which is formed by fusion of infected cells of monocyte lineage (Zink et al, 2006; Price et al, 1988; Navia et al, 1986). Other pathological features of CNS infection include perivascular infiltration of immune cells, white matter pallor, reactive astrocytosis (proliferation of astrocytes), microglial nodules, inflammation of the choroid plexus and various degrees of neuronal loss and damage to dendrites (Kaul et al, 2006; Masliah et al, 1997; Belman et al, 1988; Price et al, 1988; Wiley et al, 1986).

In summary, the neuropathological findings in HIV-1 infected brain tissue can be explained by the following process: the HIV infected monocyte (macrophage/microglia) is the initiator of the pathological process (Zink et al, 2006; Sharer et al, 1986). There is cell-to-cell communication between the HIV infected monocytes and astrocytes, which establish an interaction that amplifies production of glial proliferatory and neurotoxic factors (Ellis et al, 2009; Banks et al, 2006; Brouwers et al, 1995; Wiley et al, 1986).

The changes in the CNS are localised to sub-cortical structures, including deep white matter and basal ganglia, but may be found to some extent in the cortex (Ellis et al, 2009; Rausch and Stover, 2001; Belman et al, 1988). A striking finding in the brains of infected children is the presence of mineralisations consisting of calcium salts and iron in the basal ganglia and cerebral white matter (Rausch and Stover, 2001; Belman et al, 1988). Vascular inflammation is found frequently in children, and is characterised by lesions consisting of cuffing and infiltration of lymphocytes, monocytes and multinucleated giant cells into the walls of the small and medium sized parenchymal vessels (Rausch and Stover, 2001).

2.5 Central Nervous System Involvement in HIV

HIV encephalopathy and neurodevelopmental delay in HIV will be discussed together, as neurodevelopmental delay is a sign of encephalopathy in infants and young children.

Neurological and developmental abnormalities are frequent complications of HIV infection in children, especially in younger perinatally affected children (Potterton et al, 2009a; Potterton et al, 2009b; Baillieu and Potterton, 2008; Raskino et al, 1999; Nozyce et al, 1994; Belman, 1992). Development of encephalopathy is one of the most severe complications of HIV, but its frequency at different ages has long been disputed (Tardieu et al, 1992). HIV encephalopathy may occur very early in the course of HIV infection, and 88.1% of children who develop it do so within the first two years of life (Newell et al, 1998). This is more common in HIV infected children who present in early infancy and have rapid progression, as they may have been infected in utero during or before the last weeks of pregnancy, which is the period of fastest brain growth (Newell et al, 1998). In the absence of antiretroviral therapy and co-trimoxazole prophylaxis, about 20% of infants vertically infected with HIV develop a rapidly progressing form of the disease, and progress to AIDS or die within the first year of life (European Collaborative Study, 1994). These infants are more likely to be born to mothers with advanced disease or to be infected in utero than children whose disease course is slower (Ioannidis et al, 2004; Williams et al, 2001; Kuhn et al, 1999). Since the most rapid brain growth occurs in the first four years of life, children infected at an older age are less likely to suffer the major neuro-cognitive deficits seen among HIV infected infants and young children. In general, there is strong evidence of neuropsychological deficits associated with symptomatic paediatric HIV disease (Fowler, 1994).

In children with vertically transmitted HIV, infection of the CNS may occur early in the disease process while the CNS is still immature. The pattern of disease therefore depends on the stage of brain development when CNS infection occurs (Belman, 1990). Neurological and developmental signs are often markers of HIV-1 disease in infants and may precede other signs of disease progression (Bisiacchi et al, 2000; Belman, 1992). CNS involvement is the first AIDS-defining symptom in as many as 18% of paediatric patients (Tardieu et al, 2000; Gabuzda and Hirsch, 1986). There are

many risk factors for HIV encephalopathy, including the timing of infection, advanced stage of maternal disease at delivery, rapid disease progression with early immune suppression, high plasma viral load in infancy, as well as factors such as home environment and socioeconomic status (Robertson et al, 2008).

The timing and severity of HIV-related paediatric central nervous system disorders vary greatly (Epstein et al, 1986). The majority of HIV infected children have developmental delay and deficits in cognitive function, language and motor skills, impaired brain growth and loss of developmental milestones, as well as cerebellar, sensory and primitive reflex dysfunction (Potterton et al, 2009b; Baillieu and Potterton, 2008; Udgirkar et al, 2003; Blanchette et al, 2001; Pearson et al, 2000; Fragoso et al, 1999; Drotar et al, 1997; Wolters et al, 1997; Belman et al, 1996; Pollack et al, 1996; Chase et al, 1995; Nozyce et al, 1994; Belman, 1992; Belman et al, 1988; Epstein et al, 1986; Belman et al, 1985; Ultmann et al, 1985). Studies have shown that these delays may begin as early as four months of age, continue into the preschool years, and manifest as either global or selective delays in neurodevelopment (Chase et al, 2000; Wolters et al, 1995).

HIV-1 infection of the developing CNS of infants and children is characterised by either a progressive or static loss of previously acquired developmental milestones with cognitive, behavioural and motor manifestations (Udgirkar et al, 2003; Belman et al, 1988; Epstein et al, 1986; Sharer et al, 1986; Belman et al, 1985). Significant delays in mental and motor development of HIV infected infants have been found over the first two years of life (Blanchette et al, 2001; Chase et al, 2000; Smith et al, 2000; Gay et al, 1995). Such impairments are likely to become more severe as the children are expected to reach more complex and integrative milestones (Blanchette et al, 2001; Belman, 1992). Neural development does not cease after two years of age, and therefore further delays can be expected to evolve over time, which may be important indicators of disease progression (Blanchette et al, 2001; Chase et al, 2000; Gay et al, 1995).

2.5.1 CNS Abnormalities Associated with Encephalopathy

A variety of CNS abnormalities are associated with HIV infection. These include cortical atrophy, calcification of the basal ganglia and frontal white matter, ventricular

enlargement, and white matter low attenuation (Wolters et al, 1995; DeCarli et al, 1993). The most important finding from a developmental perspective is that of myelinopathy. During infancy the brain still experiences a period of rapid myelination, which coincides with the attainment of significant motor, cognitive, and behavioural milestones. In particular, myelination of the frontal and parietal regions of the brain continues throughout childhood and even into early adulthood. These regions are responsible for higher cortical functions, such as language, sequence and the integration of multiple stimuli. Destruction of the myelination processes of these areas by HIV will cause significant delays in higher functioning. Since many of the more advanced developmental skills do not develop until after two years of age, delays can be expected in areas such as visual–motor processing, verbal memory, processing speed, and sequential processing (Blanchette et al, 2001; Gay et al, 1995).

2.5.2 The Pathophysiology of Encephalopathy

An initial pathophysiological hypothesis is that early encephalopathy may be the consequence of pre-natal HIV-1 infection of the brain, inducing a reduction in brain cell proliferation during late pregnancy, leading to decreases in intrauterine brain growth and parenchymal atrophy (Tardieu et al, 2000; Tornatore et al, 1994). Neurodevelopmental impairment may be a marker of overall host susceptibility, as some infants may be more susceptible to HIV-related CNS disease, resulting in faster progression and greater mortality rate (Ellis et al, 2009). Infected monocytes with the ability to cross the blood-brain-barrier and adhere to neural cells may play a major role in the induction of HIV-1 related encephalopathy (Tardieu et al, 1992). An accumulation of the HIV-1 *nef* protein has been described in astrocytes from brains of children with early encephalopathy (Epstein and Gendelman, 1993), and the presence of viral protein in developing astrocytes may also limit their proliferation during late pregnancy.

2.5.3 Definition and Presentation of Encephalopathy

Encephalopathy was first described by A Belman in 1988. Since her description and definition, it has been widely studied, and more current definitions are now available. These include “The failure to attain or loss of developmental milestones or loss of intellectual ability, verified by standard developmental scales or neurophysiological tests” (Brady et al, 1995).

The WHO (2007) defines HIV encephalopathy as at least one of the following progressing over at least two months in the absence of another illness:

1) Gross discrepancy between the actual and developmental age; failure to attain or loss of developmental milestones; loss of intellectual ability; or
2) Progressive impaired brain growth demonstrated by stagnation of head circumference or
3) Acquired symmetric motor deficit accompanied by two or more of the following: paresis, pathological reflexes, ataxia, gait disturbances

Along with the above, the WHO (2007) recommends that brain computed tomography (CT) scan or magnetic resonance imaging (MRI) should be performed to exclude other causes.

Tindyebwa et al (2004), have developed a simplified definition of encephalopathy to cater for resource-poor settings, where diagnosis is dependent on the presence of two of the following for at least two months:

1) Any child who is HIV positive;
2) Lack of growth in head circumference assessed by serial measurements at least three months apart;
3) Neurodevelopmental milestones such as loss of skills (particularly motor) and lack of acquisition of skills;
4) Diffuse symmetrical hyperreflexia

Along with the above, Tindyebwa et al (2004) recommended lumbar puncture to exclude CNS infections.

The WHO (2007) and Tindyebwa et al's (2004) definitions are similar, but in resource-poor settings, a few problems may arise with the WHO (2007) definition, namely, to detect discrepancies between actual and developmental age requires standardised assessment tools which may not be available, loss of intellectual ability requires cognitive testing, which is not readily available, and requires equipment. Brain CT or MRI scans are expensive, and are not available in many rural hospitals or clinic settings. Several patterns of encephalopathy have been recognised: a) A subacute progressive course (the most severe), b) plateau (followed by deterioration or improvement) or c) static/stable course (Belman, 1994).

2.5.3.1 Subacute Progressive Encephalopathy

Subacute progressive encephalopathy, is characterised by a gradual and insidious loss of previously obtained cognitive and motor milestones, although children may remain relatively stable for extended time periods before a new loss is appreciated (Simpson and Berger, 1996; Armstrong et al, 1993).

2.5.3.2 Plateau Encephalopathy

Plateau encephalopathy refers to those children who do not progress in their development, but who may not show deterioration for lengthy periods of time. Nevertheless, these children eventually deteriorate and lose previously acquired skills. As a result, they experience severe cognitive and motor impairments, and this unique course has been associated with the presence of HIV in, or clearance from cortical and subcortical regions of the brain (Armstrong et al, 1993).

2.5.3.3 Static Encephalopathy

The third major type of encephalopathy that has been identified in children vertically infected with HIV has been termed static encephalopathy. This involves no deterioration of attained milestones, but is characterised by significantly delayed development; that is, acquisition of new skills at a much slower rate than would be anticipated based on the child's age. These children may display cognitive skills ranging from low-average to markedly impaired, depending on their individual rates of new development (Brouwers et al, 1995; Armstrong et al, 1993). In children with a static encephalopathic course, poor maternal nutrition, disturbed home environment, repeated periods of illness and hospital admissions all contribute to developmental delay and motor deficits (Ultmann et al, 1985).

2.5.4 Predisposing Factors to Encephalopathy

Risk factors for paediatric CNS involvement include the timing of infection, advanced maternal disease at delivery, rapid progression with early advanced immune suppression, high plasma viral load in infancy, and factors such as home environment and socioeconomic status (Robertson et al, 2008). In the absence of HAART, children with fast progression and more aggressive disease are likely to die before reaching the age of two years (Newell et al, 2004; Taha et al, 2000).

Those children with AIDS-defining illnesses in the first two years of life display very severe neurological and neurodevelopmental abnormalities (Chase et al, 2000; Pearson et al, 2000; Belman et al, 1996; Nozyce et al, 1994). The presence of hepatomegaly, splenomegaly, or lymphadenopathy in the first three months of life increases the likelihood of HIV encephalopathy and the risk of death 28-fold (Laufer et al, 2000).

2.5.4.1 Immunological Factors

A lower CD4 count and higher viral ribonucleic acid (RNA) load are associated with increased severity of disability, growth failure and slower attainment in milestones in infants and children with vertically transmitted HIV (Pearson et al, 2000; Belman et al, 1996; Pollack et al, 1996; Chase et al, 1995).

The risk of encephalopathy is higher among HIV infected children born to mothers with more advanced disease as measured by CD4 count and viral load at the time of delivery (Blanche et al, 1990). Children with more advanced degrees of immune suppression early in life and those with a high-plasma viral load in infancy also have higher rates of encephalopathy (Nozyce et al, 1994). Other virological factors associated with HIV-associated CNS disease are a detectable p24 antigen in mothers and infants (Tardieu et al, 2000), elevated plasma HIV RNA loads in the first year of life (Cooper et al, 2004), and, for later-onset encephalopathy, high HIV RNA loads in cerebrospinal fluid (CSF) (Tardieu et al, 2000; Sei et al, 1995). In general, higher CD4 cell counts, lower CD8+ cell counts, and higher CD4:CD8 ratios have been associated with better neurobehavioural functioning (Marcotte et al, 2003). Tardieu et al (2000) found that encephalopathy in the first year of life was associated with normal or high levels of CD4+ T-lymphocytes in contrast to later-onset encephalopathy, in which CD4+ T-lymphocytes were deficient. HIV infected infants with low percentages of activated CD8+ T-cells at ages one and two months had a better neurodevelopmental prognosis during the first three years of life than infants with high percentages of activated CD8+ cells, and this association was independent of several variables that predict outcome of HIV infection, as well as overall clinical prognosis. The implication of these findings is that the activation state of CD8+ cells is important for the effects of HIV infection on CNS development in infants. Efforts to influence this activation may have preventive or therapeutic effects (Mekmullica et al, 2009).

2.5.4.2 Environmental Risk Factors

Many social and clinical factors may also affect functional status and quality of life for HIV infected children. HIV infected children are exposed to and live with altered family dynamics, in foster homes or homes with ill parents who may also have poor access to care, or in child-headed households. The medications used to slow disease progression have many potential toxicities and side effects. In addition, both poor nutritional status and growth are common (Arpadi et al, 2000; Mckinney et al, 1998).

The cumulative effect of recurrent, frequent and severe episodes of opportunistic infections is also known to contribute to delayed motor development amongst HIV infected infants (Drotar et al, 1997). The motor skills of infants living with HIV are often delayed due to the child's decreased ability to interact with his/her environments during periods of illness. Increased hospital admissions and length of hospital stay may lead to developmental delay in these children (Cooper et al, 2004; Fiser et al, 2000). The most common reasons for hospital admissions in children with HIV are pneumonia and gastroenteritis (Meyers et al, 2000).

Two hundred million children in developing countries fail to reach their potential in cognitive development because of poverty, poor health and nutrition, and deficient care (Grantham-McGregor et al, 2007). This has far reaching implications for education, and potential earning power, and therefore may result in transmission of poverty to upcoming generations in subsequent years.

2.5.4.3 Timing of Infection

Between 20 and 40% of HIV-positive women transmit infection to their babies in utero and intrapartum periods in the absence of antiretroviral therapy, as well as postpartum via breast-feeding (Peckham and Gibb, 1995).

Vertically infected children are more likely to experience CNS disease, and children with intra-uterine infection are more likely than those with peri- or post-partum vertical infection to develop early onset, rapidly progressive and severe encephalopathy (Smith et al, 2000; Wolters et al, 1995). A positive culture result within the first 48 hours of life (indicating in-utero infection) is associated with a more rapid decline in neurobehavioural function, compared with perinatally infected children (Smith et al,

2000). McGrath et al (2006a) found that children who tested positive between birth and 21 days of life had the poorest development over the first 18 months of life.

The impact of the timing of HIV-1 infection on neurodevelopment is related to the stage of brain development (Davis et al, 1992). Early interference with myelination could affect projection fibres, as well as association and commissural connections resulting in global patterns of deficits with motor abnormalities. In contrast, later interference would largely affect the commissural connections and the connections with poles of the cerebral lobes affecting cognitive functions (Gendelman et al, 1994). Another possible mechanism of neurodevelopmental injury is HIV-enhanced apoptosis.

2.5.5 Encephalopathy in sub-Saharan Africa

It may be difficult to apply results from studies in developed countries to the populations in resource-poor settings, where there is also a high prevalence of malnutrition, malaria, opportunistic infections, as well as different child-rearing environments and a higher incidence of poverty.

Studies on neurodevelopment in HIV infected children in Africa have revealed that a large proportion of children have moderate to severe delay in mental, motor, and language development (Ferguson and Jelsma, 2009; Potterton et al, 2009b; Van Rie et al, 2009; Baillieu and Potterton, 2008; McGrath et al, 2006a; Drotar et al, 1997; Msellati et al, 1993). Young HIV infected children also demonstrate more frequent and severe delays in mental and motor development compared with older HIV infected children (Van Rie et al, 2009). It has also been shown that motor function is more severely affected than other components of neurodevelopment (Ferguson and Jelsma, 2009; Baillieu and Potterton, 2008; McGrath et al 2006a; Tindyebwa et al, 2004; Drotar et al, 1997). Contrasting results were found by Van Rie et al (2008) on a population of children in Kinshasa, in the Democratic Republic of the Congo, where mental and not motor delay was most severely affected by HIV infection. These children were aged 18-72 months, which indicates that cognitive function may be more severely impaired in older children. The high prevalence of delay in cognitive development may be due to differences in child-rearing practices and lack of exposure to educational toys. This was supported by the high frequency of severe delay in

mental development among healthy control children in their study. Van Rie et al (2008) found in the same study that motor development improved after commencement of HAART, most likely due to the positive effects on muscle strength.

2.5.6 Neurodevelopment in School-age Children

From the beginning of the epidemic, there have been consistent findings of significant neurological, developmental, cognitive, and language deficits in HIV infected infants and younger children (Coplan et al, 1998; Belman et al, 1996). Unfortunately, less is known about these children as they age. In general, the severity of neurological and neuropsychological deficit increases with the severity of HIV-related illness (Jeremy et al, 2005).

With the advent of antiretroviral treatment, a large number of HIV infected children are reaching adolescence (Brackis-Cott et al, 2009). Studies indicate that school-aged perinatally infected children present with significant learning problems, which affect their ability to perform in school, achieve appropriate developmental milestones, and function independently (Martin et al, 2006). Advances in the medical treatment of children with HIV not only have prolonged survival, but have also improved quality of life (Smith et al, 2006). Despite these advances, many children experience both the direct and indirect effects of HIV infection, which affect their ability to achieve at a school-level. Brackis-Cott et al (2009) found that a large proportion of HIV positive children were held back a grade in school, and eventually need to be placed in special education classes. The high rate of developmental delay in HIV infected children as compared with healthy controls underscores the need for screening for and prevention of neurodevelopmental delay at an early age and calls for access to early interventions, nutritional and care programs (Van Rie et al, 2007; Jin et al, 2007; Bridge et al, 2006; Powell et al, 2004; Eickmann et al, 2003).

Children with HIV infection who experienced an early AIDS-defining illness exhibit significant impairment in their overall cognitive and motor ability, relative to their HIV infected, but relatively healthy counterparts and non-infected peers (Abubakar et al, 2008a; Nozyce et al, 1994). Findings reported by Tardieu et al, (1995) suggest that other health markers, such as circulating CD4+ lymphocyte counts during the first year of life are predictive of cognitive functioning among vertically infected school-aged

children. Children who experience an AIDS-defining illness early in life score significantly lower on verbal, perceptual performance, quantitative abilities, and memory subtests than do either infected children with no AIDS-defining illness or non-infected children (Abubaker et al, 2008a).

The prevention of early AIDS-defining illnesses among young children with HIV infection is crucial for preventing poor neurodevelopmental outcomes (Abubakar et al, 2008a). With early, appropriate antiretroviral therapy, not only does the child stand a better chance of survival, but may also avoid significant developmental impairment due to decreased chances of an early AIDS defining illness. In the meantime, HIV infected children of school-going age may need additional educational and psychosocial resources in the classroom and regular re-evaluation of their progress and learning needs (Smith et al, 2006).

2.6 Socioeconomic Effects on Neurodevelopment

Poverty is associated with inadequate food, poor sanitation and poor hygiene which lead to increased infections and stunting in children. Poverty is also associated with poor maternal education, increased maternal stress and depression, and inadequate stimulation in the home (Bradley and Corwyn, 2002). All these factors detrimentally affect child development (Grantham-McGregor et al, 2007; Olness, 2003). Risk factors related to poverty frequently occur together, and the increased number of risk factors increases the developmental deficit (Gorman and Pollitt, 1996). Many children in developing countries are exposed to multiple risk factors for poor development including poverty, poor health and nutrition. A conservative estimate is that more than 200 million children under five years of age in developing countries are not developing to their full potential due to poverty and the associated health, nutrition and social factors (Grantham-McGregor et al, 2007; Walker et al, 2007). Walker et al (2007) have identified four key risk factors affecting development in children in resource-poor settings which warrant urgent intervention: stunting, inadequate cognitive stimulation, iodine deficiency, and iron deficiency anaemia.

Anaemia is a common condition in HIV infected children and contributes significantly to morbidity. The prevalence in HIV infected infants is determined to a large extent by the prevalence of other conditions that cause anaemia, such as malaria and

helminthic infections (Tindyebwa et al, 2004). The prevalence of malnutrition, especially micronutrient malnutrition, also contributes significantly to the prevalence of anaemia. HIV infected children have an equal prevalence of anaemia compared with uninfected children but have a higher case fatality rate (Tindyebwa et al, 2004). Other developmental risk factors include malaria, intrauterine growth restriction, maternal depression, exposure to violence, and exposure to heavy metals (Walker et al, 2007). Infants born to HIV infected women are more likely to suffer from congenital malaria than children born to uninfected women. Likewise, an increased frequency of malaria has been noted in HIV infected children, with associated higher levels of parasitaemia than in other children (Tindyebwa et al, 2004). Additionally HIV infected children are more likely to be anaemic during an episode of malaria compared with uninfected children (Tindyebwa et al, 2004). Children living in resource-poor settings in South Africa are at risk for all of the above-mentioned issues. Those children infected with HIV are at added risk due to immune factors, and the social factors surrounding the condition.

Developmental research has clearly established that both socioeconomic status and aspects of the home environment account for a significant proportion of the variance in cognitive functioning of healthy and preterm children (Brooks-Gunn et al, 1996; Bradley et al, 1989). Aspects of a child's home environment and their associations with CNS factors may explain some of the variability in the cognitive functioning of children with HIV-1 infection. That is, despite CNS pathology, protective mechanisms may promote cognitive development in children with HIV-1 infection or, conversely, risk factors may result in greater vulnerability to cognitive dysfunction (Coscia et al, 2001). As the child matures, the effect of poverty begins to snowball and there may be a decline in mental, motor and social/emotional development. Higher family income is associated with a more cognitively stimulating home environment, less maternal emotional stress and more positive parenting practices, which in turn are associated with higher cognitive outcomes (Linver and Brooks-Gunn, 2002). The length of time that a child is exposed to poverty may impact on the degree of developmental impairment (Duncan et al, 1994), The timing of exposure to poverty is important, with children born into a poor household being at greater risk than those exposed to poverty at a later stage of their childhood (Burgess et al, 2004).

Parents are central to child-care (Schor et al, 2003). The health and well-being of children are linked to their parents' physical, emotional and social health as well as their social circumstances and child rearing practices. Environmental stress is a known 'neurotoxin'. Children from homes with higher numbers of environmental risk factors have been shown to achieve lower scores on measures of cognitive performance (Hochhauser et al, 2008; Gutman et al, 2003; Bennett et al, 2002). A family which is preoccupied with life crises is less able to provide the stimulating and varied experiences which foster cognitive growth (McLoyd, 1998). The lack of items such as toys geared to teaching about shapes and colours has been shown to be associated with reduced Intelligence quotient (IQ) scores in children (Bradley et al, 1989). Decreases in emotional support and in cognitive stimulation have been shown to account for 33-50% of the decreases in academic performances in children under conditions of chronic poverty (Korenman et al, 1995). Socioeconomic disadvantage has been shown to have a negative educational effect on language and reading skills, as well as on the motor skills of children (McPhillips and Jordan-Black, 2007). Poor development leads to poor school achievement, and children who do poorly in school are likely to transfer poverty to the next generation. This loss of human potential is associated with more than a 20% deficit in adult income and has implications for national development (Grantham-McGregor et al, 2007).

Research from developed countries has identified three aspects of parenting that are consistently related to young children's cognitive and social-emotional competence: cognitive stimulation, caregiver sensitivity and responsiveness to the child, and caregiver affect (emotional warmth or rejection of child) (Walker et al, 2007). These child-rearing dimensions affect children from developed and developing countries in similar ways (Bradley and Corwyn, 2005). All these aspects of parenting are likely to be affected in the context of HIV, due to maternal stress and depression (Potterton et al, 2007).

Effective interventions are available that may reduce the developmental loss currently estimated to affect more than 200 million children under five years of age in developing countries (Engle et al, 2007). The most effective interventions are comprehensive programmes for younger and disadvantaged children and families that are of adequate duration, intensity, and quality, and are integrated with health and

nutrition services. Providing services directly to children and including an active parenting and skill-building component is a more effective strategy than providing information alone (Engle et al, 2007). This may be difficult to achieve in the South African context due to lack of resources, understaffing, and high patient case loads, especially in the HIV context. Interventions may need to be preventative in terms of providing information, adequate nutrition, and parental support.

2.7 Paediatric HIV and Growth

In the formative years of development, malnourishment affects all dimensions of health and development (Spurr, 1983). For children growing up in poverty, physical and mental development are intertwined. Measures such as weight, height, and head circumference are predictors of cognitive abilities, whilst morbidity is negatively associated with cognitive test scores (Bhargava, 1998).

HIV infected children in developing countries show a decline in length and weight within the first months of life, and eventually manifest a picture of chronic malnutrition (Tindyebwa et al, 2004). This may result from HIV infection itself, underlying disease such as tuberculosis, inadequate macro/micronutrient intake or a combination of any, or all of the above. At least 90% of HIV infected children experience wasting and nutritional depletions during the course of their illness. A high viral load in children is associated with increased risk of failure to thrive, while infections such as pneumonia, diarrhoea, and tuberculosis further exacerbate growth failure (Tindyebwa et al, 2004).

In 2006, the World Health Organisation (WHO 2006a) released two sets of child growth standards to replace the National Centre for Health Statistics References (Dibley et al, 1987) as a tool for growth monitoring, nutrition screening, and surveillance, as a clinical tool to assist with the diagnosis of malnutrition, and as a tool for nutrition research (Van den Broek et al, 2009; WHO, 2006a). These new standards were developed in accordance with the idea that children born in any region of the world and given an optimum start in life all have the potential to grow and develop to within the same range of height- and weight-for-age. The new WHO child growth standards, which will be used worldwide, provide a common basis for the analysis of growth data (WHO, 2006a).

Moderate malnutrition is defined as a weight-for-age between -3 and -2 standard deviations (SD) below the median of the WHO child growth standards (WHO, 2006a), similarly, moderate wasting and stunting are defined as weight-for-height and height-for-age respectively, between -3 and -2 SD (WHO, 2006a). Wasting can have a very rapid onset, but can also recover very quickly under favourable circumstances (WHO, 2006a). Stunting is an indication of slow skeletal growth and is often associated with poor socioeconomic conditions, chronic or repeated infections as well as inadequate nutrition. Stunting takes a long time to correct once adverse circumstances have been corrected as skeletal growth is a slower process than growth in body mass (WHO, 2006a). Children with growth failure may have decreased weight-for-height (wasting) and decreased height-for-age (stunting). Malnutrition initially causes wasting, while stunting is associated with prolonged malnutrition (Chantry and Moye, 2005).

Severe acute malnutrition (severe wasting and/or oedema) or severe stunting is defined as weight-for-age less than -3 SD (WHO, 2006a). Mid-upper-arm-circumference-for-age scores below -2 SD are associated with wasting which is an indication of acute malnutrition and severe growth disturbance (Abubakar et al, 2008b).

Childhood malnutrition is high among HIV infected children and the magnitude is even higher in developing countries where it is already endemic. Malnutrition may occur as a result of reduced calorie intake due to poverty or poor feeding practices, or may be due to the effect of the HIV on the gastrointestinal system which results in malabsorption (Winter and Miller, 1994). HIV infected children are at increased risk of malnutrition for many reasons, including: decreased food intake because of anorexia associated with illness, mouth ulcers, oral thrush, increased nutrient loss resulting from malabsorption, diarrhoea or HIV enteropathy (after deterioration of immune function from malnutrition or HIV, enteric pathogens may injure the intestinal mucosa causing malabsorption) (Bachou et al, 2006; Tindyebwa et al, 2004). Prolonged malabsorption leads to malnutrition which causes immunodeficiency, more rapid progression to AIDS, and increased infection by opportunistic pathogens (Winter and Miller, 1994). Certain antiretroviral drugs may contribute to malnutrition by causing diarrhoea and vomiting (Bachou et al, 2006). Increased metabolic rate because of infections, opportunistic infections, and the HIV infection itself, release cytokines

(tumour necrosis factor [TNF] alpha and cachetin) into plasma or tissues and may mediate weight loss in HIV infected children (Tindyebwa et al, 2004). The effect of malnutrition is compounded by the high burden and recurring infections and infestations in HIV infected children. In addition, HIV-positive mothers have higher rates of low-birth-weight babies and premature birth, which are risk factors for malnutrition and failure to thrive (Tindyebwa et al, 2004).

When formally assessed, growth hormone (GH) levels in HIV infected children, even in those with growth failure, are usually normal (Lepage et al, 1991; Laue et al, 1990). Potential causes of altered growth in HIV infected infants may be due to decreased peripheral sensitivity to GH, decreased levels of insulin-like growth factors (Matarazzo et al, 1994), or the release of cytokines secondary to HIV-1 infection which may then lead to ineffective action of growth hormone (Pollack et al, 1996). Cytokines are especially attractive as mediators of decreased growth because of their rapid appearance during active HIV infection. Interleuken-6 (IL-6) and macrophage inflammatory protein-1a (MIP-1a) have both been shown to be elevated during periods of rapid viral replication and also to have effects on bone activity (Cocchi et al, 1995; Fuller et al, 1995; Udagawa et al, 1995).

Regular monitoring of weight, height and head circumference is recommended in all HIV infected children (Chantry and Moye, 2005; Nachman et al, 2002; Coovadia and Meyers et al, 2001; Bobat et al, 1998). Chronic malnutrition has been associated with decreased intellectual capacity, and therefore decreased earning capacity in later life, which perpetuates the cycle of poverty (Bridge et al, 2006). In a child who is not meeting his growth milestones, vitamin and micronutrient deficiencies should be considered as well (Mintz, 1999).

A number of studies have investigated growth deficiencies in children infected with HIV. Nathan et al (2003) reported stunting, but not wasting in a sample of institutionalised HIV positive African children. Spira et al (1999) found that failure to thrive in the first few months of life was associated with increased risk of death or progression to AIDS in HIV infected infants in Rwanda. Bobat et al (2001) investigated the growth of HIV infected children born in Durban, South Africa. They found that HIV positive infants exhibited early and sustained low mean length-for-age and weight-for-

age z-scores but weight-for-length was within normal limits. The authors of this study concluded that HIV infected children have early and sustained stunting, but not wasting, although children with rapidly progressing disease are stunted and wasted and therefore have a very poor prognosis. In the European Collaborative Study (2003), growth faltering in terms of both weight and height was associated with vertically acquired HIV infection at least up to 10 years of age, extending previous findings from the European Collaborative Study (1995) and others (Arpadi, 2000) which were limited to early childhood. Differences in growth patterns between infected and exposed but uninfected children not only persisted but also significantly increased with age (Arpadi, 2000).

2.7.1 Growth and Neurodevelopment

There is an association between growth and neurodevelopmental delay in HIV positive children especially in children with advanced stages of the disease (Pollack et al, 1996). Infants with the most severe growth delay have significant cognitive and motor delay even if this only becomes apparent later (Pollack et al, 1996). Wiznia et al (1996) report a strong correlation between poor weight-for-age and decreased cognitive function and a low CD4 count. Similarly, Potterton et al (2009b) showed that poor weight-for-age was a predictor for both poor motor and cognitive function. Missmer et al (2000) found that a decreased height-for-age z-score is a strong predictor of decreased functional status, suggesting that height may be used as an indicator of disease severity. The delays in linear growth and neurodevelopment occur at different times suggesting that they represent two separate effects of high HIV viral burden or load (Pollack et al, 1996).

Stunting in early childhood has been associated with diminished adult human capital, including poorer cognitive development, behavioral problems, and lower schooling attainment, even after controlling for confounding factors such as parental schooling and socioeconomic status (Victora et al, 2008). The critical period for the effect of undernutrition on cognitive development is early in infancy (Emond et al, 2007). It has been found that weight gain from birth to two years had the strongest relationship with schooling outcomes followed by birthweight; weight gain from 24 to 48 months has a weak or no relationship to schooling outcomes. The magnitudes of the relationships are of economic and public health importance (Martorell et al, 2010). Head-

circumference-for-age is sensitive to the effects of chronic undernutrition, but is more closely related to genetic factors and may be influenced by perinatal insults such as birth asphyxia (Silva et al, 2005; Gale et al, 2004). One of the primary targets of HIV-1 infection is brain tissue, which also commonly results in reduced brain growth (Macmillan et al, 2001). Acquired microcephally has been noted by a number of authors and is usually due to impaired brain growth and cerebral atrophy secondary to encephalopathy (Mitchell, 2001; Macmillan et al, 2001; Mintz, 1999). Cortical atrophy is associated with disease progression and neurodevelopmental delay (Llorente et al, 2003; Pearson et al, 2000). Overall, the size of the head predicts the size of the brain, which is associated with Intelligence Quotient (IQ) – i.e. a bigger brain denotes better IQ (Wickett et al, 2000). The brain doubles its birth weight in the first year and triples it by age six (Gale et al, 2003). Head circumference at this age is about 93% of its final size, so measurements of head circumference made in adults are largely a reflection of brain growth during the first few years of postnatal life. The results of this study suggest that brain development during infancy and early childhood is more important than foetal growth in determining how well cognitive abilities are preserved in old age. Factors that promote brain growth during this period may help to protect against cognitive decline (Gale et al, 2003). Head growth in the first nine months of life and head growth between nine months and nine years of age are also related to cognitive function, regardless of head size at the beginning of these periods. Postnatal head growth is significantly greater in children whose mothers are educated to degree level or of higher socio-economic status (Gale et al, 2004).

2.8 The Link between Socioeconomic Status and Growth

One of the links between the effect of environmental risk and poor developmental outcome is sub-optimal physical growth (Bradley and Corwyn, 2002). The most vulnerable group of children with the highest prevalence of growth restriction are those under five years of age (WHO, 1995). Growth restriction in the early years of life increases the risk of mortality (Black et al, 2003), morbidity (Nannan et al, 2007), and leads to developmental delay and impairments (Siegel et al, 2005). Restricted social and economic resources have been associated with poor growth (Walker et al, 2000) and with poor developmental outcome (Bradley and Corwyn, 2002; 1996). Abubakar et al (2008b), and Walka and Pollitt (2000) report a significant relationship between socioeconomic status (SES) and height-for-age, and that the attainment of motor skills

is related to the child's anthropometric measurments. The psychomotor performance of children experiencing poor physical growth ranges from moderate delay for stunting, being underweight and small mid-upper-arm-circumference-for-age, to severe delay for those with poor head growth (Abubakar et al, 2008b). SES therefore has a significant effect on psychomotor development through its influence on height-for-age, weight-for-age and mid-upper-arm-circumference-for-age. This observed relationship can be interpreted by Bronfenbrenner's (1977) framework. This states that a child grows up in layered environments, ranging from proximal factors (in this case anthropometric status) to distal factors (in this case SES). Distal factors define the context for proximal factors (SES impacts on anthropometric status); and proximal factors have more impact on developmental outcome than distal factors (anthropometric status has stronger association with psychomotor development compared with SES). These findings by Abubakar et al (2008b) suggest that the relationship between SES and psychomotor outcome is mediated by anthropometrics.

There is no significant relationship between head-circumference-for-age and SES which may be explained by the large contribution of genetic factors to variance in head size. The association may mean that head circumference is the anthropometric measure that is less susceptible to influences of social and economic factors (Abubakar et al, 2008b). These findings therefore highlight the importance of monitoring growth in an infant population due to the impact of poor growth on psychomotor development (Abubakar et al, 2008b). Problems with psychomotor development in the early years may negatively impact on other aspects of child development since motor skills lead to developments in other areas such as social and communication skills (Giagazoglou et al, 2005). This is supported by earlier studies which indicate that the effect of undernutrition on locomotor development is the main pathway to other developmental deficits (Gardner et al, 1999).

Research indicates that the negative effects of poor growth can be reversed or at least minimised through early intervention (Gardner et al, 2005; Powell et al, 2004; Pollitt et al, 1997). Abubakar et al (2008b) suggest that in resource-limited contexts, interventions aimed at facilitating physical growth in children with poor growth outcomes will potentially have a significant effect on psychomotor development. Interventions among infants in resource-limited settings should address both physical

health and motor as well as psychosocial functioning (Engle et al, 2007; Walker et al, 2005; Pelto et al, 1999).

2.9 Highly Active Antiretroviral Therapy (HAART)

Guidelines for the treatment of HIV infection in children recommend the use of highly active antiretroviral therapy (HAART). This should include two nucleoside reverse-transcriptase inhibitors (NRTIs) and either a protease inhibitor (PI) or a non-nucleoside reverse transcriptase inhibitor (NNRTI) (Working Group on Antiretroviral Therapy, 2009; DOH, 2008, WHO, 2006b, Sharland et al, 2002). In HIV infected adults, combination therapy which includes a PI is associated with suppression of viral load, increased CD4 + T-lymphocyte counts, and lower mortality (Li et al, 1998). Several studies have demonstrated sustained increases in CD4+ T-cells in adults, even when virological failure has occurred after PI therapy (Deeks et al, 2000). Studies in HIV infected children also have found evidence of improvement in immunological status following therapy which included a PI (Wiznia et al, 2000). The proportion of children with virologic suppression to undetectable levels is generally smaller than the corresponding proportion of adults (Nadal et al, 2000; Nachman et al, 2000; Wiznia et al, 2000). This is due to the fact that the immune system is developing during childhood, and therefore the effects of treatment may differ between children and adults and between children of different ages.

2.9.1 Initiation Criteria for Antiretroviral Therapy

Eligibility for antiretroviral therapy (ART) is currently assessed on either clinical or immunological grounds. Treatment is started either when an individual's clinical condition deteriorates to moderate or severe disease (WHO category three or four) or when their immunological levels drop to a level indicating severe immune deficiency (WHO, 2006c). CD4 percentage is the most appropriate immunological criterion to screen HIV- infected children from birth to 12 years for ART eligibility (WHO, 2006b), as it identifies the highest proportion of children in the early stages of immunosuppression. On the other hand, CD4 count and Total Lymphocyte Count (TLC) may be less appropriate as they tend to fluctuate more (WHO, 2006b). Where access to CD4 percentage assays is not available, CD4 percentage obtained with a simple formula based on total white blood count, lymphocyte percentage and total CD4 count may improve the identification of children in need of treatment in resource

poor settings (Callens et al, 2008).

In developed countries the decision of when to initiate ART in children is based on clinical symptoms and assessment of CD4+ T-cell count or percent and HIV RNA viral load (Sharland et al, 2002). However, in many resource-limited settings these laboratory tests are not available. Therefore there is increasing interest in using alternative markers, such as total lymphocyte count (TLC) (Little et al, 2007). The WHO clinical staging system has also been widely used to initiate ARV's in resource-limited countries, particularly in the African region, and has proved useful in all levels of care (WHO, 2007). An evaluation of the Zambian guidelines revealed 20-fold higher mortality rates in children classified as being in WHO clinical stage four compared with those children in stages one and two, which confirms the association of clinical stage with prognosis (Walker et al, 2006). The WHO recommendations for assessment of stage three and four disease may require confirmatory diagnosis (WHO, 2005). One of the indicators of clinical stage four disease is HIV encephalopathy, which is a condition where clinical diagnosis can be made on the basis of signs or simple investigations. This indicates that a simple developmental assessment tool may be sufficient to diagnose this condition, and where laboratory testing is not available, ARV's could be started on this basis. The criteria for staging (WHO, 2005) are shown in Table 2.1 below:

Table 2.1 WHO Paediatric Clinical Staging System

	Condition	Diagnosis
Clinical Stage 1	Asymptomatic	Clinical
	Persistant Generalised lymphadenopathy	Clinical
Clinical Stage 2	Hepatosplenomegaly	Clinical
	Papular pruritic eruptions	Clinical
	Seborrhoeic dermatitis	Clinical
	Fungal nail infections	Nail scrape
	Angular cheilitis	Clinical
	Lincal gingival erythema	Clinical
	Human papilloma virus infection	Clinical
	Molluscum contagiousum infection	Clinical
	Recurrent oral ulceration	Clinical
	Parotid enlargement	Clinical
	Herpes zoster	Viral culture
	Recurrent respiratory tract infection	Clinical and x-ray
Clinical stage 3	Unexplained moderate malnutrition	Documentation and monitoring
	Unexplained persistent diarrhoea (>14 days)	Clinical
	Unexplained persistent fever (>month)	Clinical
	Oral candidiasis (outside first 6 weeks of life)	Microscopy or culture
	Oral hairy leukoplakia	Clinical
	Pulmonary TB	x-ray and sputum
	Severe recurrent presumed bacterial pneumonia	clinical
	Acute necrotising gingivititis, or stomatitis, or acute necrotising periodontitis	clinical
	Systemic lymphoid interstitial pneuomonitis	x-ray, oxygen saturation
	Chornic HIV-associated lung disease (including bronchiectasis)	x-ray
	Unexplained anaemia or neutropenia	lab testing
Clinical stage 4	Unexplained severe wasting or severe malnutrition	Documentation and monitoring
	Pneumocystis pneumonia	Sputum and chext x-ray
	Recurrent severe presumed bacterial infection	Clinical
	Chronic herpes simplex virus infection	Histology or culture
	Oesophogeal candidiases	Clinical
	Extrapulmonary TB	Blood culture
	Karposi's sarcoma	Biopsy
	Cytomegalovirus retinitis and infection of organs other than lymph nodes, liver or spleen	Histology
	CNS toxoplasmosis	CT scan and CSF culture
	Cryptococcal Meningitis	CSF culture
	HIV encephalopathy	Clinical and recommendation of CT scan to exclude other causes
	Disseminated mycosis	Microscopy or histology
	Candidiasis of the trachea, bronchi or lungs	Microscopy or histology
	Disseminated mycobacteriosis other than TB	Stool, blood, body fuid culture
	Cryptosporidiosis	Microscopy
	Isosporiasis	Response to therapy
	Cerebral or B cell non-Hodgkins lymphoma	CNS imaging
	Progressive multifocal leukoencephalopathy	MRI or CT scan
	Acquired HIV-Associated rectal fistula including recto-vaginal fistula	Clinical
	HIV-associated nephropathy	Urine testing
	HIV-associated cardiomyopathy	Echocardiography

WHO 2005 Interim WHO Clinical Staging of HIV/AIDS , page 30

Treatment for HIV in low-income and middle-income countries (LMICs) is at present driven by a public health approach, where the primary goal is to provide ART to as broad a population as possible, in settings in which individualised management of patients by specialised physicians is not feasible (Bartlett and Shao, 2009). The issue of when to start HAART in perinatally HIV infected children is topical (Prendergast et al, 2007). Results of a clinical trial investigating the impact of early HAART on African

children have recently been reported where early initiation of HAART resulted in a 76% reduction in mortality, and decreased disease progression by 75% at a 10 month follow-up (Violari et al, 2008). Rapid control of HIV infection is especially important in resource-limited countries with high rates of opportunistic infections. Therapy initiation guidelines would be useful for South African physicians while awaiting the results of ongoing trials of different CD4 count thresholds for therapy initiation (Walensky et al, 2009).

2.9.2 Treatment Initiation in Developed Countries

Based on the European (Sharland et al, 2002) and US guidelines (Working group on ARV Therapy, 2009), children with clinical or immunologic deterioration should receive ART. Controversy remains regarding early treatment in infants with no or mild/moderate signs or normal immunologic status. The risk of disease progression is higher in children than in adults, with15-20% of HIV-1 infected children progressing to AIDS or death within the first year of life (HIV Paediatric Prognistic Markers Collaborative Study Group 2003; Gray et al, 2001; Tovo et al, 1992). The commencement of combined ART within the first six months of life in HIV-1 perinatally infected children has been shown to result in avoidance of clinical manifestation, preserved CD4+ T-lymphocyte percentage, and no increase in CD8+ T-lymphocyte percentage, at least for the first years of life. Clinical benefits from early ART were also shown (Chiappini et al, 2006). However, in children under two years of age neither CD4+ T-lymphocyte percentage nor viral load can identify those at higher risk of disease progression, supporting the universal treatment strategy (Dunn, 2003).

Early ART has been recently proposed for the treatment of asymptomatic or moderately symptomatic infants with HIV-1 infection (Working group on ARV Therapy, 2009). As the rate of viral replication during the first months of life is correlated with disease outcome, early therapeutic intervention, which keeps the virus at low levels during primary infection, may lead to a better long-term viral suppression and preserved immune system function (Working group on ARV Therapy, 2009). Moreover, one possible goal of early ART is the prevention of HIV-mediated damage of the developing nervous system which is particularly frequent in infants (Sanchez-Ramon et al, 2003).

2.9.3 Treatment in Developing Countries

ART has been available in South Africa since 2004, although strict guidelines for initiating treatment have had to be followed due to a shortage of funding for drugs (DOH, 2008). Children, like adults, are eligible for ART on the basis of clinical symptoms or signs, or of immunological status (Dabis et al, 2000), but disease progression in children is more complex than in adults, and starting treatment at an appropriate time to ensure efficacy is often difficult. Where ART can be provided to children as they progress to moderate or severe disease it has been shown to significantly and substantially improve their chances of survival. In developing countries the rollout of a comprehensive paediatric care package encompassing regular monitoring, ART, prophylaxis and other supportive therapy is therefore likely to cause a ballooning of the demand for ART (Little et al, 2007). In developing countries, the commencement of HAART along with the frequent problems of advanced disease, poor access to healthcare services, lack of infrastructure, inadequate nutrition and poor living standards may result in poorer outcomes compared with developed regions (Cowburn et al, 2007; Feucht et al, 2007). Despite these challenges, HAART has been shown to decrease mortality and morbidity in poorly-resourced settings (Violari et al, 2008; Cowburn et al, 2007; Reddi et al, 2007).

In order to facilitate the rapid scale-up of paediatric antiretroviral treatment (ART), the World Health Organisation (WHO) published revised paediatric ART guidelines for resource-poor settings in 2006 (WHO, 2006b). The eligibility criteria for infants and children rely on clinical and/or immunological thresholds and aim to identify those children with poor prognosis if ART initiation is delayed. The WHO guidelines are as follows: The threshold CD4 levels for severe immunodeficiency are <25% for infants <11 months, <20% for children aged 12-35 months, or <15% for children aged 3 years and above (WHO, 2006b).

In comparison with World Health Organisation Guidelines, the South African guidelines have slightly lower CD4 thresholds, and therefore initiation of ART would commence later. In South Africa, children are eligible for ARV's based on the following criteria from 2008 (DOH, 2008):

Clinical Criteria

- Confirmation of diagnosis of HIV-infection **AND**
- Recurrent (>2 admissions per year) hospitalisations or prolonged hospitalisation (>4 weeks) for HIV-related illness **OR**
- The patient satisfies the WHO Stage III/IV disease **OR**
- For relatively asymptomatic patients, one can consider CD4 percentage <20% if under 18 months or <15% if over 18 months.

Social Criteria

These criteria are extremely important for the success of the programme and need to be adhered to – the principle is that adherence to treatment must be at least probable.

- At least one identifiable caregiver who is able to supervise the child for the administration of the medication.
- Disclosure to another adult living in the same house is encouraged so that there is someone else who can help with the child's ART.

The DOH goals for HAART therapy are as follows:

- To increase survival and decrease HIV-related morbidity and mortality.
- The child's CD4 count should rise and remain above the baseline count.
- The child's viral load should become undetectable <25 copies/ml and should remain undetectable on ART (DOH, 2008).

On the second of December 2009, President Jacob Zuma announced revised guidelines for ARV initation in South Africa in certain vulnerable groups of people. The most significant changes, which will commence as from the first of April 2010 are: pregnant women living with HIV will qualify for ARV's if their CD count falls below 350 cells/mm^3; and pregnant women with higher CD4 counts will be given treatment from 14 weeks of pregnancy to attempt to prevent mother-to-child transmission. All HIV positive children under the age of one year will be eligible for treatment regardless of their CD4 count. For children aged between one and five years, those with clinical stage three or four signs, or a CD4 percentage of ≤ 25 are eligible to commence ART. Children eligible to be fast tracked are those under one year of age, those with signs of stage four disease, and those with multidrug resistant tuberculosis. This will allow

children to access treatment earlier, and for many, as soon as they are diagnosed (Department of Health, 2010).

WHO immunological thresholds for total CD4+ T-lymphocyte count (CD4 count) and CD4 percentage are associated with improved prognosis, and correspond to a five percent decrease in annual mortality risk in children from the United States and Europe aged one year or older (HIV Pediatric Prognostic Makers Collaborative Study, 2005). The same level of immunosuppression in children living in resource poor settings may not have the same prognosis due to differences in environmental factors such as opportunistic infections or malnutrition (Bachou et al, 2006; Lugada et al, 2004). Embree et al (2001), found that CD4 percentage and CD4 count in healthy Kenyan children was markedly lower than European, American and West African children. Therefore the thresholds established through longitudinal studies in the United States and Europe may not be extrapolated to other regions in the world.

In the pre-HAART era, African children experienced rapid and more aggressive HIV disease with higher mortality (Obimbo et al, 2004). As programmes providing HAART scale up in Africa, survival of HIV-1 infected children is expected to improve, based on what has been observed in Western settings (Gortmaker et al, 2001). In a review of 30 studies done in sub-Saharan Africa, which described the treatment of HIV infected children (Sutcliffe et al, 2008), it was found that children are entering treatment programmes at older ages in sub-Saharan Africa. Most children in these studies were older than five years when they started antiretroviral therapy. In a pooled analysis of mortality among HIV infected treatment-naive African children, an estimated 35% and 53% died within one and two years, respectively (Newell et al, 2004). By age five years, other studies have estimated that 62–89% of children have died (van der Loeff et al, 2003). Only a few studies reported on how children came to be enrolled in the antiretroviral therapy programmes: most were identified as HIV infected and clinically in need of treatment through health-care services, rather than in infancy through Prevention-of-Mother-to-Child-Transmission (PMTCT) or voluntary counselling and testing programmes, when children are often asymptomatic (Wamalwa et al, 2007). Therefore, older children with slower disease progression are more likely to gain access to antiretroviral therapy in sub-Saharan Africa. By contrast, nearly two-thirds of HIV infected children who would have benefited from life-prolonging treatment before

reaching age five years are not being diagnosed and treated (Wamalwa et al, 2007).

By definition, children eligible for treatment by CDC or WHO guidelines have moderate to severe immunosuppression (WHO, 2006b). However, with a median CD4+ T-cell percentage of less than 10% in most studies, children in sub-Saharan Africa are starting antiretroviral therapy at more advanced stages of disease (CD4 T-cell percentage of approximately five to ten percent lower) than children in high-income countries, thereby putting them at an increased risk for treatment failure, and poor immunological response (Resino et al, 2006).

2.9.4 Regimens of ART

Antiretroviral regimens including Protease Inhibitors (PI's) have been shown to increase CD4 cell counts, even for children with advanced disease (Storm et al, 2005; Nachman et al, 2000; Wiznia et al, 2000), and studies are now beginning to document their effects on clinical outcomes. Baseline and treatment-mediated changes in immunologic and virologic markers were independent predictors of survival in a meta-analysis of paediatric antiretroviral clinical trials, and virologic markers predicted weight, growth and cognitive failure among children one year of age (Lindsey et al, 2000). The use of PI therapy has been found to reduce morbidity and mortality rate among HIV infected children and adolescents (Gortmaker et al, 2001), and has been shown to increase height and weight growth (Verweel et al, 2002; Buchacz et al, 2001).

In a study by Lindsey et al (2007), PI-based HAART was shown to have a positive but limited impact on neurodevelopmental trajectories and the rate of significant cognitive and motor impairment in HIV-positive infants and young children during the first three years of life, despite substantial effects on survival and immunologic status. This limited positive effect of HAART may be partly because infants and young children with a compromised CNS now survive the more severe early manifestations of the disease but may have persistent neurobehavioral difficulties. The effects of HAART may become more apparent as the children reach school age and beyond and as differential and characteristic patterns of strengths and weakness in domains of development are identified.

2.9.5 ART in sub-Saharan Africa

Resource-limited settings pose unique challenges to the implementation, effectiveness, and sustainability of HAART programs (Cowburn et al, 2007; Feucht et al, 2007). Limited availability of paediatric formulations of antiretroviral drugs, weak healthcare infrastructure and the widespread conviction that paediatric specialists are needed to initiate and monitor HAART due to its perceived complexity serve as barriers to comprehensive implementation (Kim and Gilks, 2005). Threats to effectiveness and sustainability in the paediatric population include morbidity and mortality of primary caregivers due to rising HIV prevalence rates, and although the expectation of poor adherence in this setting leading to widespread "antiretroviral anarchy" has not been met, treatment failure and exhaustion of available regimens are ongoing risks (Harries et al, 2001). Children are entering treatment programmes at older ages in sub-Saharan Africa, as many mothers have no access to PMTCT or voluntary counselling and testing programmes, and therefore children are only identified as being HIV-positive and clinically in need of treatment when they become symptomatic (Sutcliffe et al, 2008). Therefore, older children with slower disease progression are more likely to gain access to antiretroviral therapy in sub-Saharan Africa (Sutcliffe et al, 2008).

2.9.6 The Effects due to Timing of Initiation of ART

Soh et al (2003) found that improvements in CD4 percentage were greater in younger children, than in older children on PI-based therapy, which may substantiate the argument to initiate ART's. Early deaths occurred in children with advanced HIV disease within five months of treatment initiation. This suggests that a greater effort should be made to identify and treat HIV-positive children at an earlier stage of their infection, and that children with advanced disease should be closely monitored and supported at the onset of HAART (Reddi et al, 2007).

Although long-term HAART allows restoration of CD4+ cell percentages and control of viral loads in HIV-1 infected children, HAART initiation after severe immunosuppression may be less effective for restoration or maintenance of a normal CD4+ cell percentage. These data argue in favour of not delaying initiation of HAART in young children (Resino et al, 2006).

2.9.7 Effects of ART on Growth

Growth seems to be one of the most sensitive indicators of disease progression in children with acquired immunodeficiency syndrome (AIDS), with the absence of growth indicating a poor prognosis, even in children who are treated with antiretroviral regimes (Lindsey et al, 2000). HIV-1 infected children show a trend towards an increase in height and weight after the initiation of HAART, although in children with a better virologic response to therapy, these increases are greater (Verweel et al, 2002).

Increasing CD4+ T-cell counts favourably influence weight and body mass index (BMI). BMI z-scores increase more in children with an advanced clinical stage of infection at initiation of HAART due to the fact that the energy expenditure previously needed to combat infection was possibly used for catch-up growth (Verweel et al, 2002). Height increases after weight, which is a normal reaction to the correction of a growth-retarding disorder: catch-up growth first affects weight followed by height. Sustained low viral loads, and higher CD4 counts help to support this (which are dependent on successful application of HAART) (Verweel et al, 2002).

At the initiation of HAART, viral loads correlate negatively with height *z*-score change in the previous year and the clinical stage of infection correlates negatively with BMI z-score and weight z-score (Verweel et al, 2002). Although the number or percentage of CD4+ T-cells has a larger influence on prognosis, increasing growth rates also contribute to a better prognosis (Carey et al, 1998). The positive effects of HAART on growth can be sustained for at least 96 weeks (Verweel et al, 2002), and the effect on growth lasts longer in patients receiving HAART than in patients receiving mono or duo reverse transcriptase inhibitor therapy. There seems to be a relation between the time that antiretroviral therapy is successful in the suppression of viral load and the time that the positive effect on growth by this therapy can be maintained (Verweel et al, 2002).

Malnutrition is associated with disease progression and death in untreated children (Obimbo et al, 2004), and is associated with an increased risk of mortality among children receiving treatment (Reddi et al, 2007). In all studies that assessed this issue, nutritional status improved after antiretroviral therapy; however, the effect of malnutrition on virological success and immune reconstitution has not been fully

explored. Treatment programmes need to address malnutrition and micronutrient deficiencies through an integrated model of care or by partnering with organisations that have the resources to provide nutritional supplementation (Agnarson et al, 2007).

2.9.8 Effects of ART on the CNS and Neurodevelopment

Studies of cerebrospinal fluid (CSF) from children with HIV encephalopathy found active and persistent brain infection with HIV suggesting a need for antiretrovirals that penetrate the blood–brain barrier (Epstein et al, 1987). To aid clinicians in antiretroviral therapy decision-making for patients with neurological symptoms, this information has recently been used to develop a drug ranking system based on a drug's ability to penetrate the CNS (Patel et al, 2009). To rank antiretroviral regimens according to their ability to penetrate the CNS, a modified version of the CNS penetration-effectiveness rank developed by Letendre et al, (2008) has been created.

A scale of one (lowest penetration) to three (highest penetration) has been used to rank each antiretroviral drug. A CNS-penetration score is calculated for each antiretroviral regimen by summing the individual ranks of each antiretroviral drug included in the regimen. Antiretroviral regimens with scores of less than four are classified as low CNS-penetrating regimens, scores of four to five are classified as medium CNS-penetrating regimens, and scores greater than or equal to six are classified as high CNS-penetrating regimens. The score allocated to each drug by Patel et al (2009) is shown in Table 2.2 below:

Table 2.2 CNS Penetration Classification

1 (lowest penetration)	2 (medium penetration)	3 (highest penetration)
Didanosine (ddI)	Emtricitabine (FTC)	Abacavir (ABC)
Tenofovir (TFV)	Lamivudine (3TC)	Zidovudine (ZDV)
Zalcitabine (ddC)	Stavudine (d4T)	Delavirdine (DLV)
Nelfinavir (NFV)	Efavirenz (EFV)	Nevirapine (NVP)
Saquinavir (SQV)	Amprenavir (APV)	Amprenavir/ritonavir(APV-r)
Saquinavir/ritonavir (SQV-r)	Atazanavir (ATV)	Atazanavir/ritonavir (ATV-r)
Tipranavir/ritonavir (TPV-r)	Foasamprenavir (f-APV)	Fosamprenavir/ritonavir(f-APV)
Enfuvirtide (T-20)	Indinavir (IDV)	Indinavir/ritonavir (IDV-r)
		Lopanavir/ritonavir (LPV-r)

(Patel et al, 2009 Impact of HAART and CNS-penetrating antiretroviral regimens on HIV encephalopathy among perinatally infected children and adolescents. AIDS 23: 1 – 9 (page 3).

A ten-fold decrease in the incidence of HIV encephalopathy among a large prospective cohort of perinatally infected children enrolled over a 14-year period, 1993–2006 was found (Patel et al, 2009). This dramatic decline in incidence occurred with a concurrent increase in the use of HAART, proposing an association between HAART use and risk of HIV encephalopathy. HAART use was found to decrease the risk of HIV encephalopathy by 50% compared with no HAART use (Patel et al, 2009). Children who were on a high CNS penetrating regimen had a 41% decrease in the incidence of encephalopathy compared with those children on a low CNS penetrating regimen (Patel et al, 2009).

HAART may therefore inhibit or delay HIV dissemination in the CNS and may also decrease viral replication if an active and persistent infection is already established in the brain (Patel et al, 2009). One study found a significant decrease in CNS viral load with increasing numbers of CNS-penetrating antiretroviral drugs independent of HAART alone (DeLuca et al, 2002). Optimal levels of antiretroviral drugs within the CNS, over what some HAART regimens provide, are needed to stop the active replication of HIV within the brain that can lead to further neurological decline and eventually death (Cysique et al, 2004). Further research is required to assess the impact of antiretroviral CNS-penetration effectiveness on the various pathogenic mechanisms leading to neurological deterioration and disease. Highly CNS-penetrating regimens result in a substantial survival benefit to children with HIV encephalopathy compared with low CNS-penetrating regimens (Van Dyke et al, 2008).

2.10 Developmental Screening and Surveillance

Developmental screening is the administration of a brief standardised tool that aids the identification of children at risk for a developmental disorder. Developmental screening does not result in either a diagnosis or treatment plan but rather identifies areas in which a child's development differs from same-age norms. Screening is a brief assessment procedure designed to identify children who should receive more intensive diagnosis or assessment (Dworkin, 1993).

Developmental screening has been shown to improve the accuracy with which children's developmental delays are identified when compared with decisions based

only on clinical judgment (Dearlove and Kearney, 1990). Previous studies indicated that the most common developmental screening technique used in the primary care setting is informal clinical assessment and that few health care practitioners use standardised developmental screening tests routinely (Rydz et al, 2006). The performance characteristics of screening tools demonstrate their accuracy to be well above that of informal methods (Glascoe, 1997). Unfortunately, clinical assessment alone detects less than 30% of children with developmental disabilities (Rydz et al, 2006). In contrast, standardised developmental screening instruments have been reported to have sensitivities and specificities of 70% to 90% (Glascoe, 1997).

Surveillance is a continuous process which is performed by skilled professionals, and involves observations of children during the provision of health care. The components of developmental surveillance include attending to parental concerns, obtaining a relevant developmental history, making accurate observations of children, and sharing opinions and concerns with other relevant professionals (Dworkin, 1993). Paediatricians often use age-appropriate developmental checklists to record milestones as part of developmental surveillance.

Surveillance and screening have limitations due to the dynamic nature of development (American Academy of Paediatrics [AAP], 2006). Repeated and regular screening is more likely than a single screening to identify problems, especially in later-developing skills such as language (AAP, 2006). Waiting until a young child misses a major milestone such as walking or talking may result in late rather than early recognition, increasing parental dissatisfaction and anxiety and depriving the child and family of the benefits of early identification and intervention (AAP, 2006). The advantages of standardised developmental screening instruments are that they provide a set of norms against which to measure development, serve as a reminder to the paediatrician to observe development, are efficient ways to record observations, and identify children with delays. The major disadvantage is that they take time to administer and interpret. Therefore, developmental screening instruments are not widely used in paediatric practice (AAP, 2006; Glascoe, 2001). Although the use of screening tests would improve the rate and accuracy of identification, a recent survey demonstrated that only 23% of primary care clinicians used a standardised screening tool (Sand et al, 2005).

2.10.1 The use of Standardised Assessment Tools in sub-Saharan Africa

The adequate monitoring of disease effects and intervention among children in rural communities in Africa is hampered by a lack of appropriate assessment tools (Holding and Kitsao-Wekulo, 2004). Instruments developed in one culture cannot be readily transferred to another, despite evidence that cognitive abilities of children do not vary across cultures (van de Vijver, 1997). The shortage of personnel with a background in child assessment within an African context has been identified as another limitation to the use of tests that require prior experience in child development and extensive training (Olness, 2003).

Other challenges to the application of Western instruments in a non-Western context arise from a lack of familiarity with test demands, incomparability of samples and poor translation of test items (Holding and Kitsao-Wekulo, 2004; van de Vijver, 1997). Problems with applying and adapting standardised assessment techniques in Africa begin with the often prohibitively high price of Western materials (Aina and Morakinyo, 2001) and are compounded by the shortage of trained and qualified test administrators (Olness, 2003; Haataja et al, 2002). The numerous challenges described highlight the need to develop culture-appropriate items, administration procedures and the establishment of culture-specific norms for the interpretation of score levels (Abubakar et al, 2007).

In the absence of existing tests there are different approaches that have been applied to the production of new assessment measures (van de Vijver and Tanzer, 2004). The most common approach is translation of existing measures, but this may constrict the within-population variance and mask true group differences (Connolly and Grantham-McGregor, 1993). On the other hand, the production of a novel assessment limits the comparability of outcomes across different cultural settings (Sternberg et al, 2001). A third option is to assemble a selection of activities into a new measure, and to ensure that items included are not only acceptable to the target population but also evaluate constructs common to those measured by published tests (Abubakar et al, 2007). This is potentially important as it enables the comparison of disease effects across sites and contexts (Holding and Kitsao-Wekulo, 2004).

2.10.2 Properties of Acceptable Screening Tools

The best screening instruments have good psychometric properties, including adequate sensitivity, specificity, validity, and reliability, and have been standardised on diverse populations (American Academy of Pediatrics, 2001).

● Reliability is the ability of a measure to produce consistent results.

● Validity of a developmental screening test relates to its ability to discriminate between a child at a determined level of risk for delay (ie, high, moderate) and the rest of the population (low risk).

● Sensitivity is the accuracy of the test in identifying delayed development.

● Specificity is the accuracy of the test in identifying individuals who are not delayed (AAP, 2006)

For developmental screening tests, scoring systems should minimise both under-referrals and over-referrals. Trade-offs between sensitivity and specificity occur when devising these scoring systems in order to achieve this (AAP, 2006). None of the above properties of screening tests can be used in isolation to assess the effectiveness of a test, but together they provide the multiple perspectives for evaluating a test's validity (Meisels, 1989).

2.11 Development and Statistical Properties of Screening Tools

Aylward (2009) proposed that there is no real 'gold standard' in developmental evaluation due to the level of variation in the tests, testee, and tester. Development, when assessed using a developmental test, may be a moving target, due to two factors: (1) the Flynn effect (Flynn, 1999) and (2) revisions of developmental assessment instruments. With the former, the mean of a test increases 0.3-0.5 points per year, this roughly equating to five points per decade. Therefore, it may be better to update norms with minimal alterations in test content. This raises conceptual and financial issues and the argument that changes in knowledge of development should be reflected by changes in tests (Aylward, 2009). Regardless, clinicians should not unquestioningly accept a test as being the gold standard. Knowledge of the test's strengths, weaknesses, and relationships to other tests is critical. In the literature reviewed on the development of screening tools, a certain sequence of steps may be followed when developing a new tool or adapting an existing tool. These include item selection or generation, validation and reliability testing, and standardisation.

2.12 Item Selection

In a review of the literature on the development of screening tools, item selection is largely based on the following:

- Review of the literature (Kirby et al, 2010; Bart et al, 2010; Martin et al, 2009; Bayley, 2006a; Bayley 2006b; Bayley 2006c; Harris et al, 2003; Beckung, 2000)
- Review of similar screening tools (Kirby et al, 2010; Harris et al, 2003)
- Items taken from an existing assessment tool (Gladstone et al, 2008; Bayley, 2006c; Aylward, 1995)
- Interviews with people who have the relevant condition (Bart et al, 2010)
- Clinical experience (Bayley, 2006a; Harris et al, 2003; Beckung, 2000)

2.13 Statistical Properties of Screening Tools

Screening tools should have the following statistical attributes (American Psychological Association, 1985):

2.13.1 Validity

There are a number of tests that an assessment tool should undergo in order to determine validity. Validity is defined as “the degree to which a useful (meaningful) interpretation can be inferred from a measurement” (Rothstein et al, 1991, pg 597). According to the Standards for Measurement in Physical Therapy these include criterion-based validity, construct validity, and content validity (Rothstein et al, 1991).

2.13.1.1 Criterion Validity

Criterion Validity consists of three aspects: concurrent validity, predictive validity and prescriptive validity. Rothstein et al (1991) define these as follows:

2.13.1.2 Concurrent validity

The interpretation is justified by comparing the new measurement to be validated with another validated measurement at approximately the same time. The measurement of concurrent validity should be provided when the new test claims to be able to give information about the current status of a person at the time that the measurements are obtained (Rothstein et al, 1991). Concurrent validity is the degree to which outcomes on one test correlate with outcomes on a criterion test, when both tests are given at the same time (Portney and Watkins, 2000). Correlations of 0.6 or greater should be

obtained (American Psychological Association, 1985). Other values which can be obtained following studies for concurrent validity are sensitivity, specificity, positive predictive value and negative predicitive value.

2.13.1.3 Predictive Validity

Predictive validity examines the justification of using a measurement to say something about future events or conditions (Rothstein et al, 1991). Information about predictive validity should be provided when a measurement claims to be able to provide information about future status (Rothstein et al, 1991). Marks et al (2008) express concern at conducting predictive validity studies on developmental screening tools, as by design, screening tools are developed to provide information on the current status of a child, and not give an indication of future outcome. The American Academy of Pediatrics (2006) state that a screening tool should generate a referral to early intervention, where the child may receive more in-depth assessment, or monitoring through surveillance.

2.13.1.4 Prescriptive Validity

Prescriptive validity refers to the validity of a test when the results are used to determine what treatment the person will receive (Rothstein et al, 1991), information regarding prescriptive validity should be provided when the the results of a measurement will be used to determine treatment options.

2.13.1.5 Construct Validity

This is the theoretical basis for using a measurement for a specific purpose (Rothstein et al, 1991). Evidence for construct validity is through logical argumentation based on evidence from theory and research.

2.13.1.6 Content Validity

Content validity is an essential step in the development of new measuring devices as it represents a beginning mechanism for linking abstract concepts with observable and measurable indicators (Wynd et al, 2003). Content validity is often the first step in examining the validity of a newly developed tool and is not determined by statistical analyses, but by practical considerations (McEwan et al, 2003; American Psychological Association, 1985). This is often carried out in consultation with a panel

of experts to obtain consensus on each item, as well as the consensus on the scale as a whole (Dunn et al, 2005; Harris et al, 2003; Harris and Daniels, 1996). Specifying the objectives of the test and providing the rationale for item selection enhance the content validity of a test. A theoretical basis should be provided, as well as the source of items and proof of systematic and logical analysis of item development (American Psychological Association, 1985). When assessing content validity, it is necessary to define the domain of content of the concept being measured and then to determine whether this is adequately covered by the instrument. The more elements within the concept that are actually assessed by the instrument, the greater the instrument's content validity (Sim and Arnell, 1993).

Content validity is largely a matter of judgement, involving two distinct phases: the scale developer should enhance content validity through careful conceptualisation and domain analysis prior to item generation, and the relevance of the scale's content through expert assessment (Polit and Beck, 2006; Mastaglia, et al, 2003; Beck and Gable, 2001). Expert assessment can be carried out in a number of ways: These include averaging experts' ratings of item relevance and using a pre-established criterion of acceptability (Beck and Gable, 2001); using coefficient alpha to quantify agreement of item relevance by three or more experts, and computing a multirater kappa coefficient (Wynd et al, 2003), having a team of experts indicate whether each item on a scale is congruent with (or relevant to) the construct, computing the percentage of items deemed to be relevant for each expert, and then taking an average of the percentages across experts (Polit and Beck, 2006). A validity questionnaire may be drawn up, and responses to questions are rated on a likert scale (Parks et al, 2007; Harris and Daniels, 1996).

Other methods to obtain consensus include the Delphi Technique and Nominal Group Technique. The Delphi technique is a method for the collection of opinion on a particular topic. It is based on the premise that 'pooled intelligence' enhances individual judgement and captures the collective opinion of experts (Jones et al, 2000). It typically involves 15-30 participants (de Villiers et al, 2005). The conventional Delphi technique uses a series of questionnaires to aggregate expert opinion in an anonymous fashion. This takes place over a series of 'rounds'. Communication can take place by post or by electronic exchange (Jones et al, 1992). A question is

formulated and expanded upon in a set of assumptions, solutions or options. Secondly, an expert panel is identified and invited to provide opinions. The responses are analysed and ranked, using predetermined criteria for agreement and disagreement. A second questionnaire is developed using the results and feedback from the first round. Participants again record their opinions, which are collated and assessed for consensus (Jones and Hunter, 1995). The process terminates when an acceptable degree of consensus is reached. Three rounds are usually sufficient to achieve consensus with the largest adjustments occurring between rounds one and two (Bellamy et al, 1991).

Another method to achieve consensus is the Nominal Group Technique (NGT) (Potter et al, 2004), and has been recommended by the authors as a technique to achieve consensus amongst experts in the development and review of outcome measures. The NGT has been shown to be a reliable method to establish the content validity of new instruments (Hyrkas et al, 2003). The aim of the NGT is to generate information in response to a focused issue, and to facilitate discussion around this. This is done during a focused meeting, which lasts up to two hours, and consists of five to nine participants. There are a number of advantages to this technique: minimal preparation is required by participants, and participants input is limited to a single meeting lasting up to two hours. Task completion and immediate dissemination of results to the group lead to greater satisfaction amongst participants due to quick dissemination of information; and researcher-bias is minimised due to the highly structured nature of the process (Potter et al, 2004). The NGT protocol follows a number of steps as outlined below:

1. Introduction and explanation: the participants should be welcomed, and an explanation of the purpose and procedure of the meeting should be given.

2. Generation of ideas: each participant should be provided with a sheet of paper with the question to be addressed and participants are then asked to write down all ideas that come to mind when considering the question. During this period, participants are asked not to consult or discuss their ideas with others. Approximately ten minutes are allowed for this.

3. Sharing of ideas: participants should be invited to share the ideas they have generated. The process continues until all ideas have been presented. There is no debate about items at this stage and participants should write down any new ideas that may arise from what others have shared. This process ensures that all participants get an opportunity to make an equal contribution and provides a written record of all ideas generated by the group.

4. Group discussion: participants are invited to ask for further explanation and detail about any of the ideas that colleagues have produced.

5. Voting and ranking: this involves prioritising the recorded ideas in relation to the original question. Immediate results are available to participants so that the outcomes are reached by the time the meeting concludes. It is recommended that the process is recorded on videotape or audiotape (Potter et al, 2004).

A modified version of the NGT was successfully used during the content validation process of the Gross Motor Function Classification Measure in order to facilitate group decision-making with clearly focused outcomes (Palisano et al, 2008), and consensus was defined as agreement with a question by at least 80% of participants (Palisano et al, 2008). It has been used in a number of other studies to evaluate the content validity of new tools (Dyrbye et al, 2010; Shea et al, 2007; van der Camp et al, 2006).

2.13.2 Evaluation of Diagnostic Tests

A diagnostic test is used to screen for the presence or absence of a disease or abnormal condition. The quality of a screening test therefore should be evaluated against the findings obtained by a test which is considered to be the ‘gold standard’ (Portney and Watkins, 2000). The validity of a diagnostic test is therefore evaluated in terms of its ability to accurately assess the presence or absence of a specific condition, which is done through the evaluation of the sensitivity and specificity of the test:

2.13.2.1 Sensitivity, Specificity, Positive Predictive Value and Negative Predictive Value

The sensitivity, specificity, positive predictive value, and negative predictive value can be used to analyse the criterion validity of a screening test (Glascoe, 1997). Sensitivity and specificity of developmental screening tools are measured by comparing the test results to that of gold-standard developmental evaluation tools (Glascoe, 1997).

Sensitivity is a measure of validity of a screening procedure based on the probability that someone who has the condition will test positive (Portney and Watkins, 2000). Sensitivity is computed as: (the number of children with delays who were identified)/([those identified] + [children with delays who were not identified by the screening test]). Glascoe (1997) recommends that a screening test should have sensitivity levels of 70-80% due to the nature and complexity of measuring the continuous process of child development. Meisels (1989) suggested that more conservative criteria should be used with sensitivity levels of no less than 80% in order to avoid too many over-referrals. The value of 80% is also recommended by the American Psychological Association (1985).

Specificity is a measure of validity of a screening procedure based on the probability that someone who does not have the condition will test negative (Portney and Watkins, 2000). It is computed as (the number of normal children who also scored 'normal' on the test)/(normal children + normal children identified by a test as having delays) (Glascoe, 1991). There often is a trade-off in specificity if the sensitivity of a measure is very high. Again, Glascoe (1997) recommends that a screening test should have specificity levels of 70-80% due to the nature and complexity of measuring the continuous process of child development. Meisels (1989) suggested that more conservative criteria should be used with specifity levels of no less than 80% to avoid under-referrals. The recommendation by the American Psychological Association, (1985) is 90%. This indicates that there is a wide variation in acceptable levels of specificity, but when combining the above recommendations, it should be between 70-90%.

Positive predictive value is an estimate of the likelihood that a person who tests positive actually has the condition (Portney and Watkins, 2000). The lower the

prevalence of a disorder, however, the lower will be the positive and negative predictive values. A positive predictive value of 50% would be acceptable in developmental screening, although 30% and above is often found, because some of the low average children will be over-referrals (Aylward, 1997). The negative predictive value is the proportion of patients who are correctly diagnosed by the test as not having the condition (Altman and Bland, 1994). This result is less commonly presented.

2.13.3 Reliability

A test should be reliable – that is, it should be consistently free from error (Portney and Watkins, 2000). There are a number of types of reliability which a test should possess:

2.13.3.1 Interrater Reliability

Interrater reliability is the degree to which two or more raters can obtain the same ratings for a given variable (Portney and Watkins, 2000), and should be 0.8 or greater (American Psychological Association, 1985). Interrater reliability is best measured when all raters are able to measure a response during a single trial (Portney and Watkins, 2000). If interrater reliability has not been established, it cannot be assumed that other raters would have obtained the same results, which therefore limits the application of the findings (Portney and Watkins, 2000). Studies should have been conducted in a clinical context consistent with the intended use of the measurements (Rothstein et al, 1991). A number of raters may evaluate reliability, and may range from two to multiple numbers of raters. Videotapes may be useful in allowing multiple raters to view and rate the same performance (Smits-Engelsman et al, 2008; Bodkin et al, 2003; Portney and Watkins, 2000; Gowland et al, 1995), although this may be problematic, as it is difficult to capture all aspects of a performance, or all body parts on a video, and some studies have found that higher reliability coefficients have been found during observations rather than assessing a videotaped performance (Gowland et al, 1995).

2.13.3.2 Test-retest Reliability

Test-retest reliability is the degree to which an instrument is stable, based on repeated administrations of the test to the same individuals over a specified time period (during

which individuals are not expected to change) (Portney and Watkins, 2000) and should be at least 0.8 (American Psychological Association, 1985). Within-day and between-day studies should have been conducted in a clinical context consistent with the intended use of the measurements (Rothstein et al, 1991).

In the reviewed literature on developmental screening and performance tests, the time between testing for test-retest reliability differed significantly, and ranged from two days to thirty days (Thomas-Stonell et al, 2010; Bayley, 2006a; Saigal et al, 2005; Harris and Daniels, 2001; Squires et al, 1997; Coster, 1995; Frankenberg et al, 1992). Portney and Watkins (2000) state that the interval between tests should be considered very carefully: they should be far enough apart to avoid fatigue, learning or memory effects, but close enough to avoid changes in the variable to be measured. They state that the primary criteria for choosing an appropriate interval are the stability of the response variable, and the purpose of the test. In addition, they also state that measurements of infant development might need to be taken over a short period to avoid the rapidly occurring changes of early infancy. Caution should be given to carry-over effects, which may occur with repeated measurements, and concepts such as motor learning, which may result in improvements in the first test score (Portney and Watkins, 2000).

2.13.3.3 Intrarater Reliability

This refers to the stability of data recorded by one individual over two or more tests. This is usually assessed using trials with short intervals in-between where carryover effects are not an issue (Portney and Watkins, 2000). In a test-retest situation, when a rater's skill is relevant to the accuracy of the test, intrarater and test-retest reliability are essentially the same estimate (Portney and Watkins, 2000).

2.14 Standardisation Sampling

Normative tables should be drawn up after administering the test to a large number of normal children, whose performance becomes the test's norms. The standardisation sample should be determined by the country's census using racial, geographic, and socioeconomic distributions. At least 100 children per age interval, which is usually 6-12 months should be included (American Psychological Association, 1985).

2.15 Comparison of Screening Tools

In order to illustrate the variability in development and validation of screening tools, six commonly used screening tools have been selected, and their properties have been compared.

The following discussion attempts to compare and contrast the development and psychometric properties of the following developmental screening tests: The Ages and Stages Questionnaire (ASQ) (Squires et al, 1997); The Alberta Infant Motor Scale (AIMS) (Piper et al, 1992); Denver II (DDST) (Frankenberg et al, 1992); Harris Infant Neuromotor Test (HINT) (Harris, 1991); Bayley-III Screening Test (Bayley, 2006c), and the Bayley Infant Neurodevelopmental Screener (BINS) (Aylward, 1995).

2.15.1 Description of the Screening Tools

The Ages and Stages Questionnaire (ASQ) (Bricker et al, 1995) is a parent-completed child monitoring system, which was first published in 1995, and was a revision of the infant/child-monitoring questionnaire from the 1980's. The ASQ evaluates gross motor, fine motor, communication, problem-solving, and personal-social domains of development of children aged 4-48 months, and takes 10 to 20 minutes to complete.

The Alberta Infant Motor Scale (AIMS) (Piper et al, 1992) is a motor observational assessment requiring minimal handling of the infant, and can be administered in 10 to 20 minutes by any health professional with a background in infant motor development. The AIMS was developed to provide a valid and reliable measure of motor development for infants at high risk of motor delay (Piper et al, 1992). Unlike previous standardised developmental assessments, the AIMS focuses on the motor requirements to attain each milestone, rather than looking only at when the milestone was achieved.

The Denver Developmental Screening Test (DDST) was first standardised and published in 1967, and is used to assess multiple aspects of child development. In the early 1990's, the DDST underwent revision and re-standardisation and has since been known as the Denver-II (Frankenberg et al, 1992). The Denver-II items are assessed by observation, parental report, or direct elicitation in the following domains: personal-social, fine motor adaptive, language, and gross motor. The administration time is 10

to 20 minutes (Frankenberg et al, 1992).

The Harris Infant Neuromotor Test (HINT) is a screening test that was first developed in 1991 to identify neuromotor, cognitive, or behavioral concerns in infants between the ages of 2.5-12.5 months. The administration time is about 30 minutes. The test's reliability, validity and standardisation were completed over the following decade (Harris et al, 2003).

The Bayley Infant Neurodevelopmental Screener (BINS) (Alyward, 1995) was developed for children ages 3–24 months and is based on the Bayley Scales of Infant Development, second edition (BSID-II) (Bayley, 1996). It assesses basic neurological function/intactness, receptive function, expressive function, and cognitive processes, and takes about 10 minutes to administer.

The Bayley-III Screening Test (Bayley, 2006c) is an individually administered instrument designed to briefly assess the cognitive, language and motor functioning of infants and young children between the ages of 1-42 months of age.

2.15.2 Standardisation Samples of the Six Tests

The standardisation samples of the tests are compared and contrasted below. The numbers in the standardisation samples and the stratification of each sample are shown in Table 2.3 below:

Table 2.3 Standardisation Samples of the Six Tests

Tool	Standardisation sample
BINS	600 infants with both clinical and nonclinical concerns and was stratified according to age, sex, ethnicity, region, and parent education
ASQ	2008 children from Oregon, Hawaii, and Ohio. Included children at risk for developmental delay due to medical or environmental risk factors (81%) and children who were typically developing with no known risk factors (19%). Diverse ethnic and socioeconomic backgrounds were represented
Denver-II	2096 children from Colorado; stratified by age, gender, socioeconomic status, maternal education, culture
HINT	412 children from British Columbia, Canada. The normative sample was stratified by gender, maternal education, and ethnicity.
AIMS	Normative sample of 2202 from Alberta, Canada, Stratified by age and gender only
Bayley-III Screening Test	1675 children ages 16 days to 43 months 15 days who were demographically representative according to the United States Census, as well as children with specifically chosen conditions.

The standardisation sample for the ASQ was 2008 children from Oregon, Hawaii, and Ohio (Bricker et al, 1995). The sample included both children at risk for developmental delay due to medical or environmental risk factors (81%) and typically developing children with no known risk factors (19%). Diverse ethnic and socioeconomic backgrounds were represented, although the authors noted that the Hispanic population was underrepresented and the Native-American population was overrepresented (Bricker et al, 1995).

The AIMS was standardised on a sample of 2202 infants who were recruited exclusively from the province of Alberta, Canada. The sample was stratified by age and gender, but ethnic and socioeconomic characteristics were not reported (Coster, 1995), which could indicate that the test may not be suitable for use on all groups.

The standardisation sample for the Denver-II was a sample of 2096 children from the state of Colorado, from birth to six years of age. This sample was stratified by gender, maternal education, ethnicity, and socioeconomic status (Frankenberg et al, 1992).

Recent normative data for the HINT were collected for 412 Canadian infants from the provinces of British Columbia, Manitoba, Nova Scotia, Ontario, and Quebec (Harris et

al, 2003). The sample was stratified by gender, maternal education, and ethnicity.

The standardisation sample for the BINS consisted of 600 infants with both clinical and non-clinical concerns and was stratified according to age, sex, ethnicity, region, and parent education (Aylward, 1995).

The standardisation of the Bayley-III Screening Test consisted of 1675 children from all over the United States, between the ages of 16 days and 43 months, 15 days who were demographically representative of the US population (Bayley, 2006c). Children were excluded from participating if they had any history of medical complications, and were not currently diagnosed with mental, physical or behavioural difficulties (Bayley, 2006c). A number of special group studies were conducted in order to determine its suitability for use on these groups. The groups included children at high-risk for developmental delays due to a variety of conditions (Bayley, 2006c).

As can be seen from the above, the standardisation samples for the six tools vary in size, stratification, and time that standardisation was carried out as compared with when the tool was developed. Only the Denver-II meets all the criteria outlined for standardisation samples in terms of size and stratification, which is outlined in 2.14 above (American Psychological Association, 1985).

2.15.3 Validity of the Tests

Concurrent and preditictive validity of the six tests are described below. The values for each test are shown in Table 2.4 below:

Table 2.4 Validity of the Tests

Tool	Predictive Validity	Concurrent Validity
Bins	Assessed against BSID II and has been found to be predictive of outcome at 36 months	Mental Developmental Index of the BSID II (r = 0.40) Motor Scale of the BSID II (r = 0.35) Sensitivity: 75 – 86% Specificity: 75 – 86%
ASQ	Not assessed	Compared results of 1511 ASQ assessments and results on: Revised Gesell and Amatruda Developmental and Neurologic Examination and BSID (<3 years). For >3 years, The Stanford-Binet Intelligence Test-4th edition and the McCarthy Scales of Children's Abilities Assessment administered within 29 days of ASQ completion by parent or caregiver. Sensitivity:75% Specificity: 85% PPV 46%
Denver-II	Not assessed	Suspect w/normal group Sensitivity=56% Specificity=80% PPV= 37% Suspect w/abnormal group Sensitivity=83% Specificity=43% PPV=23%
HINT	Administered between 2.5 and 12.5 months, and compared to scores on the BSID II Mental and Motor Scales at 17 to 22mnths.Correlation between HINT and the Bayley Motor Scale=-0.49($p < 0.01$), predictive validity of the HINT and the Bayley Mental Scale was poor (r=-0.11)	With Bayley II Mental:r= -0.73 Motor: r= -0.89 Sensitivity and specificity testing underway
AIMS	164 high-risk infants recruited from Edmonton, Alberta. Correlation between AIMS, and MAI, and PDGMS at 4 months compared with evaluation by paediatrician at 18 months. At 8 months: sensitivity: 86.4% Specificity:93% PPV: 65.5%	120 infants from 0 to 13 months of age on the Motor Scale of the Bayley Scales of Infant Development (r=0.98) and the Peabody Developmental Motor Scales (r=0.97)
Bayley-III Screening Test	Not assessed	Against the BSID III on 1657 children At-Risk: moderate (41.82% - 65.91%) Emerging: accurate (63.87% - 77.78%) Competent: very accurate (83.84% - 92.11%)

2.15.3.1 Validity of the ASQ

For the ASQ, concurrent validity was calculated by comparing the results of 1511 assessments using the ASQ and the results obtained on the Revised Gesell and Amatruda Developmental and Neurologic Examination, the Bayley Scales of Infant Development II (<3 years), the Stanford-Binet Intelligence Test-4th edition and the McCarthy Scales of Children's Abilities (>3 years) (Bricker et al, 1995). The assessments were administered within 29 days of ASQ completion by parent or

caregiver. The sensitivity and specificity of the ASQ across all age intervals were determined to be 75% and 85% respectively, with a positive predictive value of 46%. In the second edition of the ASQ, published in 1999, two additional questionnaires were added for the age intervals 10, 14, 16, 22, 33, 42, 54, and 60 months (Bricker and Squires, 1999). Neither reliability nor concurrent validity of the second edition has been examined, nor has predictive validity been examined for either the first or second edition.

2.15.3.2 Validity of the AIMS

In evaluating the AIMS's concurrent validity, the authors assessed 120 infants from birth to 13 months of age on the Motor Scale of the Bayley Scales of Infant Development II and the Peabody Developmental Motor Scales. Concurrent validity was measured using the Pearson product-moment correlation. Correlation coefficients were r=0.98 for the Motor Scale of the Bayley Scales of Infant Development II and r=0.97 for the Peabody Developmental Motor Scales (PDMS) (Piper and Darrah, 1994).

The predictive validity of the AIMS was assessed in a sample of 164 infants at high risk recruited from two neonatal intensive care units in Edmonton, Alberta (Darrah et al, 1998). The predictive validity of the AIMS was compared with the predictive validity of the Movement Assessment of Infants (MAI) and the Peabody Developmental Gross Motor Scale (PDGMS). A physiotherapist administered the AIMS, MAI, and PDGMS to each infant at four and eight months of age and was blinded to the infants' medical history. At 18 months of age, a follow-up assessment was conducted by a developmental paediatrician who classified the infants as normal, suspect, or abnormal in terms of motor development. The paediatrician's evaluation was based on the following criteria: postural control, muscle tone, reflexes, and achievement of motor milestones. At 18 months, 78% of the infants were classified by the paediatrician as normal, 8.5% as suspect, and 13.4% as abnormal in their motor development. For predictive validity analyses, infants at 18 months receiving a suspect outcome were either grouped with those classified as normal or with those receiving an abnormal classification. Predictive values were determined for all three measures initially administered at four and eight months. When grouping infants with suspected delay with normal infants at the 18-month outcome assessment, the AIMS

and MAI showed similar sensitivity at four months but the MAI provided greater specificity (Darrah et al, 1998). When infants were assessed at eight months, however, the AIMS demonstrated greater specificity than the MAI for the 18-month outcomes. The PDGMS did not show an acceptable combination of sensitivity and specificity until its cutoff score was set at the 16th percentile rank at four months. When infants suspect for delay were grouped with abnormal classifications, a grouping proposed by Glascoe et al (1992), the sensitivity of the AIMS and MAI decreased for assessments at both four and eight months while the sensitivity of the PDGMS was poor at four months and its specificity was poor at eight months. Grouping suspect with abnormal outcomes maximised both the sensitivity and specificity of the AIMS: 77.3% and 81.7% at four months and 86.4% and 93.0% at eight months, respectively. The positive predictive value for the AIMS was 39.5% at four months and 65.5% at eight months.

2.15.3.3 Validity of the Denver-II

Although the revision and re-standardisation of the DDST attempted to address previous inadequacies, the Denver-II was published and distributed prior to studying its concurrent validity or predictive accuracy (Harris et al, 2005). In 1992, a study evaluating the accuracy of the Denver-II in identifying children with atypical development was published (Glascoe et al, 1992). This consisted of 17 examiners who assessed five children each. One hundred and forty one items had 100% agreement, seven had 90-99% and one had 83%. To examine the test's concurrent accuracy, Glascoe et al (1992) assessed the correspondence of categorical classifications on the Denver-II with categorical classifications on other reference screening tests in order to determine the sensitivity, specificity, and positive predictive value of the Denver-II. One hundred and four children, ranging in age from three to 72 months, were recruited. All children were assessed initially using the Denver-II. Within seven days of that assessment, each child was assessed on a battery of standardised assessments, including the Vineland Adaptive Behavior Scale, the Kaufman Assessment Battery for Children Achievement Subtests, the Fluharty Preschool Speech and Language Screening Test, and one of the following cognitive tests: the Bayley Scales of Infant Development II, the Stanford- Binet Intelligence Scale, 4th edition, or the Kaufman Assessment Battery for Children. By applying the results from this battery of tests to criteria for special education eligibility, the presence or absence

of developmental delay was determined and these classifications were compared with those from performance on the Denver-II. When questionable scores were grouped with normal scores, the Denver-II yielded a sensitivity value of 56%, specificity of 80%, and a positive predictive value of 37%. When grouping abnormal scores with suspect scores, the sensitivity of the Denver-II was 83%, specificity was 43%, and positive predictive value was 23%. Based on these findings, the American Academy of Pediatrics (2001) noted that ‘‘the Denver-II screening test is used widely but has modest sensitivity and specificity depending on the interpretation of questionable results.’’ Despite attaining the desired level of sensitivity by using this recommended grouping, more than 50% of typically developing children would then be classified as suspect for developmental delay. As a result, developmental screening with the Denver-II could lead to further diagnostic examination of about 60% of the children tested. Such low specificity of a screening test produces concern for an unnecessarily high referral rate.

A study published by Glascoe and Byrne (1993) examined the relative accuracy of the Denver-II, the Developmental Profile II, and the Battelle Developmental Inventory Screening Test when compared with scores on a battery of standardised assessments. Specificity for the Denver-II was unacceptably low (46%) when grouping questionable with abnormal scores, although sensitivity was 83%. Using this grouping, the positive predictive value was only 28%, however, when grouping questionable scores with normal scores, specificity increased to 80% but sensitivity fell to 56%. The positive predictive value using this grouping was 42%. Glascoe and Byrne (1993) concluded that the Denver-II ‘‘produced more incorrect than correct classifications.’’ Although revision and re-standardisation represented an improvement over the original DDST, the failure to address the revised test’s accuracy based on standard criteria for the development of new tests is a major shortcoming, and the suggested validity studies have not yet been undertaken.

2.15.3.4 Validity of the HINT

Concurrent validity was assessed by comparing HINT total scores with raw scores on the Mental and Motor Scales of the BSID II for 54 infants at high risk (Harris et al, 2005). Both tests were administered at the same assessment session. The Pearson r was used to calculate the relationship between the HINT and the BSID-II scales. For

the HINT and the Bayley Mental Scale, the concurrent validity was r= -0.73, whereas for the HINT and the Bayley Motor Scale, the relationship was r= -0.89.The predictive validity of the HINT, administered between 2.5-12.5 months, was assessed by comparing those scores to scores on the Bayley Mental and Motor Scales at 17 to 22 months. The predictive correlation (Pearson r) between the HINT and the later Bayley Motor Scale was -0.49 ($p < 0.01$). However, the predictive validity of the HINT for the Bayley Mental Scale was r= 0.11. Interestingly, the predictive correlation for the HINT to the later Bayley Motor Scale was stronger than the relationship between the early Bayley Motor Scale and the later Bayley Motor Scale (0.49 and 0.34). Sensitivity and specificity analyses of the HINT's predictive accuracy with the Bayley-II at 17–22 months are currently in progress and involve a sample of 119 infants, approximately half of whom were at high risk and half of whom were at low risk for developmental delays (Harris et al, 2003).

2.15.3.5 Validity of the BINS

Concurrent validity of the BINS was assessed against the BSID II, on 199 infants at four age groups. Pearson correlations were calculated in order to evaluate the agreement between the BINS and the BSID II Mental Developmental Index (MDI), and results ranged from 0.43 to 0.82 (Aylward, 1995). For agreement between the BINS and the BSID II Psychomotor Developmental Index (PDI) score, results ranged between 0.39 to 0.58. All correlations were low to moderate apart from the result of 0.82 for the 24 month age group against the BSID II MDI (Reid and Rigby, 1997). For the BINS, validity as a screening tool has been evaluated by examining sensitivity, specificity, and positive predictive value (PPV). The sensitivity of the test is 80%, and the specificity of the test is 89% (Aylward, 1995). Aylward (1995) did not report the ranges obtained during the sensitivity and specificity studies, and specificity is thought to lie between 50 and 70% (Reid and Rigby, 1997), which indicates that the BINS may overidentify children as having developmental problems. In terms of predictive validity, the BINS has been highly correlated with the BSID II at 36 months (Aylward, 2000). More validity studies need to be conducted in order to clarify all the different results obtained, and establish more precise values for validity of the BINS (Reid and Rigby, 1997).

2.15.3.6 Validity of the Bayley-III Screening Test

Validity was assessed against the BSID III on 1657 children (Bayley, 2006c). For children who obtained scores in the At-Risk category of the Bayley-III Screening Test, classification was moderate (41.82%-65.91%) – none were identified as proficient. For children who obtained scores in the Emerging category of the Bayley-III Screening Test, classification was accurate (63.87%-77.78%), with 0.82%-5.21% being identified as At-Risk. For children in the Competent category of the Bayley-III Screening Test, classification was very accurate (83.84%-92.11%), with no children being misidenitifed as At-Risk (Bayley, 2006c).

2.15.4 Reliability

Reliability of the six tests are compared and contrasted for the following: interrater reliability, intrarater reliability and test-retest reliability, which were the most commonly reported reliability properties of the tools. The values for each test are shown in Table 2.5 below:

Table 2.5 Reliability of the Tests

Tool	Interrater reliability	Intrarater reliability	Test-retest reliability
BINS	.79 to .96	Not described	.71 to .84
ASQ	94% (SEM = 0.12)	94%	94%
Denver-II	141 items had 100% agreement, 7 had 90 – 99% and one had 83%.	Not assessed	Excellent agreement for 59% of items (k≥ 0.75), 23% of items: fair to good agreement (k≥ 0.40).
HINT	0.99.	0.98 to 0.99	0.98.
AIMS	0.996 and above	0–18mo (n=195) ICC=0.9915 3–18mo (n=45) ICC=0.98–0.9952	0.86 to 0.99
Bayley-III Screening Test	Not described	Cognitive 0.85 Receptive Communication 0.88 Expressive Communication 0.88 Fine motor 0.82 Gross motor 0.86	0.80 – 0.83

2.15.4.1 Reliability of the ASQ

For the ASQ, test-retest (n=175) and interrater reliability (n=112) were based on parent-completed questionnaires (Squires et al, 1997). To evaluate test-retest reliability, scores on two identical questionnaires completed by parents at two-week intervals were compared. Based on the percentage of agreement between the outcomes of the two completed questionnaires, test-retest reliability was reported to be 94% (standard error of the mean [SEM]=0.10). Interrater reliability was determined by evaluating the percentage of agreement between the classification of children as assessed by parent-completed questionnaires and by professional examiners. Inter-rater reliability was found to be 94% (SEM=0.12). The percent of agreement will often overestimate true reliability because it fails to account for chance agreement. The Kappa statistic would have been preferable for calculating the test-retest and interrater reliability of the ASQ (Portney and Watkins, 2000).

2.15.4.2 Reliability of the AIMS

Interrater reliability was assessed through simultaneous administration by two different examiners where one examiner actively scored the test while the other observed, and scored the test independently. The scores obtained for the 221 infants included in this correlation analysis resulted in an interrater reliability coefficient of > 0.996 (Piper et al, 1992). Test-retest reliability was analysed by administering the AIMS to 253 infants on two occasions, within seven days of each other. The test-retest reliability across all ages on the AIMS ranged from 0.86 to 0.99 but the authors do not report the statistic used to calculate the reliabilities (Coster, 1995).

2.15.4.3 Reliability of the Denver-II

For the Denver-II, test-retest reliability at seven to 10 days was found to show excellent agreement for 59% of the Denver-II items ($k \geq 0.75$), whereas 23% of items demonstrated fair to good agreement ($k \geq 0.40$). Interrater reliability was (≥ 0.75) (Frankenberg et al, 1992). Interrater relibaility was assesed using 17 examiners, who each assessed five children. One hundred and forty one items had 100% agreement, seven items had 90-99% and one item had 83%.

2.15.4.4. Reliability of the HINT

Interrater reliability of the HINT was examined by five paediatric physical or

occupational therapists for 28 infants at high risk (Harris and Daniels, 2001). One therapist served as the primary examiner and administered the test, while a second therapist observed and scored the test independently. For the total HINT score, the interrater reliability intra-class correlation coefficient (ICC) was 0.99. To evaluate test-retest reliability, the primary examiner re-administered the HINT within nine days of the initial screening for 20 of the infants. The test-retest reliability ICC=0.98. Intrarater reliability was examined by videotaping the assessments of 20 infants, scoring the videotapes, and then rescoring them at least one month later. For the five therapists, the intrarater reliability and intra-class correlation coefficients (ICCs) ranged from 0.98 to 0.99 (Harris and Daniels, 2001).

2.15.4.5 Reliability of the BINS

For the BINS, test-retest reliability was conducted by correlating BINS scores for 150 infants who were tested twice, the time interval being one week apart (Aylward, 1995). The reliability coefficient scores ranged from 0.71-0.84 depending on the child's age. Interrater reliability was conducted on 90 infants, and the number of examiners ranged from 15 to 18. For each infant, one examiner assessed the child, while the other observed and scored at the same time. Reliability coefficient results ranged from 0.79-0.96 (Aylward, 1995).

2.15.4.6 Reliability of the Bayley-III Screening Test

Internal consistency of the items was tested resulting in the following scores: Cognitive=0.85, Receptive Communication=0.88, Expressive Communication=0.88, Fine Motor=0.82 and Gross Motor=0.86, which reflects a high degree of internal consistency in the items (Bayley, 2006c). Test-retest reliability was tested using 203 children who were tested twice (range of 2-30 days between tests), with stability coefficients of between 0.80-0.83 revealing that the Bayley-III Screening Test provides consistent measurements (Bayley, 2006c).

It is difficult to compare the reliability results of the screening tools, as different methods and statistical calculations have been used, Only the HINT meets the recommended reliability requirements for screening tests (Rothstein et al, 1991), as none of the other tools reported on intrarater reliability.

As can be seen from the above review of the screening tests, none of the tests reviewed satisfies the criteria of having a representative standardisation sample as well as acceptable levels of reliability, validity, sensitivity, specificity, and positive predictive values (Harris et al, 2005). The AIMS is the strongest with regards to reliability, concurrent validity, and predictive validity, but is limited in comparison with the other tests in that it assesses only gross motor development. Although the AIMS' standardisation sample is large (n=2202) and was randomly selected and stratified by age and gender, concerns exist about the failure to report ethnic and socioeconomic variables (Harris et al, 2005).

The Denver-II has questionable accuracy in concurrently identifying typical children when compared with other standardised developmental assessments. In spite of having been published more than a decade ago, the predictive validity of the Denver-II has yet to be evaluated. Furthermore, the normative sample for the Denver-II was limited to children from the state of Colorado. Similar to the Denver-II, no predictive validity studies have been conducted for the ASQ, although this was not the intended goal. The HINT has strong reliability as well as acceptable concurrent validity with the Bayley-II. The predictive relationship of the HINT to the Bayley Motor Scale was modest, albeit stronger than the early Bayley Motor Scale's relationship to the later Bayley Motor Scale. Compared with the other tests, the HINT's normative sample is the smallest but its age range is also the narrowest. The normative sample is, however, diverse and representative of a variety of ethnic groups and maternal education levels. The Bayley-III Screening Test's accuracy improves with the classification category, which is problematic, as the lowest scores are in the At-Risk group – therefore the sensitivity of this screening test is poor. In light of the fact that each of the screening tests reviewed has identified strengths and weaknesses, practicing clinicians should use stringent criteria in determining which test to use within their own setting (Harris et al, 2005).

2.16 Test Manual

The following information should be outlined in the test manual:

The purpose of the screening test, as stated (Rothstein et al, 1991). The desired qualifications of the examiner or assessor should be outlned. The age range that the test covers should be stated. The time needed to administer and score the test should

be described, and is especially important where time constraints may apply. The developmental domains encompassed by the screening test should be outlined to ensure that the correct aspects of development are being assessed (Rothstein et al, 1991). The comparability between the standardisation sample used to determine normal values of the test (eg, ethnicity, gender, demographic characteristics) and those of the infants or children being screened should be discussed (Rothstein et al, 1991). Lastly, the traditional psychometric properties of the test (eg, reliability, validity) and/or the clinical epidemiological characteristics (eg, sensitivity, specificity) should be outlined in order to determine whether the test has been well-developed and will ascertain what it was developed for (Harris et al, 2005).

2.17 Conclusions

Paediatric HIV remains a major health challenge in South Africa. HIV is known to cause neurological damage through a number of pathological mechanisms. This results in a well described encephalopathy which causes developmental complications early on in childhood, and affects all facets of development, most notably gross motor development. Children use motor skills to interact with their peers and to maintain levels of confidence and independence in daily activities. Cognitive growth and social maturation have both been associated with motor skill development (Heller, 1997). Evaluation of motor skills in children with HIV infection requires the establishment of a descriptive base of motor performance levels for this population (Smith et al, 2002). With the advent of antiretroviral therapy, many children are now living longer, but often with severe developemental delay or disability due to HIV encephalopathy.

In developed countries, HIV positive children are routinely assessed for developmental problems, and are timeously referred for intervention. In South Africa, due to under-staffing, and high patient case loads, routine developmental screening has not been incorporated into routine paediatric HIV care, despite being a recommendation in the National Strategic Management Plan for HIV/AIDS for 2007-2011 (NSP, 2007). This is due to a number of factors including a lack of awareness, a lack of trained staff, a lack of time, and most of all, a lack of suitable outcome measures within this context.

There is therefore a need for a suitable screening tool, which is able to identify those children at risk for developmental delay, who may then be referred on for comprehensive developmental assessment and treatment. The screening tool should posess the following properties recommended by Rothstein et al (1991): the tool should be culturally sensitive, reliable, valid, and have high levels of sensitivity and specificity. As stated in chapter one this study therefore undertakes to evaluate an existing screening tool to determine its suitability for use in the South African context, and should this not prove to be suitable, a new screening tool will be developed, in order to address the need for identification and referral of those children with developmental problems.

Chapter 3

MEASURING INSTRUMENTS

In this chapter the measuring instruments used in this study will be discussed. Issues of reliability and validity are highlighted.

3.1 Bayley Scales of Infant and Toddler Development III (BSID III)

The Bayley Scales of Infant Development (BSID) were first published in 1969 and have long been considered the gold standard in developmental assessments (Harris et al, 2005; Tieman et al, 2005). Three versions of the BSID exist. The Bayley Scales of Infant and Toddler Development have been found to be sensitive to developmental changes in the first two years of life of infants who are medically fragile (Niccols and Latchman, 2002), and these findings support the clinical validity of the scale, which can be applied to those with HIV.

Like its predecessors (Bayley Scales of Infant Development I and II), the Bayley Scales of Infant Development III (BSID III) is an individually administered instrument which assesses developmental functioning of infants and young children aged between one and forty-two months. The primary purpose of the BSID III is to identify children with developmental delay and to provide information for intervention planning (Bayley, 2006a). The BSID III is a revision of the BSID II, and the revision was undertaken to improve the quality and enhance the utility of the instrument. The separate areas of Expressive and Receptive Communication, and Fine and Gross Motor function are able to be analysed individually in the BSID III. The BSID III was standardised on a national sample of 1,700 children in the United States of America who were stratified by age, sex, race/ethnicity, parent education level and region (Bayley, 2006b). The BSID III is reliable and valid, and is predictive of later developmental outcome (Bayley, 2006b). Internal consistency coefficients range from r=0.86 (Fine Motor) to r=0.91 (Cognitive, Expressive Communication and Gross Motor), to Receptive Communication which is r=0.87.The test-retest reliability ranges between r=0.63-0.87. The interrater reliability coefficient is r=0.82. The concurrent validity between the BSID II and BSID III ranged between r=0.41-0.71 (Bayley, 2006b). The BSID III is not an appropriate tool to use in busy HIV clinics in South

Africa, as the administrator needs to be trained in the use of the tool and the assessment is lengthy.

The Bayley Scales of Infant Development II (Bayley N, 1993), has been extensively used in HIV/AIDS research worldwide, and has also been used in South Africa in order to be able to compare results with worldwide studies (Potterton et al, 2009a; Potterton et al, 2009b; Baillieu and Potterton, 2008; McGrath et al, 2006a; Aina and Morakinyo, 2005; Chase et al, 2000; Drotar et al, 1997; Gay et al, 1995). The BSID I was normed on over 700 South African children taken from both urban and rural areas, and was found to be suitable for use on South African infants (Richter et al, 1992; Richter and Griesel, 1988). The South African infants scored above the test norms up to 10 months of age on both the Mental and Motor scales, and from 10 months up to the middle of the second year the groups obtained very similar scores. The trend across the whole age range in the South African sample was for urban children to score higher than rural children on both the Mental and Motor scales, although this was not statistically significant across the age range (Richter and Griesel, 1988). Based on these findings, the authors recommended that the BSID be used without adaptation on South African children. A study was also done in Uganda, examining normative data for the BSID II on infants between eight weeks and 30 months (Aina and Morakinyo, 2005). They determined that the infants scored above the normal values in both the Mental and Motor scale of the BSID II, and furthermore, their scores in the motor scale were significantly higher, especially in the lower age groups of eight, 16 and 32 weeks when compared with data from the Western world (Aina and Morakinyo, 2005).

The BSID III has been tested on normal black African infants (0-18 months) in Johannesburg who come from similar socioeconomic backgrounds to this study population, and results showed that the infants did well, and were often above average, which indicates that the BSID III is suitable for use on this population (Rademeyer, 2010; Brown, 2009). The procedure for training and interrater reliability testing is outlined in chapter four.

3.2 Bayley-III Screening Test

The Bayley-III Screening Test was developed using items from the BSID III. It is designed to assess briefly the cognitive, language and motor functioning of infants and young children between one and forty-two months. The primary purpose is to determine whether a child shows competence in age-appropriate tasks, shows evidence of emerging age-appropriate skills, or shows evidence of being at-risk for developmental delay. It was standardised on a sample of 1,675 children in the United States of America and included representative proportions of infants according to selected demographic variables. It has been shown to be reliable and valid (Bayley, 2006c). Internal consistency ranges between r=0.82-0.88, and test-retest reliability ranges between r=0.73-0.89. Validity was examined through test content, and classification accuracy against the BSID III. Classification accuracy ranges between 41.82% to 92.11% (Bayley, 2006b) Due to the fact that this tool was developed directly from the BSID III, which has long been considered to be the 'gold standard' in infant developmental assessment (Harris et al, 2005; Tieman et al, 2005), and the BSID I and II have shown to be suitable for use on a South African population, it was concluded that the Bayley-III Screening Test would be the most appropriate screening test to evaluate for use in the screening of infants infected with HIV.

3.3 Household Economic and Social Status Index (HESSI)

The Household Economic and Social Status Index (HESSI) was used to collect demographic information in this study. The HESSI was developed and standardised in Soweto, South Africa (Barbarin and Khomo, 1997).

The HESSI has the following characteristics:

- Simplicity in administration; ease of calculation;
- Scaled to be sensitive to variations in economic status at the low end of the economic scale;
- Reliance on information which is accessible to the informant;
- High correlations with widely used methods for assessing economic well-being; and
- Dimensions which have cross-cultural, cross-regional and cross-national reliance (Barbarin and Khomo, 1997).

Barbarin and Khomo (1997) recognised that financial resources alone are not an accurate measure of poverty in developing countries, and therefore rather than relying on self-reported income, the scale focuses on hunger, housing, utility expenses, possession of durable consumer goods, and accumulation of assets, social status and education (Barbarin and Khomo, 1997). Although these indicators may lack the financial specificity for studies of economic development, they are useful for studies of individual behaviour and child development (Barbarin and Khomo, 1997). The HESSI has been used in longitudinal studies in Soweto in order to evaluate socioeconomic status on a similar population (Potterton et al, 2009a; Potterton et al, 2009b; Barbarin and Richter, 2001; Barbarin and Khomo, 1997).

3.4 Conclusion

Based on the literature reviewed it was determined that the BSID III was the most appropriate developmental assessment tool for use in this study. The Bayley-III Screening Test was developed directly from the BSID, and it was therefore determined that this would be the most suitable screening tool to evaluate HIV-infected infants. The HESSI was considered to be an appropriate tool for assessing socioeconomic status in this study population as it was developed in Soweto, South Africa and focuses on socioeconomic factors which may impact on a child's development.

Chapter 4

STUDY ONE: THE EVALUATION OF THE BAYLEY-III SCREENING TEST AGAINST THE BAYLEY SCALES OF INFANT AND TODDLER DEVELOPMENT III

Study one consisted of the testing of the Bayley Scales of Infant and Toddler Development III (BSID III), against the Bayley-III Screening Test, to evaluate whether the Bayley-III Screening Test was suitable for use on this study population. The methodology and results for this stage will be presented in this chapter.

The objectives of this study were to:

- Test the Bayley-III Screening Test against the Bayley Scales of Infant and Toddler Development III in order to determine its suitability for use in South Africa
- Analyse demographic, anthropometric, and immunologic data in order to describe the health and socioeconomic status of the sample

4.1 Location

This study was conducted at the Harriet Shezi Children's Clinic, which is a public sector paediatric HIV clinic at Chris Hani Baragwanath Hospital in Soweto, South Africa. This is an outpatient clinic, which provides a variety of services for HIV positive children. Children are managed by a multi-disciplinary health team including paediatricians, medical officers, primary health care sisters, professional nurses, lay counsellors, dieticians, psychologists and pharmacists. Children are referred to the Harriet Shezi Children's clinic as soon as they are diagnosed with HIV. They are followed up at 3-6 monthly intervals depending on how well they are, and may be seen more regularly if they are ill.

The doctors at the clinic see over 100 HIV positive patients per day, ranging in age from birth to 16 years. Harriet Shezi clinic is an example of a relatively well-staffed, well-resourced urban clinic, and children who attend this clinic have access to many more services than are available in most other clinics in South Africa. In outlying

clinics, due to a shortage of doctors, children may be seen by nurses, who may be in charge of administering antiretrovirals, as well as making complex clinical decisions. All clinics, regardless of services offered, are understaffed due to the high numbers of children who attend the clinics each day. Harriet Shezi clinic was chosen as the location for data collection, as it is representative of a paediatric clinic in South Africa in terms of patient characteristics and volumes.

4.2 Ethical Considerations

Ethical clearance was obtained from the Committee for Research on Human Subjects at the University of the Witwatersrand (Clearance certificate M070817) (See Appendix I).

Permission to conduct the study was obtained from the hospital Chief Executive Officer at Chris Hani Baragwanath Hospital, as well as the head of the Harriet Shezi Clinic (See Appendix II and III).

Informed, written consent for the child was obtained from his/her primary caregiver prior to him/her being enrolled in the study. The study was explained to the caregiver and a written information sheet was provided. Caregivers were given the opportunity to ask questions before giving their consent (See Appendix IV).

The results of the assessment were discussed with the primary caregiver, and where the child was found to be delayed and in need of therapy, appropriate referrals were made. Home advice was also given to each caregiver regarding appropriate activities to do with the child in order to facilitate development.

4.3 Study Design

The design for study one was a cross-sectional study.

4.3.1 Subjects

Consecutive subjects were recruited from the Harriet Shezi Children's Clinic at Chris Hani Baragwanath hospital, Soweto. The caregivers of all children who met the inclusion criteria were asked to participate in the study.

4.3.2 Sample Size

For study one, the sample size calculation was performed as follows: the expected proportion of successes was 0.8 and testing was at the 0.05 level of significance. A sample size of 112 had a 90% power to detect a kappa of 0.75 (H_1: kappa $\geq$ 0.75) when testing the null hypothesis that kappa was 0.4 (H_o: kappa = 0.4). The sample size when comparing the Bayley Scales of Infant Development III with the Bayley-III Screening Test was therefore 112 infants.

Inclusion criteria:

- HIV positive infants aged between six and 18 months

Exclusion criteria:

- Infants who had severe enough physical and/or sensory impairments, which would not allow a norm-referenced score to be obtained on the Bayley Scales of Infant and Toddler Development III as outlined in the manual (Bayley, 2006a)

4.4 Materials and Measurements

4.4.1 Bayley Scales of Infant and Toddler Development (3rd version)(BSID III)

The Bayley Scales of Infant and Toddler Development III (Bayley, 2006a) have been found to be suitable for use on infants who are medically fragile (Niccols and Latchman, 2002). The BSID III assesses Cognitive function, Expressive and Receptive Communication, and Fine and Gross Motor function (Bayley, 2006a).

4.4.2 Bayley-III Screening Test

The Bayley-III Screening Test items are a subset of items of the Bayley Scales of Infant and Toddler Development, Third Edition (Bayley, 2006a; Bayley, 2006c). The test is designed to briefly assess the cognitive, language and motor functioning of infants and young children between 1-42 months of age and to determine whether a child shows competence in age-appropriate tasks, shows evidence of emerging age-appropriate skills, or shows evidence of being at-risk for developmental delay (Bayley, 2006c).

4.4.3 Household Economic and Social Status Inventory (HESSI)

The HESSI was used to give an indication of socioeconomic status in this population. This tool was developed in Soweto, South Africa (Barbarin and Khomo, 1997).

4.4.4 Anthropometric Measurements

Height/length and weight were measured routinely by the clinic nurses prior to the children seeing a doctor. Head circumference was measured in centimeters by the researcher, using a tape measure. The tape measure was placed just above the eyebrows and ears, and around the occipital bulge at the back of the child's head. The measure was recorded to the nearest 0.1cm.

Children who could stand independently had their height measured using a wall mounted scale (Hottain Ltd, Crymych, Dyfed). The children stood against the measure, bare foot or in socks, with their heels against the wall, and had the horizontal plate lowered until it rested on top of their heads. Height was recorded in centimeters. Infants who could not stand had their length measured in supine on a "Morena Baby Measuring Mat". Their heads were placed against the vertical board and the measurement was taken at the heel. Length was recorded in centimeters. Children who could stand had their weight measured on a Standing Digital Scale (Seca). Infants were weighed on a Tanita 1584 Baby Scale. Weight was recorded in kilograms to two decimal places.

4.4.5 Immunologic and Virologic Data

The child's clinical file was examined and the most recent CD4 count and viral load were noted, as well as their antiretroviral regimen and how long they had been on antiretroviral therapy.

4.5 Procedure

4.5.1 Training and Reliability

The administration of the BSID III is not limited to specific professions, however, paediatric experience is recommended. The researcher and the two research assistants who participated in the study all have more than five years of paediatric clinical experience and are well versed in the administration of standardised assessments. The researcher and two research assistants went through the

assessment thoroughly, and each item was discussed. The BSID III was administered to two normal children in order to practice administration of the items. Prior to commencement of data collection, interrater reliability was conducted between the researcher and the two research assistants on both the Bayley Scales of Infant and Toddler Development III and the Bayley-III Screening Test. Ten children (just under ten percent of the sample size for the first stage of the study) were assessed by the researcher and the two research assistants (one administered the items and allocated scores, while the other two observed and allocated scores). The scores were compared, and results are depicted in Table 4.1 and Table 4.2 below:

Table 4.1 Agreement between Researcher and Research Assistants on the BSID III (n=10)

Domain	Agreement between researcher and research assistant 1	Agreement between researcher and research assistant 2
Cognitive	100 %	100 %
Expressive Communication	99 %	100 %
Receptive Communication	100 %	99 %
Fine Motor	100 %	100 %
Gross Motor	100 %	100 %

The agreement between the researcher and the research assistants on the BSID III was excellent. The main discrepancies occurred in the Language domains, and therefore the responses were examined, and the reasons as to the differing scores were discussed and consensus on scoring was reached.

Table 4.2 Agreement between Researcher and Research Assistants on the Bayley-III Screening Test (n=10)

Domain	Agreement between Researcher and Research Assistant 1	Agreement between Researcher and Research Assistant 2
Cognitive	100 %	100 %
Expressive Communication	100 %	99 %
Receptive Communication	99 %	98 %
Fine Motor	100 %	100 %
Gross Motor	100 %	100 %

Again, agreement between the researcher and research assistants on the Bayley-III Screening Test was excellent, and only small discrepancies in scores were encountered in the Language domain. The results and responses were discussed, and consensus was reached as to what the correct responses from the child should be. The agreement between the researchers was excellent, and therefore data collection could proceed following the reliability testing.

4.6 Data Collection

The files for each child attending the clinic that day were laid out as soon as the caregiver registered at reception. The files were screened to determine the ages of the infants who were registered for the clinic, and all infants between the ages of six and 18 months were selected. The caregivers of the children who met the inclusion criteria were approached, and the caregiver was given a letter of information regarding the study. Where the caregiver could not speak or understand English, a counsellor or nurse was asked to translate. Once the caregiver agreed to participate, he/she was asked to give written informed consent for the child and his/herself (see Appendix IV). The researchers waited for the children to be weighed and measured before taking the child and caregiver to a quiet room to be assessed. The caregiver was asked to complete the HESSI form (and again where English was not understood, a counsellor or nurse was asked to translate for the assessment). The child's height, weight, CD4 count and viral load were noted from the file, as well as the combination of antiretroviral drugs the child was on, and length of time that they had been on

treatment. The child's head circumference was measured by the researcher with a tape measure – this procedure is described in 4.4.4 above.

The child was assessed on the BSID III and the Bayley-III Screening Test according to the guidelines outlined in the manuals. Once the assessments had been scored, the results were given to the caregiver, and the child's developmental status was discussed, along with the impact of HIV on neurodevelopment. When the child obtained a low score, he/she was appropriately referred, and where home intervention was appropriate, activities and advice were given to the caregiver. One hundred and twelve (112) infants between the ages of six and 18 months were assessed, and the data were entered into Excel spreadsheets.

4.7 Statistical Analysis

Statistical analysis was performed in consultation with the Medical Research Council of South Africa using STATA (version 10). The agreement between the Bayley Scales of Infant and Toddler Development III and the Bayley-III Screening Test was analysed using a Kappa statistic for ordinal data.

The Kappa scores were interpreted as follows:
<0.4 = poor agreement
≥ 0.4-0.74 = moderate agreement
≥ 0.75 = excellent agreement (Landis and Koch, 1977)

Demographic, anthromometric, immunologic and virologic data collected during study one were descriptively analysed using means, standard deviations and percentages. Z-scores were used for the following: height-for-age, weight-for-age, body mass index (BMI), head circumference, and weight-for-height.

4.8 Results

The results for this stage of the study will be presented and briefly discussed. Demographic information, anthropometric data, immunological data and developmental scores will be presented to give an indication of the health and socioeconomic status of the sample.

4.8.1 Demographic Information

Demographic information was collected using the HESSI questionnaire. Means and standard deviations or percentages were used to summarise the data. The data are displayed in Table 4.3 below:

Table 4.3 Demographic Information of the Sample (n=112)

Age (months)	11.38 (±3.9)
Primary Caregiver	
Mother	91 (81.3%)
Grandmother	16 (14.3%)
Father	1 (0.8%)
Aunt	2 (1.8%)
Other	2 (1.8%)
Education of primary caregiver	
<grade 5	6 (5.3%)
grade 5 – 6	7 (6.4%)
grade 7 – 9	25 (22%)
grade 10 – 11	40 (35%)
grade 12	28 (25%)
tertiary	7 (6.3%)
Total number of people living in the house	4.8 (±2.1)
Total number of adults living in the house	3.1 (±1.3)
Total number of children living in the house	2.1 (±1.2)
Means of income	
1 adult bringing in income	51 (45%)
2 adults bringing in income	6 (5.3%)
Grant dependent families	27 (24.7%)
No income	28 (25%)
Housing	
Shack	51 (45%)
Room	9 (8%)
Flat	6 (5.3%)
Home shared with other family	32 (29.2%)
Own home	14 (12.5%)

The mean age of the sample was 11.38 (±3.9) months. Eighty-one percent of primary caregivers were mothers, with grandmothers taking on this responsibility where mothers were no longer alive. Most primary caregivers had not completed school, and

therefore maternal education levels were low. When taking the average number of people per household (4.8 people) together with the fact that 45% of the families lived in a shack, and in 45% of families there is only one adult bringing in an income, the socioeconomic status of the sample was determined to be low.

4.8.2 Anthropometric Data

The anthropometric data are presented below, and are represented by means and standard deviations of z-scores which were calculated using height, weight, head circumference, age and gender. The results are shown in Table 4.4 below:

Table 4.4 Anthropometric Data of the Sample (n=112)

	6 – 8 m (n=38)	9–11 m (n = 23)	12 – 14 m (n = 21)	15 – 18 m (n = 30)	Total Group (n= 112)
Head circumference	-3.36(± 2.25)	-2.66 (±1.4)	-2.99 (±3.0)	-2.84 (±2.1)	-3.01 (±2.2)
Height-for-age	-2.26 (±1.5)	-1.60 (±1.6)	-1.82 (± 1.3)	-2.87(± 2.0)	-2.21 (±1.7)
Weight-for-age	-2.29 (±1.8)	-1.46 (±1.6)	-1.36 (±1.6)	-1.94 (±2.0)	-1.85 (±1.8)
Weight-for-height	-1.10 (±1.99)	-0.75 (±1.5)	-0.64 (±1.6)	-0.74(±1.9)	-0.85 (±1.7)
BMI	-1.39 (±2.09)	-0.73 (±1.4)	-0.41 (±1.6)	-0.35(±1.9)	-0.79 (±1.8)

In 2006, the World Health Organisation released two sets of child growth standards to replace the National Centre for Health Statistics References as a tool for growth monitoring, nutrition screening, and surveillance, as a clinical tool to assist with the diagnosis of malnutrition, and as a tool for nutrition research (Van den Broek, 2009; WHO, 2006a). The summary statistics can be compared with the reference, which has an expected mean z-score of 0 and a SD of 1.0 for all normalised growth indices.

4.8.2.1 Malnutrition, Stunting and Wasting

Moderate malnutrition is defined as a weight-for-age between -3 and -2 z-scores below the median of the WHO child growth standards (WHO, 2006a), similarly, moderate wasting and stunting are defined as a weight-for-height and height-for-age respectively, between -3 and -2 z-scores (WHO, 2006a). Severe acute malnutrition (severe wasting and/or oedema) or severe stunting is defined as height-for-age less than -3 z-scores (WHO, 2006a).

All data used to determine whether there is malnutrition, wasting or stunting present in this sample are presented in Table 4.4 above. In terms of malnutrition in this sample, it can be seen that only the younger group (6-8 months) is affected, with a mean weight-for-age of -2.29, which is indicative of moderate malnutrition.

Stunting is a problem in this sample, and is most severe in the 6-8 month old age group, where there is moderate stunting (height-for-age =-2.26). The 15-18 month age group also shows moderate stunting (height-for-age =- 2.87).

4.8.2.2 Body Mass Index (BMI)

Body Mass Index (BMI) is a simple index of weight-for-height that is commonly used to classify underweight, overweight and obese children. It is defined as the weight in kilograms divided by the square of the height in metres (kg/m^2) (WHO, 2006a). The only group with a low body mass index is the 6-8 month age group, again, due to their low weight measurements.

4.8.2.3 Head Circumference

The children in this sample had head circumference z-scores of between -2 and -3, indicating that they are below the WHO average, when taking the cut-off score of less than -2. Head circumference is known to correlate closely with brain volume and can therefore be used to measure brain growth (Wickett et al, 2000). It has been found that postnatal brain growth is more important than prenatal brain growth in determining higher mental function at the age of nine years (Gale et al, 2004). This is of concern for these children, as they may be affected cognitively when reaching school age.

4.8.3 Immunological and Antiretroviral Status

The immunological status of the infants is consistently monitored at the clinic. Baseline laboratory tests including CD4 cell count, CD4 percentage (percentage of total lymphocyte count) and full blood count are performed for each child. Children not yet meeting the criteria for HAART initiation are followed up at varying intervals which are dependent on age, immune status and clinical disease category, until they meet the criteria for HAART. Prior to HAART initiation, a viral load test is performed and lay counsellors provide counselling and information on HAART and adherence. Children undergo a clinical examination at four and 12 weeks after HAART initiation, and are assessed at 12 weekly intervals thereafter. CD4 count, CD4 percentage, and viral load tests are performed 24 weeks after starting HAART, and are repeated every 24 weeks thereafter.

CD4 percentage and viral load are used by the clinic to monitor immune status and adherance or response to HAART. According to the South African Department of Health guidelines (DOH, 2008), treatment failure is defined as the following:

- no improvement in CD4 percentage after 24 weeks of HAART
- return of CD4 percentage (repeated after one month of treatment) to baseline or below, in the absence of concurrent illness
- more than 50% decline in CD4 percentage from peak (confirmed within one month) in the absence of concurrent illness
- rebound of viral load to baseline

Immunological data including CD4 percentage and viral load were collected for each child in this study in order to assess their current immunological status. The latest available values were recorded from the file. The immunological and virological status of the sample is shown in Table 4.5 below:

Table 4.5 Immunologic and Antiretroviral Status of the Sample (n=112)

Age	n	Number of children on HAART	Average length of time on HAART	CD4%	Viral Load
6–8 m	38	24 (63.2%)	6.3 weeks (±6.4)	19.6%(±9.6)	96262 (±26)
9–11 m	23	18 (78.3%)	16.8 weeks (±11.4)	23.0% (±12)	38037 (±75.2)
12–14 m	21	13 (61.9%)	15 weeks (± 17.2)	22.2%(±9.6)	11124(±116.7)
15–18 m	30	22 (73.3%)	15.9 weeks (±15.4)	28.0%(±0.5)	5568.6(±78.7)
Total	112	77 (68.8%)	18 weeks (±2.8)	23.0% (±9.9)	21823.2 (±3.5)

The immunological status of this sample was rather varied, and no real conclusions could be drawn from it. The number of children overall who are receiving HAART is not optimal (68.8%). The mean viral load drops with increasing age, and the mean CD4 percentage increases with increasing age.

4.8.5 BSID III Composite Scores

The mean composite scores achieved by the sample are presented below. These scores give an indication of where the infant's development lies relative to the average score (Bayley, 2006b). Composite scores have a range of 40 to 160, a mean of 100 (which is the average), and a standard deviation of 15 (Bayley, 2006b) (See 'Definition of Terms' on page xviii for interpretation of the scores). The composite scores are presented as means with the standard deviation for each age group and for the total group in Table 4.6 below:

Table 4.6 BSID II Composite Scores and Interpretation (n = 112)

Age	n	Cognitive	Language	Motor
6 – 8 months	38	85 (± 20.85)	95 (± 18.68)	76 (± 23.95)
9 – 11 months	23	86 (± 14.20)	93 (± 15.47)	74 (± 20.59)
12 – 14 months	21	75 (± 17.41)	81 (± 15.80)	71 (± 8.72)
15 – 18 months	30	74 (± 15.22)	79 (± 14.19)	68 (± 15.59)
Total Group	112	80 (± 18.04)	88 (± 17.68)	72(± 20.32)

The composite scores give an indication as to the developmental status of the sample. Language function is the area which is least affected in the sample, although the only scores which fell into the Average range in the Language domain were those in the two younger age groups. Cognitive function is less affected than motor function in this sample, although the scores are low and fall into the 'Low Average' and 'Borderline' categories. Motor function is the area of development which is most consistently severely affected, and falls into 'Borderline' or 'Extremely Low' categories. In all domains, the scores worsen as the children get older. The percentage of infants in each classification category for Cognitive, Language and Motor function is shown in Table 4.7 below:

Table 4.7 Percentage of Infants in each Classification Category for Cognitive, Language and Motor Development

Composite	Classification	Cognitive	Language	Motor
130 and above	Very Superior	0.9%	0%	0%
120 – 129	Superior	0.%	1.8%	0%
110 – 119	High Average	3.6%	8%	2.7%
90 – 109	Average	31.3%	41.1%	19.7%
80 – 89	Low Average	17.9%	17.9%	17%
70 – 79	Borderline	17%	16.1%	18.8%
69 and below	Extremely Low	28.6%	15.2%	42%

In each area of development, there is a large percentage of infants who scored in the Extremely Low range (69 and below). This is most evident in the Motor facet, where

41.96% of infants obtained scores in the Extremely Low range. Within the Cognitive facet, infants most commonly obtained scores in the Average range, followed by the Extremely Low range. In the Language domain, infants most commonly obtained scores in the Average range.

4.9 Agreement between BSID III and the Bayley-III Screening Test

The scaled score of the BSID III represents the child's performance on a subtest relative to his or her peers of the same age (Bayley, 2006b), and may be converted to the category obtained on the Bayley-III Screening Test (Bayley, 2006c) (See 'Definition of Terms' on page xix for conversion of the scores).

The data for each child were entered into Excel, and the results for the BSID III were converted to the same format as the Bayley-III Screening Test (At-Risk, Emerging, and Competent) for comparison. In order to enter the data in Excel, each category was assigned values as follows: the category 'At-Risk' was assigned a value of one, 'Emerging' was assigned a value of two, and 'Competent' was assigned a value of three. Agreement between the results was tested using a Kappa statistic for ordinal data. The results are presented in Table 4.8 below:

Table 4.8 Agreement between the Bayley Scales of Infant and Toddler Development III and the Bayley-III Screening Test (n=112)

Age group	n	Cognitive	Expressive Communication	Receptive Communication	Fine motor	Gross motor
6–8 months	38	0.41	**0.75**	0.55	0.28	0.61
9–11 months	23	0.51	0.48	**0.86**	0.50	0.43
12–14 months	21	0.69	**0.79**	**0.78**	0.51	**0.92**
15–18 months	30	0.70	**0.89**	**0.87**	0.49	0.41
Total group	112	0.58	**0.82**	**0.76**	0.44	0.57

The above scores were obtained when each child was assessed on the BSID III, and the Bayley-III Screening Test, and were analysed as follows: the BSID III scaled score was used to compare the result with the Bayley-III Screening Test. The figures in bold represent the Kappa scores which denote excellent agreement. The agreement in each area of development is presented below:

4.9.1 Cognitive Development

The agreement between the Bayley-III Screening Test and the BSID III in the Cognitive facet ranged between 0.41 and 0.70. None of the Kappa scores in the Cognitive facet were higher than 0.70, representing moderate agreement across the Cognitive facet. The overall agreement in the Cognitive facet was 0.58, which represents moderate agreement (≥0.4 -<0.75).

4.9.2 Language Development

The agreement between the Bayley-III Screening Test and the BSID III was better in the Language facet. The Expressive and Receptive Communication domains had excellent agreement in all age groups apart from the 9-11 month group for Expressive Communication, and the 6-8 month age group for Receptive Communication. The overall agreement between the tools in the Expressive Communication facet was 0.82, and in the Receptive Communication facet was 0.76, representing excellent agreement (≥ 0.75).

4.9.3 Motor Development

The agreement between the Bayley-III Screening Test and the BSID III in the Motor facet ranged between 0.28 and 0.92. The agreement in the Fine Motor facet was poor to moderate for all age groups, ranging between 0.28 and 0.50. In the Gross Motor facet the agreement ranged between 0.41 and 0.92, with most age groups demonstrating moderate agreement (≥0.4 -<0.75). The overall agreement in the Fine Motor facet was 0.44, and in the Gross Motor facet, was 0.57 which represents moderate agreement in both (≥0.4 -<0.75).

4.9.4 Agreement Compared with Bayley-III Screening Test Standardisation Sample

The classification accuracy of the Bayley-III Screening Test with the BSID III is outlined in the Bayley-III Screening Test technical manual (Bayley, 2006c). In order to

establish whether the present study had similar trends in classification accuracy, the data were analysed in the same way as is outlined in the technical manual. The percentage of children correctly identified by the Bayley-III Screening Test was calculated, and could then be compared with the Bayley-III Screening Test standardisation sample. This was done for each possible category of the test namely At-Risk, Emerging, and Competent. The classification accuracy of the Bayley-III Screening Test with the BSID III for At-Risk cases is shown in Table 4.9 below:

Table 4.9 Classification Accuracy of the Bayley III Screening Test with the BSID III for At-Risk Cases

	No of Cases on BSID III	No of Cases on Bayley-III Screening Test	Percentage Identified by Screening Test	Screener Classification of Outlying Cases
Cognitive	36	28	77.8%	5=Emerging (13.9%) 3=Competent (8.3%)
Receptive Communication	17	14	82.4%	3=Emerging (17.6%)
Expressive Communication	18	15	83.3%	2=Emerging (11.1%) 1=Competent (5.6%)
Fine Motor	34	20	58.8%	14=Emerging(41.2%)
Gross Motor	56	35	62.5%	18=Emerging(32.1%) 3=Competent (5.4%)

Table 4.9 shows the classification accuracy for the children identified by the BSID III as being At-Risk. In the technical manual (Bayley, 2006c), it is stated that classification accuracy was “moderate” for children with scaled scores of 1-4. The range for the standardisation sample was 41.82% for Fine Motor function to 65.9% for Receptive Communication. None of the children in the standardisation sample were identified as Competent. In the present study, the range is 58.8% for Fine Motor function to 83.3% for Expressive Communication. The range is wider, and Expressive Communication had better accuracy than Receptive Communication, but only by one percent. There were a number of cases where children were identified as Competent,

which did not occur in the standardisation sample (Bayley, 2006c). The range in the present study for children misidentified as Competent was 5.4% for Gross Motor function to 8.33% for Cognitive Function. The classification accuracy of the Bayley-III Screening Test with the BSID III for Emerging cases is shown in Table 4.10 below:

Table 4.10 Classification Accuracy of the Bayley III Screening Test with the BSID III for Emerging Cases

	No of Cases on BSID III	No of Cases on Screening Test	Percentage Identified by Screening Test	Screener Classification of Outlying Cases
Cognitive	32	19	59.4%	9=Competent (28.1%) 4=At-Risk (12.5%)
Receptive Communication	33	31	93.9%	2=Competent (6.1%)
Expressive Communication	19	13	68.4%	5=Competent (26.3%) 1=At-Risk (5.3%)
Fine Motor	30	17	56.7%	8=Competent (26.7%) 5=At-Risk (16.7%)
Gross Motor	28	20	71.4%	8=Competent (28.6%)

Table 4.10 shows the classification accuracy for the chidlren identified by the BSID III as having Emerging skills. In the technical manual (Bayley, 2006c), it is stated that classification accuracy was “accurate” for children with scaled scores of 5-7. The range for the standardisation sample was 63.9% for Cognitive function to 77.8% for Receptive Communication (Bayley, 2006b). The percentage of children misidentified as being At-Risk was very low (0.8%-5.2%) (Bayley, 2006c). In this study, the range is 56.7% for Fine Motor function to 93.9% for Receptive Communication. The range is wider, and Fine Motor function had the lowest percentage of classification accuracy, but this was only two percent lower than Cognitive function. There were a number of cases where children were identified as At-Risk, which did not occur in the

standardisation sample (Bayley, 2006c). The range in this study for children misidentified as At-Risk was 5.3-16.7% which is much higher than the standardisation sample. Conversely, the range for children misidentified as Competent in the present study was 6.1% to 28.6% which indicates that the problems with which these children present with may be missed. The classification accuracy of the Bayley-III Screening Test with the BSID III for Competent cases is shown in Table 4.11 below:

Table 4.11 Classification Accuracy of the Bayley III Screening Test with the BSID III for Competent Cases

	No of Cases on BSID III	No of Cases on Screening Test	Percentage Identified by Screening Test	Screener Classification of Outlying Cases
Cognitive	44	34	77.3%	9=Emerging (20.5%) 1=At-Risk (2.3%)
Receptive Communication	62	55	88.7%	7=Emerging (11.3%)
Expressive Communication	76	72	94.7%	4=Emerging (5.3%)
Fine Motor	48	33	68.8%	15=Emerging (31.3%)
Gross Motor	28	25	89.3%	3=Emerging (10.7%)

Table 4.11 shows the classification accuracy for the children identified by the BSID III as having Competent skills. In the technical manual (Bayley, 2006c), it is stated that classification accuracy was “very accurate” for children with scaled scores of 8-19. The range for the standardisation sample was 84% for Cognitive function to 92.1% for Receptive Communication (Bayley, 2006c). None of the children in the standardisation sample were misidentified as being At-Risk (Bayley, 2006). In the present study, the range is 68.8% for Fine Motor function to 94.7% for Expressive Communication. In this case, Expressive Communication was higher than Receptive Communication, and Fine Motor function had the lowest accuracy. There were a few cases (2.3%) where children were identified as At-Risk, which did not occur in the standardisation sample

(Bayley, 2006c). Similarly, the range for children misidentified as having Emerging skills was 5.3-31.3%.

4.9.4 Conclusion

As can be seen from the above results, the Kappa scores are poor to moderate for most domains, apart from the Language facets, and therefore the Bayley-III Screening Test could not be considered optimal for clinical use, as Kappa scores of ≥0.75 are preferred for clinical suitability (Landis and Koch, 1977). The scores with the highest level of agreement are those in the Expressive and Receptive Communication domains, but since Motor and Cognitive function are more severely affected in HIV-infected children, the screening tool needs to be sensitive enough to detect these problems, which it did not.

The Bayley-III Screening Test was therefore unsuitable for use in this context, as it did not sufficiently detect the main problems that occur in these children. A new screening tool therefore needed to be developed, which would be more suitable for use in this context. The development of a new test is therefore reported on in chapter five.

Chapter 5

STUDY TWO: DEVELOPMENT OF A NEW SCREENING TOOL

5.1 Development of a New Screening Tool

The Bayley Scales of Infant and Toddler Development III and the Bayley-III Screening Test did not correlate sufficiently for clinical use (Kappa <0.75), and therefore a new screening tool had to be developed. The objectives for study two were as follows:

- To identify items which would be suitable for inclusion in a new screening tool
- To categorise items into age group suitability
- To develop a scoring system for each age group
- To develop a new screening tool based on the above three objectives

5.2 Initial Identification of Developmental Facets to be Included in a New Screening Tool

Developmental scores from the original sample were analysed to determine which facets of development were most severely affected in this group, and therefore which facets of development should be included in a screening tool. Motor development was consistently the lowest score obtained in all age groups (see Table 4.5). The scaled scores obtained on the BSID III were also examined in order to distinguish between Gross and Fine Motor function, and Expressive and Receptive Communication function, which the composite scores do not. The scaled scores (means and standard deviations) are presented in Table 5.1 below:

Table 5.1 Scaled Scores Obtained by the Sample on the BSID III (n=112)

Age	n	Cognitive	Expressive Communication	Receptive Communication	Fine Motor	Gross Motor
6–8m	38	7.07 (±4.2)	8.55 (±3.2)	9.81(±3.7)	7.26 (±4.6)	5.15(±3.7)
9–11m	23	7.00(±2.7)	9.17(±3.2)	8.82(±2.4)	7.39(±2.6)	5.65(±4.2)
12–14m	21	5.14(±3.4)	7.09 (±3.8)	6.47(±2.1)	6.14 (±3.2)	4.28 (±3.7)
15–18m	30	4.96 (±3.0)	7.36 (±2.9)	5.70 (±2.3)	5.20 (±2.6)	4.33 (±2.9)
Total	112	6.13 (±3.6)	7.88 (±3.4)	8.08 (±3.3)	6.52 (±3.6)	4.88(±3.6)

Gross Motor function was consistently found to be the developmental facet that was most severely affected, followed by Cognitive function. In order to evaluate cognitive function, a large range of equipment is required, which would be costly to produce, and would require more time during evaluation, as more subtle responses are required. A screening tool needs to fulfil the following criteria in a busy clinic in a developing country:

a) It should identify those children who are experiencing developmental problems, so that they may be referred for further assessment
b) It should be sensitive enough to distinguish between children who are experiencing problems and those who are not
c) It should be quick and easy to administer due to time constraints
d) It should not require profession specific training, so that counsellors and nurses may also administer it, should the need arise
e) It should be inexpensive enough to reproduce, and should not require any special equipment

After careful examination of the above data, it was evident that a gross motor screening tool should be developed. Gross motor function is relatively quick and easy to assess, as no equipment is needed, which is a major consideration for clinics in South Africa and other developing countries, as time is limited, and budgets cannot cater for expensive assessments requiring equipment. The identification of gross motor function as being the most severely affected facet of development is in line with results of studies done globally as well as in South Africa (Potterton et al, 2009a; Potterton et al, 2009b; Ferguson and Jelsma, 2009; Baillieu and Potterton, 2008; Drotar, 1997; Chase et al, 1995; Nozyce et al, 1994).

5.3 Process of Development

The process of development of the new screening tool is outlined in Figure 5.1 below:

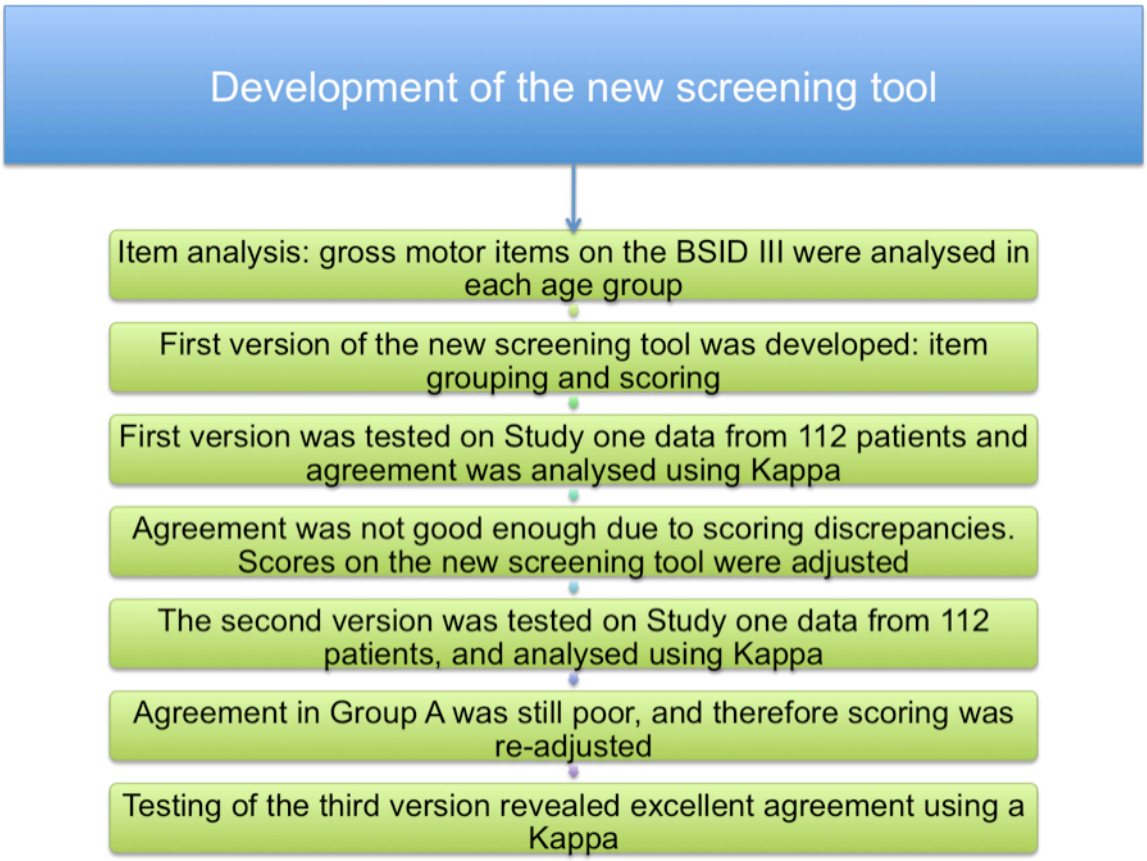


Figure 5.1 Process of development of the new screening tool

5.3.1 Item Analysis

Data from the BSID III Gross Motor scale were entered into an Excel spreadsheet. The scores for each item on the BSID III were entered as follows:

1 = credit

0 = no credit

2 = finished the facet (after obtaining five zero's in a row)

The data were entered according to age group and were grouped according to the score obtained on the BSID III namely 'At-Risk', 'Emerging', and 'Competent' (See Appendix V for example).

Percentages obtained by the group for each item were obtained for each score group (At-Risk, Emerging and Competent), and six items were then selected for each age group according to the percentages obtained. This was done as follows and is illustrated in Table 5.2:

Table 5.2 Item Selection

Item	Percentage Obtained in the At-Risk Group	Percentage Obtained in the Emerging Group	Percentage Obtained in the Competent Group
1	<20	100	100
2	<20	100	100
3	0	≤50	100
4	0	≤50	100
5	0	0	As close to 100% as possible
6	0	0	As close to 100% as possible

Two items which discriminated the At-risk, from Emerging and Competent groups (less than 20% in the At-Risk group, and 100% in the other groups) were selected. Then, items which discriminated between Emerging and Competent groups were selected (0-5% of children who were At-Risk obtained credit, 30-50% of the Emerging group obtained credit, and 100% of the Competent group obtained credit). Lastly, two items were selected which discriminated the Competent group from the other two groups (100% or as high as possible in the Competent group, and 0% in the other groups).

The selected items were grouped according to the age groups for study one (6-8 months, 9-11 months, 12-14 months, 15-18 months). Scoring was determined as follows: the oldest children in the group should obtain credits for all items to be considered Competent; and one less credit was needed by each subsequent age group to be considered Competent. This was done following a literature search to determine which activities were appropriate per age group, and therefore which items a child should get credit for to be placed in the Competent or At-Risk categories. The Emerging category was assigned a score in the middle of the other two. Table 5.3 illustrates this with an example from the 6-8 month age-group:

Table 5.3 Scoring

Age	Score Obtained to be placed in the Competent category	Score Obtained to be placed in the Emerging category	Score Obtained to be placed in the At-Risk category
6 months	4	3	≤2
7 months	5	4	≤3
8 months	6	5	≤4

The initial outcomes of the new screening tool were similar to those used in the Bayley-III Screening Test, namely At-Risk, Emerging and Competent, so that comparisons between the tools could be made. The first version of the new screening tool was then tested against the data collected in study one using the motor section of the assessment form for the 112 infants. Only the Gross Motor facet of the BSID III data was used, as the new screening tool was a gross motor assessment tool only, and the new screening tool was scored against all 112 infants. The results are presented in chapter six.

Chapter 6

STUDY THREE: TESTING OF THE NEW ASSESSMENT TOOL

The objective for study three was as follows:

- Testing and refinement of the first version of the screening tool until it satisfied statistical criteria (K≥0.75).

6.1 Testing of the First Version of the New Assessment Tool

Once the items had been assigned to age groups, and the scoring had been formulated, the first version was ready for testing against data collected for study one, using the assessment forms from the BSID III. This was done using the Gross Motor scale only, and the results from the new screening tool were compared with the scaled score obtained from the BSID III, which had been converted into categories (At-Risk, Emerging and Competent). One hundred and twelve BSID III assessments were scored against the new screening tool in the following age groups, which were stratified as outlined in Table 6.1 below:

Table 6.1 Age Groups and Stratification for Testing of First Version of the New Screening Tool (n = 112)

Age	n
6 – 8 months	38
9 – 11 months	23
12 – 14 months	21
15 – 18 months	30

6.1.1 Statistical Analysis

The data were entered into an Excel spreadsheet, and the statistical analysis was done in consultation with a biostatistician from the Medical Research Council of South Africa, using Stata (version 10). Correlations between the new screening test and the items on the BSID III were carried out using a Kappa statistic for ordinal data. The results are presented in Table 6.2 below:

Table 6.2 Agreement between the BSID III and the First Version of the New Assessment Tool (n=112)

Age Group	Kappa	Percentage of Children whose Scores Correlated on both Tools
6 – 8 months	0.34	63.2%
9 – 11 months	0.63	78.3%
12 – 14 months	0.33	61.9%
15 – 18 months	**0.75**	86.7%
Total group	0.51	72.3%

Agreement was poor in all age groups aside from the 15-18 month group where K=0.75 indicates excellent agreement (Landis and Koch, 1977). Therefore the data from age-groups 6-8 months, 9-11 months and 12-14 months needed to be analysed in order to determine where the scores did not agree and why this had occurred.

6.1.2 Analysis of Results

Once the data had been analysed to determine where the discrepancies lay, it was determined that the screening tool was not sensitive enough in terms of scoring for the middle age category in age-groups 6-8 months, 9-11 months and 12-14 months, and therefore scores were adjusted as follows:

The two youngest age groups were combined for scoring in age-groups 6-8 months, 9-11 months and 12-14 months in order to demand less of the infants in the middle age category in that age group (e.g. for seven month old, a score of five was required to denote competence in the first version, and in the second version, only four was required). This was done in age groups 6-8 months, 9-11 months and 12-14 months only, as the Kappa for the 15-18 month age-group was 0.75, which indicates excellent agreement (Landis and Koch, 1977).

Following further data analysis, it was revealed that the last two items in the age group 6-8 months were unsuitable for this age category, as they were too difficult for even the Competent infants, and did not differentiate between ‘Competent’ infants and

those with ‘Emerging’ skills. The last two items in the age group 6-8 months were therefore replaced with two items which were sourced from the item analysis. These two new items were better suited to the age category 6-8 months, and ensured that those children who were Competent could achieve all items. This is shown in Figure 6.1 below; Omitted items are deleted, and the highlighted portions indicate where changes were made. The second version emerged through the changes outlined above, and was ready for testing against the BSID III.

FIRST VERSION	SECOND VERSION
6 - 8 months 1) controls head at 90 degrees while lying on stomach 2) Lifts chest while lying on stomach (with extended arms) 3) plays with feet 4) rolls from back to stomach ~~5) makes stepping movements~~ ~~6) crawls on stomach~~ **6 months** ≥4: Competent; 3: Emerging, ≤ 2: At-Risk **7 months** ≥5: Competent, 4: Emerging, ≤ 3: At-Risk **8 months** 6: Competent, 5: Emerging, ≤ 4: At-Risk	**6 - 8 months** 1) controls head at 90 degrees while lying on stomach 2) Lifts chest while lying on stomach (with extended arms) 3) plays with feet 4) rolls from back to stomach 5) sits alone 6) turns body while seated **6 +7 months** ≥4: Competent; 3: Emerging, ≤ 2: At-Risk **8 months** 6: Competent, 5: Emerging, ≤ 4: At-Risk
9 - 11 months 1) turns body while seated 2) makes stepping movements 3) moves from sitting to being on hands and knees 4) crawls at least 1 ½ metres on hands and knees 5) pulls up to standing position 6) walks sideways with support **9 months** ≥4:Competent; 3: Emerging, ≤ 2: At-Risk **10 months** ≥5:Competent, 4: Emerging, ≤ 3: At-Risk **11 months** 6:Competent, 5: Emerging, ≤ 4: At-Risk	**9 - 11 months** 1) turns body while seated 2) makes stepping movements 3) moves from sitting to being on hands and knees 4) crawls at least 1 ½ metres on hands and knees 5) pulls up to standing position 6) walks sideways with support **9 +10 months** ≥4: Competent; 3: Emerging, ≤ 2: At-Risk **11 months** 6: Competent, 5: Emerging, ≤ 4: At-Risk
12 – 14 months • pulls up to standing position • bounces while standing • sits down from standing in a controlled manner • stands independently • walks alone with coordination • squats without support **12 months** ≥4:Competent; 3:Emerging, ≤ 2:At-Risk **13 months** ≥ 5:Competent, 4:Emerging, ≤ 3:At-Risk **14 months** = 6:Competent, 5:Emerging, ≤ 4:At-Risk	**12 – 14 months** 1) pulls up to standing position 2) bounces while standing 3) sits down from standing in a controlled manner 4) stands independently 5) walks alone with coordination 6) squats without support **12 +13 months** ≥4:Competent; 3:Emerging, ≤ 2:At-Risk **14 months** 6:Competent, 5:Emerging, ≤ 4:At-Risk
15 – 18 months 1) sits down from standing in a controlled manner 2) stands independently 3) stands up with no assistance 4) walks alone with coordination (at least 5 steps) 5) squats without support 6) runs with coordination **15 months** ≥4:Competent; 3:Emerging, ≤2:At-Risk **16 months** ≥5:Competent, 4:Emerging, ≤3:At-Risk **17 – 18 months** 6:Competent, 5: Emerging, ≤ 4:At-Risk	**15 – 18 months** 1) sits down in a controlled manner 2) stands independently 3) stands up with no assistance 4) walks alone with coordination (at least 5 steps) 5) squats without support 6) runs with coordination **15 months** ≥4:Competent; 3: Emerging, ≤2: At-Risk **16 months** ≥5:Competent, 4: Emerging, ≤3: At-Risk **17 – 18 months** 6: Competent, 5: Emerging, ≤ 4: At-Risk

Figure 6.1 Development of the second version of the new screening tool

6.2 Testing of the Second Version of the New Screening Tool

Once changes had been made to the first version, the second version was developed, and is shown in Figure 6.1. Testing of the second version was conducted against previous data collected for study one, using the BSID III Gross Motor facet. The same procedure that was outlined in 6.1 above was followed.

6.2.1 Statistical Analysis

Statistical analysis was carried out as described in 6.1.1. The results are presented in Table 6.3 below:

Table 6.3 Agreement Between the BSID III Gross Motor Scaled Score and the Second Version of the New Screening Tool (n=112)

Age group	Kappa	Percentage of Children whose Scores Correlated on both Tools
6 – 8 months	0.56	73.7%
9 – 11 months	**0.85**	91.3%
12 – 14 months	**0.84**	90.5%
15 – 18 months	**0.88**	93.3%
Total group	**0.76**	85.7%

The agreement in the age group 6-8 months was still not sufficient. All other age groups had excellent agreement between the BSID III and the new screening tool ($k>0.75$) (Landis and Koch, 1977). Therefore further analysis was undertaken to determine why the agreement in the 6-8 month age group was low.

6.2.2 Analysis of Results

It was determined that the scoring and items for this age group were not appropriate, and therefore were analysed and adjusted. The credits obtained by each child on the BSID III and the new screening tool were analysed to determine whether the child obtained credits for the same items on both tools, and was placed in the same score category by each tool (i.e. At-Risk, Emerging, or Competent). Item number six ('turns body while seated') in the age group 6-8 months was too difficult, and therefore even if the child was Competent on the BSID III, he/she would not obtain a Competent score

on the screening tool, as he/she could not obtain six credits. Item six was removed, and the scores were adjusted as follows: an eight month old should obtain five credits in order to be classified as Competent, and a seven and six month old should obtain a score of four (illustrated in Figure 5.2 below).

The last item was removed in the second age category “walks sideways with support” (9-11 months), and age groups were combined to form the following groups: 6-8 months, 9-12 months, 13-16 months, and 17-18 months. This was done to avoid over-penalising the older children (e.g. where a 14 month old was expected to obtain six credits in order to be deemed Competent through being the oldest in the group, the child now fell into the middle of the group and was only expected to obtain four credits). This was also done to ensure that items fell into clinically relevant age groups, as on the BSID III, items are not allocated to specific age groups, and therefore some of the items may have been too difficult for the groups they were allocated to on the new screening tool (illustrated in Figure 6.2 below). The item ‘bounces while standing’ was removed, as most children tested in the 12-14 month age group had outgrown this activity. Once the above changes had been made, the third version was ready for testing against the BSID III. The third version is illustrated in Figure 6.2 below

SECOND VERSION	THIRD VERSION
6 - 8 months 1) controls head at 90 degrees while lying on stomach 2) lifts chest while lying on stomach (with extended arms) 3) plays with feet 4) rolls from back to stomach 5) sits alone ~~6) turns body while seated~~ **6 +7 months** ≥4: Competent; 3: Emerging, ≤2: At-Risk **8 months** 6: Competent, 5: Emerging, ≤ 4: At-Risk	**6 - 8 months** 1) controls head at 90 degrees while lying on stomach 2) lifts chest while lying on stomach (with extended arms) 3) plays with feet 4) rolls from back to stomach 5) sits alone **6 +7 months** ≥ 4: Competent; 3: Emerging, ≤ 2: At-Risk **8 months** 5: Competent, 4: Emerging, ≤ 3: At-Risk
9 - 11 months 1) turns body while seated 2) makes stepping movements 3) moves from sitting to being on hands and knees 4) crawls at least 1 ½ metres on hands and knees 5) pulls up to standing position ~~6) walks sideways with support~~ **9 +10 months** ≥ 4: Competent; 3: Emerging, ≤ 2: At-Risk **11 months** 6: Competent, 5: Emerging, ≤ 4: At-Risk	**9 - 12 months** 1) turns body while seated 2) makes stepping movements 3) moves from sitting to being on hands and knees 4) crawls at least 1 ½ metres on hands and knees 5) pulls up to standing position **9 +10 months** ≥ 4: Competent; 3: Emerging, ≤ 2: At-Risk **11 +12 months** 5: Competent, 4: Emerging, ≤ 3: At-Risk
12 – 14 months 1) pulls up to standing position 2) bounces while standing 3) sits down from standing in a controlled manner 4) stands independently 5) walks alone with coordination 6) squats without support **12 +13 months** ≥4: Competent; 3: Emerging, ≤ 2: At-Risk **14 months** 6: Competent, 5: Emerging, ≤ 4: At-Risk	**13 – 16 months** 1) pulls up to standing position ~~2) bounces while standing~~ 3) sits down from standing in a controlled manner 4) stands independently 5) walks alone with coordination 6) squats without support **13 months** ≥3: Competent; 2: Emerging, ≤ 1: At-Risk **14 months** ≥4: Competent; 3: Emerging, ≤ 2: At-Risk **15 + 16 months** 5: Competent, 4: Emerging, ≤ 3: At-Risk
15 – 18 months 1) sits down from standing in a controlled manner 2) stands independently 3) stands up alone 4) walks alone with coordination (at least 5 steps) 5) squats without support 6) runs with coordination **15 months** ≥4: Competent; 3: Emerging, ≤2: At-Risk **16 months** ≥5: Competent, 4: Emerging, ≤3: At-Risk **17 – 18 months** 6: Competent, 5: Emerging, ≤ 4: At-Risk	**17 and 18 months** 1) sits down from standing in a controlled manner 2) stands independently 3) stands up alone 4) walks alone with coordination (for at least 5 steps) 5) squats with support 6) runs with coordination **17 + 18 months**: 6: Competent, 5: Emerging, ≤ 4: At-Risk

Figure 6.2 Development of the third version of the new screening tool

6.3 Testing of the Third Version of the New Screening Tool

Once the third version of the new screening tool had been developed as outlined above, it was tested against the BSID III Gross Motor facet as outlined in 6.1 above.

6.3.1 Statistical Analysis

The statistical analysis was carried out as outlined in 6.1.1 above. The results are presented in Table 6.4 below:

Table 6.4 Agreement between the BSID III Gross Motor Scaled Score and the Second Version of the New Screening Tool (n = 112)

Age group	Kappa	Percentage of Children whose Scores agreed on both Tools
6 – 8 months	**0.86**	92.1%
9 – 11 months	**0.85**	91.3%
12 – 14 months	**0.84**	90.5%
15 – 18 months	**0.88**	93.3%
Total Group	**0.87**	92%

After three rounds of refinement and further testing, the agreement between the BSID III Gross Motor facet and the new screening tool was excellent in all age groups ($k \geq 0.75$). Through the processes outlined in this chapter, the item selection for each age group and the scoring was refined and tested against the BSID III Gross Motor facet. The new screening tool was refined as much as possible through the process outlined above, and at this stage needed external comment and further testing. The next step in the process was to invite a panel of experts in the field of paediatrics to evaluate the tool for content validity. This process is outlined in chapter seven.

Chapter 7

STUDY FOUR: EXPERT PANEL

The objectives for study four were as follows:

- To have the new screening tool examined by a panel of experts for content validity
- To adjust the new screening tool according to recommendations made by the expert group

7.1 Procedure

A panel discussion was conducted with seven experts in the field of paediatrics. A modified version of the Nominal Group Technique (NGT) was used in order to facilitate discussion and obtain consensus (Potter et al, 2004). Participants were considered to be experts if they met the following criteria:

a) A minimum of 10 years of experience in paediatrics
b) A medical/physiotherapy/occupational therapy degree
c) Currently practicing in the field of paediatrics
d) Preferably involved in research
e) Familiar with the administration of standardised assessment tools

A letter of invitation was sent to each person, outlining the proposed session, and asking whether they would like to participate in the session. The session was held in the staff room of the physiotherapy department at the University of the Witwatersrand, and was recorded using a dictaphone. The participants met all the criteria outlined above and were representative of the following professions: six paediatric physiotherapists and a developmental paediatrician. The physiotherapists were working in the following settings: one in a tertiary government hospital, two in private paediatric practice, and three in academia.

The NGT protocol states that the steps below should be followed (as discussed in chapter two):

1. Introduction and explanation: The participants should be welcomed, and an

explanation of the purpose and procedure of the meeting should be given.

2. Generation of ideas: Each participant should be provided with a sheet of paper with the question to be addressed and are then asked to write down all ideas that come to mind when considering the question. During this period, participants are asked not to consult or discuss their ideas with others. Approximately ten minutes are allowed for this.

3. Sharing of ideas: Participants should be invited to share the ideas they have generated. The process continues until all ideas have been presented. There is no debate about items at this stage and participants should write down any new ideas that may arise from what others have shared. This process ensures that all participants get an opportunity to make an equal contribution and provides a written record of all ideas generated by the group.

4. Group discussion: Participants are invited to ask for further explanation and detail about any of the ideas that colleagues have produced.

5. Voting and ranking: This involves prioritising the recorded ideas in relation to the original question. Immediate results are available to participants so that the outcomes are reached by the time the meeting concludes. It is recommended that the process is recorded on videotape or audiotape (Potter et al, 2004).

This format was followed, and the following issues were addressed:
1) Appropriateness of each item selected (the item itself, and item wording)
2) Appropriateness of the age groups which had been selected
3) Appropriateness of the items for the age-group in which they fell
4) Appropriateness of scoring

Each participant was provided with the proposal for the study, the methodology used in developing the new screening tool, and a copy of the new screening tool. Participants were given these documents on arrival, and were asked to look through them, so as to familiarise themselves with the process. The participants were

welcomed, introduced to one another, and a verbal explanation of the study, and the objectives to be addressed were given.

This technique has been used in other content validity studies (Palisano et al, 2008), and consensus was defined as 80% of participants being in agreement. The same percentage of agreement was used in this study. All concerns were discussed and appropriate changes were suggested by participants. There was 100% agreement for objectives two, three and four, and 71.4% agreement for objective one. After discussion about appropriate changes to be made, and documenting these, agreement for objective one was 100%. Table 4.1 shows the agreement for each objective:

Table 7.1 Agreement between Participants for Each Question

Question	Agreement	Agreement after changes suggested
Appropriateness of each item selected	71.4%	100%
Appropriateness of the age groups which had been selected	100%	n/a
Appropriateness of the items for the age-group in which they fell	100%	n/a
Appropriateness of scoring	100%	n/a

The concerns raised about question one were discussed, and are outlined in 7.1 below:

7.1 Issues Emerging from the Panel Discussion

During the panel discussion, two issues were raised: it was suggested by one member of the expert panel that item number two in the 6-8 month age group ‘lifts chest while lying on stomach (with extended arms)’ be replaced with a new item ‘pull to sit’. This was due to the concern that should a child have increased extensor tone, they would

appear to be able to elevate their trunk while lying in prone. It was concluded that this change should not be made for the following reasons:

- 'pull-to-sit' did not differentiate as well between the groups during the item analysis: 13% of the 'At-Risk' group obtained a credit for 'pull-to-sit', whereas 0% of the 'At-Risk' group obtained a credit for 'elevates trunk while prone on extended arms'. 'Pull-to-sit' is often a difficult aspect to assess unless the assessor is skilled, as there are a number of compensatory strategies which the child may use.
- The item 'elevates trunk while prone on extended forearms' is also used in the Gross Motor Function Measure (GMFM) (Rosenbaum et al, 1990; Russell et al, 1989) which is specifically designed to assess children with cerebral palsy. This assessment tool is known as the 'gold standard' for assessment of children with cerebral palsy, and is a valid and reliable tool. This indicates that the item is reliable for use in children who have increased tone, and may be used on HIV-infected children who present with increased tone due to HIV encephalopathy.

b) The wording for item number four in the 9-12 month age group was queried. The initial wording was 'crawls forward at least 1½ metres on hands and knees'. It was suggested that the words "move forward" replace the word "crawls", as children may do a number of different forms of crawling, which are all still normal. The revisions are presented in Figure 7.1 below:

THIRD VERSION	FOURTH VERSION
6 - 8 months 1) controls head at 90 degrees while lying on stomach 2) lifts chest while lying on stomach (with extended arms) 3) plays with feet 4) rolls from back to stomach 5) sits alone	**6 - 8 months** 1) controls head at 90 degrees while lying on stomach 2)lifts chest while lying on stomach (with extended arms) 3) plays with feet 4) rolls from back to stomach 5) sits alone
6 +7 months ≥ 4: Competent; 3: Emerging, ≤ 2: At-Risk **8 months** 5: Competent, 4: Emerging, ≤ 3: At-Risk	**6 +7 months** ≥ 4:Competent; 3: Emerging,≤ 2: At-Risk **8 months** 5:Competent, 4: Emerging, ≤ 3: At-Risk
9 - 12 months 1) turns body while seated 2) makes stepping movements 3) moves from sitting to being on hands and knees 4) crawls at least 1 ½ metres on hands and knees 5) pulls up to standing position	**9 - 12 months** 1) turns body while seated 2) makes stepping movements 3) moves from sitting to being on hands and knees 4) crawls /moves forward at least 1 ½ metres 5) pulls up to standing position
9 +10 months ≥ 4: Competent; 3: Emerging, ≤ 2: At-Risk **11 +12 months** 5: Competent, 4: Emerging, ≤ 3: At-Risk	**9 +10 months** ≥ 4: Competent; 3: Emerging, ≤ 2: At-Risk **11 +12 months** 5: Competent, 4: Emerging, ≤ 3: At-Risk
13 – 16 months 1) pulls up to standing position 2) sits down from standing in a controlled manner 3) stands independently 4) walks alone with coordination 5) squats without support	**13 – 16 months** 1) pulls up to standing position 2) sits down from standing in a controlled manner 3) stands independently 4) walks alone with coordination 5) squats without support
13 months ≥3: Competent; 2: Emerging, ≤ 1: At-Risk **14 months** ≥4: Competent; 3: Emerging, ≤ 2: At-Risk **15 + 16 months** 5: Competent, 4: Emerging, ≤ 3: At-Risk	**13 months** ≥3: Competent; 2: Emerging, ≤ 1: At-Risk **14 months** ≥4: Competent; 3: Emerging, ≤ 2: At-Risk **15 + 16 months** 5: Competent, 4: Emerging, ≤ 3: At-Risk
17 and 18 months 1) sits down from standing in a controlled manner 2) stands independently 3) stands up with no assistance 4) walks alone with coordination 5) squats with support 6) runs with coordination	**17 and 18 months** 1) sits down from standing in a controlled manner 2) stands independently 3) stands up with no assistance 4) walks alone with coordination 5) squats with support 6) runs with coordination
17 + 18 months: 6: Competent, 5: Emerging, ≤ 4: At-Risk	**17 + 18 months**: 6: Competent, 5: Emerging, ≤ 4: At-Risk

Figure 7.1 Development of the fourth version of the new screening tool

After the appropriate changes had been made following the expert panel discussion, the fourth and final version of the new screening tool was ready for preliminary testing to determine whether it was suitable for clinical use. Preliminary testing of the screening tool against the Bayley Scales of Infant and Toddler Development-III Gross Motor facet was then carried out; this is outlined in chapter eight.

Chapter 8

STUDY FIVE: PRELIMINARY TESTING OF THE FOURTH VERSION OF THE SCREENING TOOL

Following the expert panel group, the new screening tool was ready for preliminary testing against the BSID III. The objective of study five was to determine whether there was acceptable agreement between the BSID III and the new screening tool.

The objectives for study five were as follows:

- To test the fourth version of the new screening tool against the Bayley Scales of Infant and Toddler Development III Gross Motor facet (n = 60) in order to establish concurrent validity
- To establish diagnostic accuracy of the new screening tool through sensitivity and specificity values (n=60)
- To establish interrater reliability, test-retest reliability and intrarater reliability for the new screening tool (n=15 for test-retest reliability, n=15 for interrater reliability)

8.1 Concurrent Validity of the New Screening Tool against the BSID III

8.1.1 Sample Size Determination for Evaluation of Concurrent Validity

A minimum sample size of 60 infants is necessary to satisfy Nunnally and Bernstein's (1994) requirements for complex statistical procedures. The requirements state that the facet on the assessment tool which contains the highest number of items should be used, and that this number should be multiplied by ten to obtain the sample size. The highest number of items on the screening tool is six. Therefore 60 consecutive children between the ages of six and 18 months were tested. The sample was again drawn from Harriet Shezi Children's Clinic. The age groups were stratified in order to ensure that sufficient data were obtained in all age groups, and therefore 15 children per age group were assessed. This is illustrated in Table 8.1 below:

Table 8.1 Stratification of Sample for Study Five (n=60)	
Age groups	**n per group**
6 – 8 months	15
9 – 12 months	15
13 – 16 months	15
17 – 18 months	15
Total	60

8.1.2 Procedure

The files for each child attending the clinic that day were laid out as soon as the caregiver registered at reception. The files were screened to determine the ages of the infants who were registered for the clinic, and all infants between the ages of six and 18 months were selected. The caregivers of the children who met the inclusion criteria were approached, and the caregiver was given a letter of information regarding the study (See Appendix VI). Where the caregiver could not speak or understand English, a counsellor or nurse was asked to translate. Once the caregiver agreed to participate, he/she was asked to give written informed consent for the child and his/herself (see Appendix VI). The child was assessed on both the Bayley Scales of Infant Development III Gross Motor facet, and the new screening tool.

8.1.3 Statistical Analysis

Once scores for both assessment tools were obtained, the results were entered into an Excel spreadsheet. Statistical analysis was done in conjunction with the Medical Research Council of South Africa using Stata. The agreement between the tools was analysed using Kappa. This was done twice: the first analysis used three possible scoring outcomes (At-Risk, Emerging and Competent) (shown in Table 8.2), and the second analysis used two (At-Risk and Emerging) (shown in Table 8.3).

For conversion to two scoring outcomes, the following procedure was used: all the data (n=60) were analysed, and items for the children who fell into the ‘Emerging’ category on either the BSID III or the new screening tool were examined. After careful analysis, the BSID III scaled scores which translated into the ‘Emerging’ category were dealt with as follows:

- Scaled scores of 5 and 6 were classified as “At-Risk”
- Scaled scores of 7 were classified as “Competent”

This was done through thorough item analysis and the use of clinical judgement to determine whether a child who could not perform the items that they received a ‘no credit’ for on the BSID III Gross Motor scale would be considered to be developmentally competent, or at risk for developmental problems. On the new screening tool, the Emerging category was always converted to “At-Risk”.

8.1.4 Results

The agreement between the BSID III and the new screening tool was tested using a Kappa statistic for ordinal data. The results are presented below.

Table 8.2 shows the agreement between the BSID III Gross Motor facet, and the fourth version of the new screening tool with **three** scoring outcomes. Table 8.3 shows the agreement between the BSID III Gross Motor facet, and the fourth version of the new screening tool with **two** scoring outcomes.

Table 8.2 Results of the BSID III Gross Motor Facet versus the New Screening Tool with Three Scoring Outcomes (n=60)

Age group	n	Percentage of Children whose Scores Correlated on both Tools	Kappa
6 – 8 months	15	73.3%	0.56
9 – 11 months	15	80%	0.66
12 – 14 months	15	80%	0.51
15 – 18 months	15	86.7%	0.69
Total group	60	80%	0.63

The agreement for the version with **three** score categories was moderate in all age groups. The new screening tool assessed the children more conservatively than the BSID III, and therefore where there were discrepancies in scoring, the new screening tool score was always within one category of the BSID III, and was always one score

category below, indicating that it assesses children more strictly. For example, for a child whose scores differed on the two tools: where a score on the BSID III was in the Competent category, the new screening tool score would place the child in the Emerging category. Sensitivity, specificity, positive predictive value and negative predictive value cannot be calculated for three outcomes (Aylward, 2000).

Table 8.3 Results of the BSID III Gross Motor Facet versus the New Screening Tool with Two Scoring Outcomes (n=60)

Age group	n	Percentage of Children whose Scores Correlated on both Tools	Kappa
6-8 months	15	86.7%	0.73
9-12 months	15	100%	**1**
13-16 months	15	100%	**1**
17-18 months	15	86.67%	0.60
Total group	60	93.3%	**0.85**

The overall Kappa for agreement between the BSID III and the new screening test with **two** scoring outcomes was excellent (0.85) (Landis and Koch, 1977). Within the age categories, there was 100% agreement in the 9-11 month group and the 12-14 month group. Kappa was moderate in the 6-8 month group, and the 15-18 month group. Tables of symmetry for the 6-8 month group and the 9-11 month group are presented below in Table 8.4 in order to explain the slightly lower Kappa values:

Table 8.4 Symmetry between the New Screening Tool and the BSID III for 6-8 Months (n=15)

Screening Tool	BSID III		
	At-Risk	**Competent**	**Total**
At-Risk	6	1	7
Competent	1	7	8
Total	7	8	15

In Table 8.4, the agreement between the two tools is shown along the diagonal (i.e. both tools scored six children as being At-Risk, and both scored seven as being Competent). There were only two cases where the tools did not agree. In the first case, the BSID III scored the child as being Competent, and the new screening tool placed the child in the At-Risk category. In the second case, the new screening test placed the child in the Competent category, and the BSID III placed them in the At-Risk category. Table 8.5 below shows the symmetry between the tools in the 17-18 month age group.

Table 8.5 Symmetry between the New Screening Tool and the BSID III for 17-18 Months (n=15)

Screening Tool	**BSID III**		
	At-Risk	**Competent**	**Total**
At-Risk	11	2	13
Competent	0	2	2
Total	11	4	15

In the Table 8.5 above, again the agreement is shown along the diagonal (both tools scored 11 children as being At-Risk, and both scored two children as being Competent). There were two cases where the tools did not agree. The BSID III scored the children as being Competent, and the new screening tool placed them in the At-Risk category. This indicates that the new screening tool may be slightly more conservative in terms of scoring. Sensitivity and specificity were therefore tested in order to assess these capabilities of the new screening tool further.

8.2 Evaluation of the Diagnostic Accuracy of the New Screening Tool

The diagnostic accuracy of the new screening tool was evaluated by examining sensitivity, specificity, positive predictive value (PPV), and negative predictive value (NPV) against the BSID III.

Sensitivity is defined as the proportion of children who are delayed and are correctly identified as delayed by the screening test, the ‘‘true positive’’ group. Specificity is defined as the proportion of children who are not delayed and are correctly identified

as not delayed by the screening test, the ‘‘true negative’’ group. The PPV is the probability that a child is truly developmentally delayed when results of the screening test are positive. Unlike the sensitivity and specificity of a screening test, which can be thought of as characteristics of the test itself, the PPV is dependent upon the prevalence of the disorder in the population tested (Frankenberg et al, 1992). Therefore, the sensitivity of the test may be a better measure to evaluate a screening tool for developmental delay than its PPV (Hess, 2004). The negative predictive value of a test is the proportion of patients with a negative screening result who do not have the condition

8.2.1 Diagnostic Properties of the New Screening Tool

The diagnostic properties of the new screening tool are presented in Table 8.6 below:

Table 8.6 Diagnostic Properties of the New Screening Tool

Sensitivity	97.4%
Specificity	85.7%
Positive predictive value	92.7%
Negative predictive value	94.7%

The sensitivity of the new screening tool is high, and indicates that the new screening tool identified 97.4% of the children who were delayed according to the BSID III. The specificity is also high, and the new screening tool identified 85.7% of children who were not delayed as being Satisfactory. The specificity is slightly lower than the sensitivity, as the new screening tool was slightly more conservative than the BSID III in classifying children as being At-Risk. The PPV is high at 92.7%, which indicates that there is a probability of 92.7% that children who were classified as having gross motor delay in this population were correctly diagnosed – the fact that gross motor delay in this population is very prevalent has to be taken into account, and therefore accounts for the high PPV. The negative predictive value (NPV) overall is high at 94.7%, which indicates that there is a probability of 94.7% that children who were classified as Competent do not have gross motor delays.

8.2.2 Diagnostic Properties per Age Group

Table 8.7 illustrates the diagnostic properties of the new screening tool for the 6-8 month age group

Table 8.7 Diagnostic Properties of the New Screening Tool for 6-8 Months	
Sensitivity	85.7%
Specificity	87.5%
Positive predictive value	85.7%
Negative predictive value	87.5%

For the age group 6-8 months, the sensitivity of the new screening tool is adequately high, and indicates that the new screening tool identified 85.7% of the children that were delayed according to the BSID III. The specificity is also high, and the new screening tool identified 87.5% of children who were not delayed as being satisfactory. The specificity is slightly higher than the sensitivity, but only marginally so. The PPV is 85.7%, which indicates that there is a probability of 85.7% that children who were classified as having gross motor delay in this population were correctly diagnosed. The NPV indicates that 87.5% of children who were classified as Competent do not have gross motor delay. Table 8.8 illustrates the statistical properties of the new screening tool for 9-12 month age group

Table 8.8 Diagnostic Properties of the New Screening Tool for 9-12 Months	
Sensitivity	100%
Specificity	100%
Positive predictive value	100%
Negative predictive value	100%

For the age group 9-12 months, the sensitivity of the new screening tool is excellent, and indicates that the new screening tool identified 100% of the children that were delayed according to the BSID III. The specificity is also excellent, and the new

screening tool again identified 100% of children who were not delayed as being so. The PPV is also 100%, which indicates that there is a probability of 100% that children who were classified as having gross motor delay in this population were correctly diagnosed. Similarly, the NPV is 100%, which indicates that there is a probability of 100% that children who were classified as Competent do not have gross motor delay. Table 8.9 illustrates the diagnostic properties of the new screening tool for the 13-16 month age group:

Table 8.9 Statistical Properties of the New Screening Tool for 13-16 Months

Sensitivity	100%
Specificity	100%
Positive predictive value	100%
Negative predictive value	100%

For the age group 13-16 months, the sensitivity of the new screening tool is excellent, and indicates that the new screening tool identified 100% of the children who were delayed according to the BSID III. The specificity is also excellent, and the new screening tool again identified 100% of children who were not delayed as being so. The PPV is also 100%, which indicates that there is a probability of 100% that children who were classified as having gross motor delay in this population were correctly diagnosed. Similarly, the NPV is 100%, indicating that there is a probability of 100% that children who were classified as being Competent do not have gross motor delay. Table 8.10 below illustrates the diagnostic properties of the new screening tool for the 17-18 month age group:

Table 8.10 Diagnostic Properties of the New Screening Tool for 17-18 Months

Sensitivity	100%
Specificity	50 %
Positive predictive value	84.6%
Negative predictive value	100%

For the age group 17-18 months, the sensitivity of the new screening tool is excellent, and indicates that the new screening tool identified 100% of the children who were delayed according to the BSID III. The specificity was low, and the new screening tool identified 50% of children who were not delayed as being so. This is due to the manner in which specificity is calculated: (the number of normal children who also scored 'normal' on the test)/(normal children +normal children identified by a test as having delays) (Glascoe, 1991).

For this test, the calculation is as follows: (the number of normal children, who received a Competent score = 2)/(the number of normal children as identified by the BSID III [2]+the number of normal children who were identified by the new screening tool as having delays [2]) = 50%. The PPV is 84.6%, which indicates that there is a probability of 84.6% that children who were classified as having gross motor delay in this population were correctly diagnosed. The NPV is 100%, which indicates that there is a probability of 100% that children who were classified as being Competent do not have gross motor delay.

8.3 Final Version of the New Screening Tool

The two final versions (one with three scoring outcomes, and one with two scoring outcomes) of the new screening tool were taken to an expert in the field of paediatric HIV, in order to discuss which one would be more approriate for use in the HIV clinics. After careful analysis of both versions, it was concluded that a screening tool with only two possible scoring outcomes ("At-Risk" and "Satisfactory") would be more appropriate, as an "Emerging" category may lead to loss to follow-up of children, and less vigilance in terms of sending those children with any risk of developmental delay for further assessment.

8.4 Reliability of the New Screening Tool

The reliabilty of the final version of the new screening tool was evaluated as follows:

8.4.1 Test-retest Reliability and Intrarater Reliability

Test-retest reliability was evaluted at Rahima Moosa Hospital in Johannesburg, Gauteng. Consent was obtained from the hospital CEO in order to carry out data collection (see Appendix VII). Fifteen infants (which is 25% of the sample size for study one) between the ages of 6-18 months who had been admitted to the wards as

in-patients were assessed twice within two days. The reasons for the use of this site, and the time interval were as follows: children attending HIV clinics as outpatients are scheduled for follow-up visits every three to six months, during which time their development would change considerably. Bringing the child back to the clinic within one to two weeks following their scheduled visits is very difficult for a number of logistical reasons: a) transport is costly to the patients, and therefore a transport fee would have to be given b) it is often difficult for the parent/caregiver of the child to attend more regularly due to work commitments. Inpatients are therefore easier to evaluate, and this is still within the intended context of use for the tool, as HIV positive infants who are inpatients will also be assessed using the tool. In terms of the time interval, due to the fact that the infants were inpatients, and were likely to be in varying stages of recovery, a short time period between assessments was selected to ensure that the infants did not change clinically. These patients were admitted with chronic conditions, and therefore no clinical improvement was expected. Once informed consent had been obtained from the child's caregiver (see Appendix VIII), the assessment was carried out by a paediatric physiotherapist, who had undergone training on the new assessment tool, and worked at Rahima Moosa Hospital, so knew the ward routines and procedures well. The assessments for each child were carried out at the same time of the day in order to minimise effects by medication, and caregiving routines.

In a test-retest situation, when a rater's skill is relevant to the accuracy of the test, intrarater and test-retest reliability are essentially the same estimate (Portney and Watkins, 2000). In this study, the rater's skill was relevant to the accuracy of the test, and therefore test-retest reliability and intrarater reliability are the same estimate. The results were entered into an Excel speadsheet, and the agreement between the score obtained on the first and second assessment was evaluated using the Pearson product-moment correlation coefficient, as the data is nominal data (Portney and Watkins, 2000). The results of the test-retest and intrarater reliability study are shown in Table 8.11 below:

Table 8.11 Test-retest and intrarater Reliability for the New Screening Tool (n=15)	
Test-retest and Intrarater reliability	r = 0.98

There were two cases where the scores differed by one point, and the score on the first assessment was higher than the second in both cases. The category in which the child was placed was not affected by the differences in scores, and in all cases, the developmental category (At-Risk or Satisfactory) was the same on the first and second assessments. The value of r=0.98 indicates excellent test-retest and intrarater reliability (Portney and Watkins, 2000), and shows that the new screening tool is stable, and obtains the same results with repeated administrations of the test, as well as stability of data obtained by one rater over two assessments (Portney and Watkins, 2000).

8.4.2 Interrater Reliability

Interrater reliability was conduted at Harriet Shezi Clinic at Chris Hani Baragwanath Hospital on 15 consecutive infants (which is 25% of the sample size from study one) between the ages of 6 and 18 months. Once informed consent had been obtained from the caregiver of the child (see Appendix IX), two raters observed the gross motor behaviour of the infant and allocated scores similtaneously. One rater was the researcher, and the other a paediatric physiotherapist who had undergone training on the new screening tool. The results were entered into an Excel spreadsheet, and the agreement between the score obtained by the first rater and the second rater was analysed using the Pearson product-moment correlation coefficient, for nominal data (Portney and Watkins, 2000). The results are shown in Table 8.12 below:

Table 8.12 Interrater Reliability for the New Screening Tool (n=15)	
Interrater reliability	r = 1

The results for interrater reliability were excellent (r = 1) (Portney and Watkins, 2000), and indicates that both raters obtained the same score on each assessment.

8.5 Conclusion

The results of this study indicate that children infected with HIV are at risk for developmental problems from an early age, This, along with the effects of HIV on growth, and the low socioeconomic status of this population present added risk factors for developmental delay. There is therefore a need for a screening tool which is able to detect these problems so that children may be referred on for more in-depth assessment and management.

The findings of this study indicate that the results of the Bayley-III Screening Test and the BSID III do not correlate well, and therefore that the Bayley-III Screening Test is not suitable for use on HIV infected infants, as it does not adequately identify the children's problems. A new screening tool was therefore developed, and the process has been outlined above. The preliminary testing for validity, reliability and diagnostic properties of the new screening tool were excellent. The implications of the above findings, as well as the development of the screening tool will be discussed in more detail in chapter nine.

Chapter 9

DISCUSSION

The results of this study will be discussed in more detail in this chapter. The main outcomes will be discussed as well as important clinical data which were collected. A number of suggestions for future research and clinical management will be put forward. The new screening tool will now be referred to as "The Infant Gross Motor Screening Test".

9.1 Classification Accuracy of the Bayley-III Screening Test

One of the main objectives of this study was to examine the suitability of the Bayley-III Screening Test for use on HIV positive infants in South Africa. The need for developmental screening of HIV positive infants has been highlighted through many studies, which have been conducted in South Africa (Potterton et al, 2009b; Ferguson and Jelsma, 2009; Baillieu and Potterton, 2008). The development of a screening tool which is suitable for use in the paediatric HIV setting has been incorporated into the National Strategic Plan (NSP) for management of HIV for 2007-2011, and the South African government has made a commitment to ensuring that all HIV positive children under three years of age undergo developmental screening (NSP, 2007-2011). This has been reiterated in the South African Department of Health Antiretroviral Treatment Guidelines (2010), where developmental monitoring is recommended in order to assess children for ARV eligibility, as well as being one of the factors which is used to monitor response to treatment (DOH, 2010). The agreement between the Bayley-III Screening Test and the BSID III was not good enough ($k < 0.75$) for clinical use. The reasons for this will be outlined and discussed.

9.1.1 Agreement between the Bayley-III Screening Test and the BSID III

The Bayley-III Screening Test is designed to briefly assess the cognitive, language and motor function of infants and young children between 1-42 months of age (Bayley, 2006c). The primary purposes are to determine whether a child is progressing according to normal expectations, to determine the degree of risk for developmental delay, and to determine if further, more comprehensive evaluation is needed. The

purpose of using a screening tool on this population of HIV positive infants would be similar – children with developmental problems who have been identified by the screening tool would need to be referred for more comprehensive testing. The characteristics of the sample population have been outlined in chapter four.

The Bayley-III Screening Test was selected, due to the fact that it was developed from the Bayley Scales of Infant and Toddler Development (Bayley, 2006c), which has long been considered the 'gold standard' of infant developmental assessment (Harris et al, 2005; Tieman et al, 2005). The BSID has been used in a number of studies in South Africa to evaluate the neurodevelopmental outcomes of HIV positive infants (Potterton et al, 2009b, Ferguson and Jelsma, 2009; Baillieu and Potterton, 2008), and the BSID III has been shown to be suitable for use on black South African infants from similar socioeoconomic backgrounds (Rademeyer, 2010; Brown, 2009). The Bayley-III Screening Test was therefore compared with the BSID III in order to determine its suitability for use in this population.

The agreement between the Bayley-III Screening Test and the BSID III in the Cognitive facet ranged between K=0.41 and K=0.70. None of the Kappa scores in the Cognitive facet were higher than 0.70, representing moderate agreement across the Cognitive domain (Landis and Koch, 1977). The overall agreement in the Cognitive domain was K=0.58, which represents moderate agreement (≥0.4-<0.75) (Landis and Koch, 1977). When analysing the classification accuracy of the Bayley-III Screening Test for Cognitive function, the screening test identified 77.8% of At-Risk cases, 59.4% of Emerging cases, and 77.3% of Competent cases.

The agreement between the Bayley-III Screening Test and the BSID III was better in the Language facet. The Expressive and Receptive Communication domains had excellent agreement in all ages, groups apart from the 9-11 month group for Receptive Communication, and the 6-8 month age group for Expressive Communication. The overall agreement between the tools in the Expressive Communication facet was 0.76, and in the Receptive Communication facet 0.82, representing excellent agreement (≥ 0.75). When analysing the classification accuracy of the Bayley-III Screening Test for Language function, the screening test identified 82.4% of At-Risk cases, 94% of Emerging cases, and 89% of Competent cases for

Receptive Communication, and 83.33% of At-Risk cases, 68.4% of Emerging cases, and 94.7% of Competent cases for Expressive Communication.

The agreement between the Bayley-III Screening Test and the BSID III in the Motor facet ranged between K=0.28 and K=0.92. The agreement in the Fine Motor facet was poor to moderate for all age groups, ranging between K=0.28 and K=0.50. In the Gross Motor facet the agreement ranged between K=0.41 and K=0.92, with most age groups demonstrating moderate agreement (≥0.4 to <0.75). The overall agreement in the Fine Motor facet was K=0.44, and in the Gross Motor facet, was K=0.57 which represents moderate agreement in both (≥0.4 to <0.75). When analysing the classification accuracy of the Bayley-III Screening Test for Motor function, the screening test identified 58.8% of At-Risk cases, 56.7% of Emerging cases, and 68.8% of Competent cases for Fine Motor function, and 62.5% of At-Risk cases, 71.4% of Emerging cases, and 89.3% of Competent cases for Gross Motor function.

The agreement of the Bayley-III Screening Test with the BSID III is not optimal (K≤0.75) due to problems with classification accuracy. The Bayley-III Screening Test is more accurate for children who are classified as being in the Competent category, than for those in the At-Risk category, as stated in the manual (Bayley, 2006c). The population of children tested in this study are at high risk for developmental problems, especially for cognitive and motor function (Potterton et al, 2009b, Ferguson and Jelsma, 2009; Baillieu and Potterton, 2008), and therefore classification accuracy of the Bayley-III Screening Test will be moderate (Bayley, 2006c).

9.1.2 Classification Accuracy of the Bayley-III Screening Test

The trends in this study with regards to classification accuracy are similar to those in the Bayley-III Screening Test standardisation sample (Bayley, 2006c), and will be highlighted below. Bayley (2006c) stated that the Bayley-III Screening Test was more accurate in detecting those children classified as Competent, than those classified as At-Risk, and where the BSID III and Bayley-III Screening Test classified children differently, it was always within one category. For example, in the standardisation sample, children classified as At-Risk on the BSID III would be misclassified as Emerging by the Bayley-III Screening Test, but never as Competent (Bayley, 2006c).

9.1.2.1 Classification Accuracy of the Bayley-III Screening Test for Children Identified by the BSID III as At-Risk

For children in the At-Risk category, classification accuracy of the Bayley-III Screening Test was 'moderate' in the standardisation sample (Bayley, 2006c). The range of correct classification of children as At-Risk by the Bayley-III Screening Test compared with the BSID III for the standardisation sample was 41.8% for Fine Motor function to 66.9% for Receptive Communication. Not one of the children in the standardisation sample was misidentified as Competent on the Bayley-III Screening Test, and where the Bayley-III Screening Test and the BSID III did not agree, the Screening Test classified children in the Emerging category (Bayley, 2006c).

In this study, the range of correct classification of children who are in the At-Risk category by the Bayley-III Screening Test as compared with the BSID III was 58.8% for Fine Motor function to 83.3% for Expressive Communication. The range in this study was wider than that of the standardisation sample (41.8%-65.9%), and Expressive Communication had better accuracy than Receptive Communication, but only by one percent. There were a number of cases where children were misidentified as Competent by the Bayley-III Screening Test, where the BSID III classified them as At-Risk, which did not occur in the standardisation sample (Bayley, 2006c), indicating that possibly on higher-risk populations such as HIV infected infants, the Bayley-III Screening Test is not sufficiently sensitive. The range in this study for children misidentified as Competent was 5.4% for Gross Motor function to 8.3% for Cognitive function. There were more children in this study who were correctly identified as being At-Risk (83.3% for Expressive Communication, as opposed to 65.9% for Receptive Communication in the standardisation sample), although the children in this study are all at high risk for severe developmental problems (Potterton et al, 2009b, Ferguson and Jelsma, 2009; Baillieu and Potterton, 2008), whereas the standardisation sample was healthy children, stratified by socioeconomics, ethnicity, gender, age and maternal education (Bayley, 2006c). The fact that despite the risk factors for the current study population, 5.4%-8.3% were misidentified as being Competent by the Bayley-III Screening Test, whereas the BSID III classified them as being At-Risk demonstrates the inherent problems with classification accuracy in the Bayley-III Screening Test. Due to the high risk nature of the HIV infected paediatric population

for developmental problems, a screening tool which is more sensitive would be necessary to detect those children who are At-Risk for developmental problems.

9.1.2.2 Classification Accuracy of the Bayley-III Screening Test for Children Identified by the BSID III as Emerging

For children in the Emerging Category, classification accuracy was 'accurate' for children with scaled scores of 5-7 in the standardisation sample (Bayley, 2006). The range of correct identification of children with Emerging skills as compared with the BSID III for the standardisation sample was 64% for Cognitive function to 77.8% for Receptive Communication (Bayley, 2006). The percentage of children misidentified as being At-Risk was very low in the standardisation sample (0.8%-5.2%) (Bayley, 2006c).

In this study, the range for correct identification of children with Emerging Skills as compared with the BSID III is 56.7% for Fine Motor function to 94% for Receptive Communication. This indicates that fewer children on the lower end of the range in this study were correctly identified as being in the Emerging Category (56.7% in this study as opposed to 64% in the standardisation sample, and there were a number of cases where children were misidentified as being in the At-Risk category, which did not occur in the standardisation sample (Bayley, 2006c). The range in this study for children misidentified as being in the At-Risk category by the Bayley-III Screening Test where the BSID III classified them as being in the Emerging category was 5.3% to 16.7%, which is much higher than the standardisation sample (0.8%-5.2%). This indicates that a much higher percentage of children were misclassified in this study, which has important implications for the clinic. These children would have to be referred on for further assessment, which places an increased burden on already overloaded services. On the other hand, children misidentified as being in the Competent category by the Bayley-III Screening Test where they were classified as being in the Emerging category by the BSID III ranged from 6.1% for Receptive Communication to 28.6% for Gross Motor function, which indicates that at the upper end of the range, almost a quarter of children who should be monitored would be misclassified, and would no longer be monitored or assessed, and therefore may be lost to the system. The fact that at the upper end of the range, the highest number of misclassifications occurred in the Gross Motor facet is especially important seeing that

gross motor function is most severely affected in this population of children (Potterton et al, 2009b; Bailieu and Potterton, 2008).

9.1.2.3 Classification Accuracy of the Bayley-III Screening Test for Children Identified by the BSID III as Competent

Classification accuracy of the Bayley-III Screening Test was 'very accurate' for children with scaled scores of 8-19 in the standardisation sample (Bayley, 2006c). The range of correct classification of children in the Competent category by the Bayley-III Screening Test as compared with the BSID III for the standardisation sample was 84% for Cognitive function to 92.1% for Receptive Communication (Bayley, 2006c). None of the children in the standardisation sample who were classified by the BSID III as Competent were misidentified as being At-Risk on the Bayley-III Screening Test (Bayley, 2006c).

In this study, the range of correct classification of children in the Competent category on the Bayley-III Screening Test as compared with the BSID III is 68.8% for Fine Motor function to 94.7% for Expressive Communication. In this study, children with Expressive Communication in the Competent range were correctly identified more often than those with Receptive Communication skills in the Competent range. Fine Motor function had the lowest accuracy, which is different to the standardisation sample where the Bayley-III Screening Test identified children with cognitive skills in the Competent range least accurately. In this study, far fewer children were correctly identified on the lower end of the range (68.8% as compared with 84% in the standardisation sample), which brings into question the statement that Classification accuracy of the Bayley-III Screening Test was "very accurate" for children with scaled scores of 8-19 in the standardisation sample (Bayley, 2006c), as 68.8% is not sufficiently accurate. There were a few cases (2.3%) where children were misidentified as being in the At-Risk category by the Bayley-III Screening Test where the BSID III identified the children as being Competent, which did not occur in the standardisation sample (Bayley, 2006c). The range for children misidentified by the Bayley-III Screening Test as having Emerging skills was 5.3%-31.3% which is much higher. This is also important, as these children would be monitored or evaluated further, which again places an added burden on the system. Therefore in this study,

the Bayley-III Screening Test was not found to be sufficiently accurate in identifying children with skills in the Competent category.

Overall, the Bayley-III Screening Test did not show adequate agreement with the BSID III in order to be used clinically on this population. The reasons for this are as follows: Classification accuracy of the Bayley-III Screening Test is highest for those children in the Competent category, and lowest for those children in the At-Risk category according to testing against the BSID III in the standardisation sample (Bayley, 2006c), which is problematic for the population of HIV positive infants for which we are evaluating its suitability, as they are at increased risk for developmental problems in all areas (Potterton et al, 2009a; Potterton et al, 2009b; Baillieu and Potterton, 2008; Udgirkar et al, 2003; Blanchette et al, 2001; Pearson et al, 2000; Fragoso et al, 1999; Drotar et al, 1997; Wolters et al, 1997; Belman et al, 1996; Pollack et al, 1996; Chase et al, 1995; Nozyce et al, 1994; Belman, 1992; Belman et al, 1988; Epstein et al, 1986; Belman et al, 1985; Ultmann et al, 1985), and therefore this screening tool would not be accurate enough to consistently detect these problems. The use of percentages is not recommended, as it often overestimates true validity, as it fails to account for chance - therefore it would have been more useful for the authors of the Bayley-III Screening Test to use Pearsons, Spearmans, or Kappa to evaluate concurrent validity (Portney and Watkins, 2000).

In this study population, mean BSID III scores for Cognitive function were in the ‘Low Average’ to ‘Borderline’ range, and mean Motor scores were in the ‘Borderline’ to ‘Extremely Low’ range, indicating that cognitive and motor function are severely compromised in this population. It would therefore be expected that even though classification accuracy for the Bayley-III Screening Test may be compromised in a healthy population, when used on children with severe delays, that accuracy may improve due to extremely low scores. The fact that there were a number of misclassifications, and that the Bayley-III Screening Test identified between 58.8%-83.3% of children who were in the At-Risk category on the BSID III indicates that far too many children would be missed if this tool were to be implemented. It has been demonstrated over and over again that neurological and developmental abnormalities are frequent complications of HIV infection in children, especially in younger perinatally affected children (Potterton et al, 2009b; Baillieu and Potterton, 2008;

Raskino et al, 1999; Nozyce et al, 1994; Belman, 1992), therefore most children in the study population would be at risk for developmental problems, and this tool would not be accurate enough to detect these.

9.2 Development of a New Screening Tool

Following the analysis of the agreement between the BSID III and the Bayley-III Screening Test, it was clear that a new screening test should be developed, which was more appropriate for the paediatric HIV infected population. The choice of a tool in a clinical context needs to reflect the purpose for which the tool is to be used. If the purpose is to screen for problems in high-risk populations, a tool is needed that is quick and easy to administer, yet has good sensitivity and specificity (Scott et al, 2007).

As described in chapter eight, the following aims and objectives for the Infant Gross Motor Screening Test were drawn up in order to ensure that it would be suitable for use in a busy clinic in a developing country:

- It should identify those children who are experiencing developmental problems, so that they may be referred for further assessment
- It should be sensitive enough to distinguish between children who are experiencing problems and those who are not
- It should be quick and easy to administer due to time constraints
- It should not require profession specific training, so that counsellors and nurses may also administer it, should the need arise
- It should be inexpensive enough to reproduce, and should not require any special equipment

HIV clinics in South Africa have a number of human resource challenges which include shortages of staff, and little mentorship and training of doctors and nurses (Meyers et al, 2007). It is therefore vital that a screening tool which is developed for use in these clinics meets the criteria outlined above.

9.2.1 Developmental Domains of the Infant Gross Motor Screening Test

Through analysis of the BSID III composite scores, and of the literature on HIV and neurodevelopment, a gross motor screening tool would be most suitable for this population. This is discussed in 9.2.1.1 below:

9.2.1.1 Analysis of Literature and BSID III Composite Scores

It has been consistently found that motor development, more specifically gross motor development is the facet of development which is most severely affected in HIV positive infants. Motor development is compromised from an early stage and HIV positive infants and children may start presenting with motor impairments in the first few months of life (Potterton and Eales, 2001; Chase et al, 1995; Nozyce et al, 1994). Muscle strength is decreased in children infected with HIV (Blanchette et al, 2001; Pearson et al, 2000), which may be a factor contributing to the fact that gross motor development is often more affected than fine motor development (Baillieu and Potterton, 2008; Potterton and Eales, 2001; Parks and Danoff, 1999).

The composite scores of the infants in this study give an indication as to the developmental status of the sample. Language function is the area which is least affected in the sample, although the only scores which fell into the Average range in the Language domain were those in the two younger age groups. This may be that language function at this stage is less demanding, and language relies mainly on receptive function at younger ages, which is consistently less affected in HIV positive children than expressive language (Wolters et al, 1997; Wolters et al, 1995). The scores in the Language domain were highest throughout the groups, which is consistent with studies done worldwide, where language function is the developmental domain which is least affected (Wolters et al, 1997; Wolters et al, 1995).

Cognitive function is less affected than motor function in this sample, although the scores are still very low and fall into the ‘Low Average’ and ‘Borderline’ categories. Motor function is the area of development which is consistently most severely affected, and falls into the ‘Borderline’ or ‘Extremely Low’ categories. This is also consistent with the findings of previous studies on similar populations, where gross motor function is the area which is most severely affected (Potterton et al, 2009b; Baillieu and Potterton, 2008).

9.2.1.2 Analysis of Factors Leading to Poor Developmental Scores

In this study, in all developmental domains, the developmental scores worsened as the children got older. It was found that immunological and virological status improved with age, which is most likely due to the effect of HAART, which has a direct effect on immunological status, especially HAART containing a PI, which the regimen used at Harriet Shezi contains. The PI used at Harriet Shezi is Kaletra ® (Lindsey et al, 2007). Lindsey et al (2007) found that HAART containing a PI had a limited positive effect on neurodevelopment in the first two years of life, and that infants with severe immune suppression were at the highest risk for poor neurodevelopmental outcomes. This indicates that preventing immune suppression may be better than treatment which is initiated once the immune system is already suppressed. Patel et al (2009) examined the CNS penetration capacities of antiretroviral drugs and ranked each drug according to its penetration capacity (See Table 2.2). Antiretroviral regimens with scores of less than four are classified as low CNS-penetrating regimens, scores of four to five are classified as medium CNS-penetrating regimens, and scores greater than or equal to six are classified as high CNS-penetrating regimens (Patel et al, 2009). Children who were on a high CNS-penetrating regimen had a 41% decrease in the incidence of encephalopathy compared with those children on a low CNS-penetrating regimen (Patel et al, 2009).

Children who are treated at Harriet Shezi clinic receive a first-line treatment regimen that has a CNS penetration score of seven, indicating that it is a high CNS-penetrating regimen. Despite this, infants in this study had very low developmental scores, which is most likely due to late HAART initiation once irreversible damage of the CNS had already occurred. This highlights the need for a screening tool, which is able to detect developmental problems early, and help to identify children in need of developmental and medical assistance.

9.2.1.3 Cultural Considerations in the Development of Screening Tools

Culture has an effect on the appropriateness of assessment tools. Gladstone et al (2008) adapted the Denver-II for use in Malawi, and state that in all domains of Western tests, there are items which are culturally inappropriate for a rural African population. Gladstone et al (2008) found that items such as ‘‘prepares cereal’’ or ‘‘plays board games/card games’’ are uncommon activities for children in rural Africa.

They also found that the pink doll in the Denver II kit was terrifying to most children. There are similar items in the BSID III, and the Bayley-III Screening Test, and the same problems have been encountered in South Africa (Potterton et al, 2009b; Baillieu and Potterton, 2008).

Language assessment is also difficult, and many assessment tools have pictures of objects that children in a peri-urban/rural African setting have never seen before (Gladstone et al, 2008). This makes it difficult to name objects, especially as many children have also never seen a book, or pictorial representations of these objects. Gladstone et al (2008) found that social skills were the least universal across cultures, and gross motor skills were the most universal when comparing African and Western outcomes on an assessment tool, which has been developed for children growing up in a Westernised setting.

Abubakar et al (2007) found that the lack of culturally appropriate and standardised measures of childhood outcome has been a major impediment in adequately studying the effects of disease exposure in sub-Saharan Africa. Their results indicate that it is feasible to develop culturally appropriate measures that are sensitive to the effects of HIV. This therefore reitrates that a gross motor screening tool would be appropriate for use in the paediatric HIV setting, as gross motor function is universal, and less culturally sensitive than any other developmental skills (Kelly et al, 2006).

9.2.2 Characteristics of the Standardisation Sample of the New Screening Tool

According to the American Psychological Association (1985) the normative sample should be determined by the country's census using racial, geographic, and socioeconomic distributions.

The Infant Gross Motor Screening Test was developed specifically for the study population: that is, HIV positive infants living in low urban socioeconomic conditions, and aged between 6-18 months, and therefore testing was carried out on this specific population against the BSID III. The infants varied in their developmental, virological and immunological status. The socioeconomic status of the sample was low, but this sample is representative in terms of socioeconomic background of HIV positive infants attending urban paediatric HIV clinics in the government sector in Gauteng. Maternal

education of the sample was low, with only 31.3% having completed school. The sample consisted primarily of black African infants, who live in Soweto, in Gauteng. The Infant Gross Motor Screening Test was developed for use on this population, and therefore should it be considered for use on other populations, further standardisation will need to be carried out.

9.2.2.1 Risk-factors Affecting Gross Motor Development in the Standardisation Sample

Socioeconomic disadvantage has been shown to have a negative educational effect on language and reading skills, as well as on the motor skills of children (McPhillips and Jordan-Black, 2007). In this study 81% of primary caregivers were mothers, with grandmothers taking on this responsibility where mothers were no longer alive. Most primary caregivers had not completed school, and therefore maternal education levels were low. When taking the average number of people per household (4.8 people) together with the fact that 45% of the families live in a shack, and in 45% of families there is only one adult bringing in an income, the socioeconomic status of the sample was determined to be low.

The mediating effect of socioeconomic status (SES) on neurodevelopment was examined by Abubakar et al (2008b), where it was found that SES has a significant effect on psychomotor development through its influence on height-for-age (stunting) and weight-for-age (malnutrition) which were both examined in this study. In terms of malnutrition in this sample, only the youngest group (6-8 months) was affected, with a mean weight-for-age of -2.29, which is indicative of moderate malnutrition. The older groups fall within the normal range (0 to -2) for this measure. This may be that the younger group has not had access to antiretroviral therapy, or nutritional counselling, having most likely been recently diagnosed, and therefore would not have attended many clinic sessions.

Stunting was a problem in this study population, and was most severe in the 6-8 month old age group, where there is moderate stunting (height-for-age = -2.26). The children in the 9-14 month age group were not affected by stunting, but the children in the 15-18 month old age range also had moderate stunting (height-for-age = - 2.87). It has been found that HIV infected infants are more commonly stunted than wasted

(Tindyebwa, 2004). There is likely to be an effect due to the above anthropometric factors on gross motor development. This is due to the mediating effect of SES on growth, and the subsequent effect of poor growth on motor development (Abubakar et al, 2008b). Due to the fact that muscle strength is decreased in children infected with HIV (Blanchette et al, 2001; Pearson et al, 2000), gross motor function may also be affected. Again this highlights the fact that this population is at great risk for gross motor delay, and the Infant Gross Motor Screening Test must therefore have excellent classification accuracy in order to identify these potential problems.

9.2.3 Selection of Items

As discussed in chapter two, a review of the literature on the development of screening tools revealed that item selection is based on the following:

- Review of the literature (Bart et al, 2010; Kirby et al, 2010; Martin et al, 2009; Bayley, 2006c; Harris et al, 2003; Beckung, 2000)
- Review of similar screening tools (Kirby et al, 2010; Harris et al, 2003)
- Items taken from an existing assessment tool (Gladstone et al, 2008; Bayley, 2006c; Aylward, 1995)
- Interviews with people who have the relevant condition (Bart et al, 2010)
- Clinical experience (Bayley, 2006c; Harris et al, 2003; Beckung, 2000)

Gladstone et al (2008) adapted the Denver-II for use in Malawi, and found that many items from Western tools can work well when adapted and translated for other settings. This is due to the fact that they have already been rigorously tested for reliability and validity. They found that when creating new items, these items were less reliable, and had poorer goodness of fit in logistic regression analysis (Gladstone et al, 2008). Due to the fact that gross motor function is universal, and all children develop the same milestones and components of movement, there is also a limited pool of gross motor assessment items. When examining the gross motor components of assessment tools, there are very similar items used in each one (Bayley, 2006c; Harris et al, 2003; Aylward, 1995; Bricker et al, 1995; Frankenberg et al, 1992; Piper et al, 1992). In this study, item selection was based on items from the BSID III. These items were selected according to how well they distinguished between children who were At-Risk, children with Emerging skills and children who were Competent. Items had undergone rigorous reliability and validity testing (Bayley, 2006c), and were

therefore more likely to be robust in use (Gladstone et al, 2008).

9.2.4 Scoring Outcomes of the Infant Gross Motor Screening Test

The first version of the Infant Gross Motor Screening Test had three scoring outcomes, namely At-Risk, Emerging and Competent, so that the results could be compared with those obtained on the BSID III. Aylward (2000) states that the moderate level of risk was developed to include infants whose developmental level is in the 'grey zone' at the time of testing, and who would require, at minimum, developmental surveillance. Aylward (2000) states that three levels of risk preclude the determination of sensitivity or specificity or estimations of relative risk based on odds ratios. However, three levels of risk can help determine which infants need to be monitored (infants obtaining a score in the Emerging group) and which infants should be enrolled in intervention programs (infants in the At-Risk group). However, in favour of scoring with only two levels of risk, Aylward (2000) states that caregivers and parents are more likely to understand a two-level outcome of either At-Risk or Competent. Sensitivity and specificity are also able to be calculated with only two levels of risk.

Aywlard et al (2000) found that the moderate-risk group had the most variation in risk consistency over the first three years and that the trend in infants falling into the moderate-risk groups was towards improvement, which may be related to the positive effects of the intervention which many of them received. Longitudinally, infants in the moderate-risk group were found to be more similar to their low-risk counterparts than to those with a high-risk classification (Aylward, 2000). In South Africa, the trend of HIV infected infants in the moderate-risk group may be towards decline, and increased similarity to high-risk counterparts as intervention is not given, and HIV infected infants have many other risk factors including socioeconomic status, the effects of HIV on neurodevelopment, and repeated hospital admissions with long stays which may lead to regression in development (Cooper et al, 2004; Fiser et al, 2000). In a longitudinal randomised control trial conducted in South Africa, where half the children were given a home stimulation programme, although their developmental scores improved over time, at the end of a year's intervention, their developmental scores were still well below normal levels (Potterton et al, 2009a).

Given that intervention may be more beneficial to infants with less severe delays, identification of moderate-risk infants, and provision of intervention may move these infants to a more optimal level of function (Aylward, 2000). This may be possible in well-resourced developed countries where early intervention may be provided to all infants, even those who are not at high-risk, but this would not be feasible in a developing country such as South Africa, where all therapy departments are understaffed and cannot manage with the high case loads they have. Therefore, the value of having an "Emerging" category in a developing country may not be as great, as there are no adequate resources for these children, and only those At-Risk would be considered to be eligible for intervention.

9.2.4.1 Conversion of the New Screening Tool from Three Scoring Outcomes to Two Scoring Outcomes

Aylward et al (2000) converted the BINS into two scoring outcomes, and tested the predictive validity of the screening tool against the version with three scoring outcomes. When converted to two scoring outcomes, the BINS produced similar results to the version with three outcomes – an early classification of high-risk was indicative of risk status at 36 months, and therefore changing the scoring categories did not change the developmental predictory outcome (Aylward, 2000). This indicates that having two scoring outcomes is not detrimental to the tool's discriminatory ability. Therefore in this study, the new screening tool was also converted to two scoring outcomes for the following reasons:

- The results are easier for parents and health professionals to understand
- The moderate risk group is not lost to follow-up, and there is a clear management plan for those in the At-Risk category, and those in the Competent category
- Sensitivity and specificity can be calculated.

In converting the BINS to a two-outcome measure, the following procedure was undertaken (Aylward et al, 2000): the moderate-risk group was subdivided, based on the cut-off score in the BINS manual that offered the best measures of sensitivity and specificity. Moderate-risk scores falling between the cut-off point and the BINS high-risk category were combined into a moderate-/high-risk group, and named 'High Risk'. Moderate-risk scores falling between the cut-off point and the BINS low-risk category

were combined with the low-risk category to form a 'Low Risk' group (Aylward et al, 2000).

A similar procedure was undertaken in this study but because there were no measures of sensitivity and specificity as there were in the BINS manual, the BSID III scaled scores which translated into the 'Emerging' category were dealt with as follows:

- Scaled scores of 5 and 6 were classified as "At-Risk"
- Scaled scores of 7 were classified as "Satisfactory"

The word "Satisfactory" was chosen due to the wide range of normal development that is possible in children between 6 and 18 months. The conversion of the "Emerging' scores was done through thorough item analysis and the use of clinical judgement to determine whether a child who could not perform the items that they received a 'no credit' for on the BSID III Gross Motor scale would be considered to be developmentally Satisfactory, or At-Risk for developmental problems. On the Infant Gross Motor Screening Test, the Emerging category was always converted to 'At-Risk'. This is due to the fact that HIV infected infants are at increased risk for developmental delays, and even though there may be an improvement in development over time, their scores are likely to be well below average (Potterton et al, 2009a).

9.3 Statistical Properties of the New Screening Tool

It is essential to establish whether a study will provide evidence of the validity of a measuring instrument and its reliability. A high level of reliability gives no firm evidence that the instrument is measuring what it is supposed to, and therefore the validity also needs to be examined. It is equally important to consider the clinical context to which any findings are likely to be applied (Sim and Arnell, 1993). Screening tools should be both valid and reliable. The definitions of the types of validity and reliability tests have been outlined in chapter two. The statistical properties of the Infant Gross Motor Screening Test with regards to validity and reliability are discussed below.

9.3.1 Validity

A measure should be valid for the purpose for which it was developed (Rosenbaum, 1998). The Infant Gross Motor Screening Test was developed in order to identify gross motor delays in HIV positive infants at the time of screening. Testing of the validity of the Infant Gross Motor Screening Test was undertaken to determine the content validity and criterion validity of the Infant Gross Motor Screening Test. Rothstein et al (1991) in the Standards for Tests and Measurements in Physical Therapy Practice states that content validity and criterion-related validity should be evaluated when developing a new tool.

9.3.1.1 Content Validity

Content validity is an essential step in the development of new measuring tools as it represents a starting mechanism for linking abstract concepts with observable and measurable indicators (Wynd et al, 2003). Content validity is often the first step in examining the validity of a newly developed tool (McEwan et al, 2003). This may be done through consultation with a panel of experts (Harris and Daniels, 1996) to obtain consensus on each item, as well as the consensus on the scale as a whole. When assessing content validity, it is necessary to define the domain of content of the concept being measured (in this case, gross motor function) and then to determine whether this is adequately covered by the instrument.

Content validity is largely a matter of judgement, involving two distinct phases: the scale developer should enhance content validity through careful conceptualisation prior to item generation, and the relevance of the scale's content should be assessed through expert assessment (Beck and Gable, 2001; Mastaglia, et al, 2003; Polit and Beck, 2006). During the development of the Infant Gross Motor Screening Test, items based on the BSID III were selected due to the fact that they had undergone validity testing, and came from the Gross Motor facet of the BSID III, indicating that they were developed to test that construct (Bayley, 2006b). This gives evidence for conceptualisation prior to item generation.

Expert assessment may be carried out in a number of different ways. A modified form of the Nominal Group Technique (NGT) was used to determine content validity of the new screening tool (Potter et al, 2004). The NGT has been shown to be a reliable

method to establish the content validity of new instruments (Hyrkas et al, 2003). This was successfully used during the content validation of the Gross Motor Function Classification System (GMFCS) (Palisano et al, 2008), and provides immediate feedback to the participants, thus increasing satisfaction with participation (Potter et al, 2004).

In this study, as discussed in chapter seven, a panel consisting of seven experts was used, and experts were asked to examine and provide feedback on each of the following questions:

- Appropriateness of each item selected (item itself, and item wording)
- Appropriateness of the age groups which had been selected
- Appropriateness of the items for the age-group in which they fell
- Appropriateness of scoring

The NGT protocol, which is described in chapter two was followed. All concerns from participants were discussed and appropriate changes were suggested. There was 100% agreement for objectives two, three and four, and 71.4% agreement for objective one. The changes suggested and made are outlined in chapter seven. Once the concerns and suggested changes had been outlined, there was 100% agreement from all the participants for all objectives. This meets the defined percentage of consensus (80% of participants) as outlined by Palisano et al (2008). The NGT (Potter et al, 2004) has been successfully used to obtain consensus during content validity of new tools (Dyrbye et al, 2010; Palisano et al, 2008; Shea et al, 2007; van der Camp et al, 2006), and was used successfully in this study in order to obtain consensus during content validity.

9.3.1.2 Criterion Validity

The American Physical Therapy Association's Task Force on Standards for Measurement in Physical Therapy distinguishes three varieties of criterion-related validity: (1) concurrent validity, (2) predictive validity, and (3) prescriptive validity (Rothstein et al, 1991). Criterion-related validity underpins quantitative research and is obtained by comparing the measurements obtained on the new instrument with a measurable criterion that is accepted as a standard indicator of a concept or variable (Rothstein et al, 1991). If the instrument gives an accurate representation of the

concept or variable, criterion-related validity has been demonstrated (Sim and Arnell, 1993). In this study, criterion validity was evaluated using the BSID III, which has long been considered the 'gold standard' in infant assessment (Harris et al, 2005; Tieman et al, 2005).

9.3.1.3 Concurrent Validity

Concurrent validity is the degree to which outcomes on one test correlate with outcomes on a criterion test, when both tests are given at the same time (Portney and Watkins, 2000). This is obtained by comparing the measurement being validated with another validated measurement at approximately the same time. Concurrent validity should be assessed when the new test claims to be able to give information about the current status of a person at the time that the measurements are obtained (Rothstein et al, 1991). Correlations of 0.6 or greater should be obtained (Glascoe, 1991; American Psychological Association, 1985). In this study, concurrent validity of the Infant Gross Motor Screening Test was assessed against the Gross Motor facet of the Bayley Scales of Infant Development III. This was done using two versions of the Infant Gross Motor Screening Test: the first had three scoring categories, and the second version had two scoring categories.

Concurrent Validity of the Infant Gross Motor Screening Test with Three Scoring Outcomes

For the version with three scoring outcomes, the agreement was analysed using a Kappa Statistic and was K=0.63 for the total group, which indicates that there is moderate agreement between the tools (Landis and Koch, 1977). The agreement ranged from K=0.51 to K=0.69 per age group, again, indicating only moderate agreement (Landis and Koch, 1977). Some of these values were below the recommended cut-off point by Glascoe (1991), and the American Psychological Association (1985), and due to having three categories, sensitivity, specificity and predictive value could not be calculated (Aylward, 1995). A Kappa value of 0.75 or above which indicates excellent agreement (Landis and Koch, 1977) would be acceptable in order to indicate that the Infant Gross Motor Screening Test is suitable for clinical use.

The Infant Gross Motor Screening Test assessed the children more conservatively

than the BSID III, and therefore where there were discrepancies in scoring, the Infant Gross Motor Screening Test score was always within one category of the BSID III, and was always one score category below. For example, for a child whose scores differed on the two tools: where a score on the BSID III was in the competent category, the Infant Gross Motor Screening Test score would place the child in the Emerging category. All scores between the tests which did not correlate were within one category of the other test's score, and there were no infants who were placed in the At-Risk category on one measure and in the Competent category on the other measure. The Emerging category was the category which was most often involved in the scores which did not correlate between the BSID III and the Infant Gross Motor Screening Test, indicating that this was a problematic category on both tests.

Concurrent Validity of the Infant Gross Motor Screening Test with Two Scoring Outcomes

After consultation with staff at the Harriet Shezi clinic, and examination of the literature, it was determined that a screening tool with two scoring outcomes would be more suitable in this setting, as the management of those with Emerging skills is difficult in a poorly resourced setting. Therefore, both scales were converted to two scoring categories: At-Risk, and Competent, as described in chapter eight (8.3). Again, concurrent validity was assessed between the Infant Gross Motor Screening Test and the BSID III, this time with two scoring categories.

The concurrent validity for the tools with two scoring categories was much higher – a Kappa value of 0.85 was obtained overall, which indicates excellent agreement (Landis and Koch, 1977). The Kappa scores per age group ranged between K=0.60 to K=1. All these scores are above the recommended cut-off point of 0.60 (Glascoe, 1991; American Psychological Association, 1985). The outlying scores were examined to determine why they did not agree. Only four out of sixty cases did not agree, and in three of those cases, the BSID III classified the child in the Competent category, and the Infant Gross Motor Screening Test classified them as At-Risk. There was only one case where the Infant Gross Motor Screening Test classified the child as being Competent, where the BSID III score was At-Risk. Although this may tend towards over-referrals on the part of the Infant Gross Motor Screening Test, the population is at extremely high risk of gross motor delay (Potterton et al, 2009b; Ferguson and

Jelsma, 2009; Baillieu and Potterton, 2008), and therefore in terms of further management, it would be better for the child to be referred for further assessment based on an over-referral, than wait three to six months for a further screening at the next clinic appointment, and then be classified as At-Risk, especially for those with progressive encephalopathy.

Glascoe (2001) found that children who receive false-positive results in screening tests perform substantially lower than children with true-negative scores on measures of intelligence, language, and academic achievement, which are the three best predictors of school achievement. These children also carried more psychosocial risk factors, such as limited parental education. Thus, she concluded that children with false-positive screening results are an At-Risk group for whom diagnostic testing may not be an unnecessary expense but rather a beneficial and needed service that may help focus intervention efforts (Glascoe, 2001).

The population for which the Infant Gross Motor Screening Test was developed has been shown to have motor, cognitive and language developmental delays in studies carried out worldwide (Potterton et al, 2009b; Baillieu and Potterton, 2008; Udgirkar et al, 2003; Blanchette et al, 2001; Pearson et al, 2000; Fragoso et al, 1999; Drotar et al, 1997; Wolters et al, 1997; Belman et al, 1996; Pollack et al, 1996; Chase et al, 1995; Nozyce et al, 1994; Belman, 1992; Belman et al, 1988; Epstein et al, 1986; Belman et al, 1985; Ultmann et al, 1985), as well as socioeconomic risk-factors and are therefore at risk for lower scores on all domains of developmental testing. In keeping with Glascoe's (2001) conclusions, again, for those infants who receive false positive results on the new screening test diagnostic testing may be beneficial.

The results of the concurrent validity of the Infant Gross Motor Screening Test compared with other commonly used screening tests are shown in Table 9.1 below:

Table 9.1 Concurrent Validity of the Infant Gross Motor Screening Test Compared with Other Commonly Used Screening Tools

Tool	Concurrent validity
BINS	With BSID II: Motor scale r = 0.35 Mental scale r = 0.40
ASQ	Only sensitivity, specificity and PPV available – no concurrent validity for second version
Denver II	Only sensitivity, specificity and PPV available – no concurrent validity for second version
AIMS	BSID II motor scale r = -0.98 Peabody Developmental Motor Scale r = -0.97
HINT	BSID II motor scale r = -0.89 BSID II mental scale r = -0.73
Bayley III Screening Test	Against the BSID III At-Risk: moderate (41.82% - 65.91%) Emerging: accurate (63.87% - 77.78%) Competent: very accurate (83.84% - 92.11%)
Infant Gross Motor Screening Test	Against the BSID III Gross Motor facet K = 0.85

The Infant Gross Motor Screening Test fares well in terms of other commonly used screening tests, with concurrent validity of K=0.85 against the BSID III Gross Motor facet. The BINS (Aylward, 1995) has low concurrent validity with the BSID II on both the Motor (r=0.35) and Mental scale (r=0.40), even though the BINS was based on the BSID II, and items were taken directly from it. The problem with the concurrent validity could stem from the fact that new items were added along with the items taken from the BSID II, and this has been shown to affect validity (Gladstone et al, 2008).

The ASQ (Bricker et al, 1995), and the Denver II (Frankenberg et al, 1992) have both been revised, but concurrent validity for the revised, and now more commonly used versions has not been carried out. Only sensitivity, specificity and positive predictive values are available for these versions. Concurrent validity of the AIMS (Piper et al, 1992), against the BSID II Motor scale (r = -0.98), and Peabody Developmental Motor Scale (r=-0.97) was very high which is most likely due to the fact that the AIMS only

measures motor development, and that there are a limited number of motor assessment items, which are common across most motor assessment tools. The concurrent validity of the Infant Gross Motor Screening Test is not as high, but it has fewer items than the AIMS, and therefore less chance of correlating on all items.

Concurrent validity of the HINT (Harris et al, 2003) against the BSID II Motor scale (r = -0.89) and BSID II Mental scale (r=-0.73) was also high, and the authors state that the HINT is predominantly a neuromotor screening test, which is why the correlation with the BSID II Motor scale was higher (Harris et al, 2003). Concurrrent validity of the Infant Gross Motor Screening Test is similar to that of the HINT.

The Bayley-III Screening Test used percentages of agreement against the BSID III as follows:
At-Risk: moderate (41.8%-65.9%)
Emerging: accurate (63.9%-77.8%)
Competent: very accurate (83.8%-92.1%)
The use of percentages is not recommended, as it will often overestimate true reliability because it fails to account for chance agreement (Portney and Watkins, 2000).

Therefore, against the seven other screening tests, the Infant Gross Motor Screening Test's concurrent validity is excellent, considering that for some of the commonly used tests such as the Denver-II and ASQ (2nd version), concurrent validity studies have not yet been conducted.

9.3.1.4 Sensitivity, Specificity, Positive Predictive Value and Negative Predictive Value

Screening tests must offer information on sensitivity, which is how well they detect the people who have the condition being screened for, and the specificity of the screening tool which is the ability to accurately rule out those who do not have the problem. The significance of these statistical considerations is both economic and clinical, as over-referrals lead to further, often more expensive and time consuming, in-depth assessments and may be upsetting to parents (Rosenbaum, 1998).

Sensitivity and specificity of developmental screening tools are measured by comparing the test results to those of gold-standard developmental evaluation tools (Glascoe, 1997). In this study, this was done against the BSID III (Bayley, 2006a), as the Bayley Scales of Infant Development have long been considered the 'gold standard' of infant assessment (Harris et al, 2005; Tieman et al, 2005).

Sensitivity is a measure of validity of a screening procedure based on the probability that someone who has the condition will test positive (Portney and Watkins, 2000). The overall sensitivity for the Infant Gross Motor Screening Test was 97.4%. For each age group, the sensitivity was tested, and the following results obtained: 6-8 months: 85.7%, 9-12 months: 100%, 13-16 months: 100% and 17-18 months: 100%. Glascoe (1997) recommends that a screening test should have sensitivity levels of 70 – 80% due to the nature and complexity of measuring the continuous process of child development. Meisels (1989) suggested that more conservative criteria should be used with sensitivity levels of no less than 80% in order to avoid too many over-referrals. A value of 90% is recommended by the American Psychological Association (1985). Therefore the overall sensitivity of the Infant Gross Motor Screening Test, as well as the sensitivity per age group is over the suggested figures, which indicates that children who have gross motor delays will be identified by the screening tool.

Specificity is a measure of validity of a screening procedure based on the probability that someone who does not have the condition will test negative (Portney and Watkins, 2000). There often is a trade-off in specificity if the sensitivity of a measure is very high. Again, Glascoe (1997) recommends that a screening test should have specificity levels of 70-80% due to the nature and complexity of measuring the continuous process of child development. Meisels (1989) suggested that more conservative criteria should be used with specifity levels of no less than 80% to avoid under-referrals. The recommendation by the American Psychological Association, (1985) is 90%. This indicates that there is a wide variation in acceptable levels of specificity, but when combining the above recommendations, it should be between 70 -90%. The overall specificity for the Infant Gross Motor Screening Test was 85.7%. For each age group it was as follows: 6-8 months: 87.5%, 9-12 months: 100%, 13-16 months: 100% and 17-18 months: 50%. Overall, the Infant Gross Motor Screening Test has excellent levels of specificity apart from the 17-18 month age group where

only four children were Competent, and the new screening tool correctly identified two out of the four. The levels achieved overall are acceptable according to suggested levels of between 70 and 90% (Glascoe, 1997).

Screening tests need to give information as to the probability that the test will give the correct diagnosis (Altman and Bland, 1994). This information cannot be obtained from sensitivity and specificity values, and therefore the predictive values need to be examined in order to obtain this information (Altman and Bland, 1994). Positive Predictive Value (PPV) is an estimate of the likelihood that a person who tests positive actually has the condition (Portney and Watkins, 2000). The lower the prevalence of a disorder, however, the lower will be the positive and negative predictive values. A positive predictive value of 50% would be acceptable in developmental screening, although 30% and above is often found, because some of the low average children will be over-referrals (Aylward, 1997). The positive predictive value of the Infant Gross Motor Screening Test is 92.7% and for each age-group 6-8 months 85.7%, 9-12 months 100%, 13-16 months 100% and 17-18 months 84.6%. This indicates that the PPV value in the Infant Gross Motor Screening Test is high, and that the likelihood that the children who are diagnosed as At-Risk on the new screening tool will have gross motor delay. This value is very high, due to the high prevalence of gross motor delay in this population.

Negative predictive value (NPV) is the proportion of patients who are correctly diagnosed by the test as not having the condition (Altman and Bland, 1994). This result is less commonly presented, and ideal values for this result could not be found. The negative predictive values are: Overall: 94.7%. Per age group: 6-8 months: 87.5%, 9-11 months: 100%, 13-16 months: 100 % and 17-18 months: 100%. The negative predictive values for the Infant Gross Motor Screening Test are also very high, indicating that overall, the proportion of children who were correctly diagnosed as having no gross motor delay, did not have the condition. The prevalence of the condition is the probability before the test is carried out that the subject has the condition (this is known as the prior probability of the disease) (Altman and Bland, 1994). The positive and negative predictive values act as revised estimates of the same probability for the patients, and are known as posterior probabilities (Altman and Bland, 1994). The difference between the prior and posterior probabilities is one way

of assessing the usefulness of the test (Altman and Bland, 1994). Based on previous studies where it was found that up to 95% of the children in a sample had gross motor delay (Baillieu and Potterton, 2008), the Infant Gross Motor Screening Test is suitable for use in this population, as it detects this high prevalence. The sensitivity, specificity and PPV of the Infant Gross Motor Screening Test are compared with those of commonly used screening tests in Table 9.2 below:

Table 9.2 Sensitivity and Specificity of the Infant Gross Motor Screening Test Compared with Other Commonly Used Screening Tools

Tool	Sensitivity, specificity and PPV
BINS	Sensitivity: 75-86% Specificity: 75-86% PPV: 0-52%
ASQ	Sensitivity: 75% Specificity: 85% PPV 46%
Denver II	Suspect w/normal group: Sensitivity=56% Specificity=80% PPV= 37% Suspect w/abnormal group: Sensitivity=83% Specificity=43% PPV=23%
AIMS	Not available
HINT	Testing currently underway
Bayley III Screening Test	Not available
Infant Gross Motor Screening Test	Sensitivity: 97.4% Specificity: 85.7% PPV: 92.7% NPV: 94.7%

The sensitivity of the Infant Gross Motor Screening Test is 97.4%, which is higher than any of the other commonly used screening tests as shown in the table above. The specificity of the Infant Gross Motor Screening Tool is 85.7%, which is similar to that of the ASQ and the BINS. The combination of sensitivity and specificity that the Infant Gross Motor Screening Tool has achieved is excellent, as in many cases (such as the Denver II), there is a trade-off between the two if one is very high. The ASQ and the BINS have also achieved a good combination of sensitivity and specificity.

The PPV of the Infant Gross Motor Screening Test is 92.7%, which is extremely high, due to the high prevalence of gross motor delay in the HIV infected infant population for which the tool was developed. The ASQ has a much lower PPV, due to the samples on which the validity studies were conducted, where the prevalence of motor delay was lower. The Denver-II and BINS have achieved particularly low PPV, even in high-risk populations, which should be investigated further (Hess et al, 2004; Glascoe and Burne, 1993). The BINS sample was at environmental risk (families with low income) rather than biological risk, which may contribute to the low PPV, and therefore the BINS may be better suited for use on a population with biological risk, and better-defined developmental problems (Hess et al, 2004).

9.3.1.5 Predictive Validity

Predictive validity examines the justification of using a measurement to say something about future events or conditions (Rothstein et al, 1991). Information about predictive validity should be provided when a measurement claims to be able to provide information about future status (Rothstein et al, 1991). Marks et al (2008) express concern at conducting predictive validity studies on developmental screening tools, as by design, screening tools are developed to provide information on the current status of a child, and not give an indication of future outcome. The American Academy of Pediatrics (2006) states that a screening tool should generate referrals to early intervention, where the child may receive more in-depth assessment, or monitoring through surveillance. Therefore it is not necessary to evaluate predictive validity at this time, as this screening instrument was not developed to predict later functional status of the child, and will be used to assess the developmental status of the child at a particular point in time, based on the score of the screening tool achieved. The Infant Gross Motor Screening Test may be used for surveillance purposes, that is, it may be administered at each clinic visit in order to monitor development.

9.3.1.6 Prescriptive Validity

Prescriptive validity refers to the validity of a test when the results are used to determine what treatment the person will receive (Rothstein et al, 1991). Information regarding prescriptive validity should be provided when the results of a measurement will be used to determine treatment options. Therefore, due to the fact that at present

rehabilitation options are not based on the results of this screening test, it is not necessary to evaluate prescriptive validity at this point.

9.3.2 Reliability

Measures should be reliable, which means that they give consistent results from one administration time to another, and from one assessor to another, when the person being assessed has not changed (Rosenbaum, 1998). The instrument might be reliable but untrained assessors may not be capable users of the instrument. Therefore, before deciding to use any clinical assessment instrument, would-be users should obtain appropriate training in both the application and interpretation of the measures they wish to apply (Rosenbaum, 1998). Reliability is defined as the consistency or repeatability of measurements (Rothstein et al, 1991). According to Rothstein et al (1991) the following forms of reliability should be considered when developing a test:

9.3.2.1 Internal Consistency

Internal consistency is defined as the extent to which items or elements that contribute to a measurement reflect one basic phenomenon or dimension. The Infant Gross Motor Screening Test only assesses gross motor function, and items were based on those from the BSID III Gross Motor scale, on which internal consistency had been conducted (r=0.92) (Bayley, 2006b).

9.3.2.2 Interrater Reliability

Interrater reliability is the consistency of measurements when more than one person takes the measurements at the same time, and indicates the agreement of measurements taken by different examiners (Rothstein et al, 1991). Interrater reliability is best measured when all raters are able to measure a response during a single trial (Portney and Watkins, 2000). If interrater reliability has not been established, it cannot be assumed that other raters would have obtained the same results, which therefore limits the application of the findings (Portney and Watkins, 2000). Studies should be conducted in a clinical context consistent with the intended use of the measurements (Rothstein et al, 1991).

In this study two raters evaluated preliminary interrater reliability, and did so during a single trial as suggested by Portney and Watkins (2000). The child's response was scored by both raters simultaneously. The context in which interrater reliability was conducted was consistent with the intended use of the measurements in that it was conducted at Harriet Shezi Children's clinic, which is a paediatric HIV clinic. Children on which the reliability testing was conducted were all HIV positive, and attending the clinic for scheduled visits. The interrater reliability was excellent (r=1) (Portney and Watkins, 2000), and indicates that different raters would obtain the same score when using the Infant Gross Motor Screening Test. The scoring on the Infant Gross Motor Screening Test is simple, and the outcomes to be assessed are clear, which contributes to the high interrater reliability.

9.3.2.3 Test-retest Reliability

Test-retest reliability is the consistence of repeated measurements over time where the subject has not changed (Rothstein et al, 1991). Within-day and between-day studies should be conducted in a clinical context consistent with the intended use of the measurements (Rothstein et al, 1991). In the reviewed literature on developmental screening and performance tests, the time between testing for test-retest reliability differed significantly, and ranged from two days to thirty days (Thomas-Stonell et al, 2010; Bayley, 2006b; Bayley 2006c; Saigal et al, 2005; Harris and Daniels, 2001; Squires et al, 1997; Coster, 1995; Frankenberg et al, 1992).

In this study, preliminary test-retest reliability was carried out in the wards at Rahima Moosa Hospital. The context in which test-retest reliability was carried out was consistent with the intended use of the measurement instrument as suggested by Rothstein et al, (1991), was achieved by that fact that the children assessed were HIV positive, and the future use for the tool would be in both in-patient and outpatient paediatric HIV settings. Portney and Watkins (2000) state that the interval between tests should be considered very carefully: they should be far enough apart to avoid fatigue, learning or memory effects, but close enough to avoid changes in the variable to be measured. They state that the primary criteria for choosing an appropriate interval are the stability of the response variable, and the purpose of the test. In addition, they also state that measurements of infant development might need to be taken over a short period to avoid the rapidly occurring changes of early infancy. In

this study, the interval of two days between tests was chosen for the following reasons: it was far enough apart to avoid fatigue, but close enough to avoid changes in the patient, as in this case, the patients were in hospital, and were subject to change due to improvements in their condition. Care was taken to assess children at the same time of the day, so that the effects of medication could be taken into consideration.

The results of the test-retest reliability study were excellent (r=0.98) (Portney and Watkins, 2000), and indicate that the Infant Gross Motor Screening Test is stable over time, and produces the same test results. In the two cases in which the score did not agree, the score on the first assessment was one point higher, which may be due to the fact that the child's medical condition had deteriorated slightly, therefore producing a lower score on the second assessment. The category in which the child was placed by the Infant Gross Motor Screening Test was the same, despite the one point difference in scores, indicating that the child would have received the same management following administration of the Infant Gross Motor Screening Test.

Portney and Watkins (2000) state that caution should be given to carry-over effects, which may occur with repeated measurements, and concepts such as motor learning, which may result in an improvement on the first test score. The Infant Gross Motor Screening Test is an observational test, and therefore motor learning would not occur, as there is no facilitation or handling of the infant.

9.3.2.4 Intrarater Reliability

Intrarater reliability refers to the stability of data recorded by one individual over two or more tests. This is usually assessed using trials with short intervals in-between where carryover effects are not an issue (Portney and Watkins, 2000). In a test-retest situation, when a rater's skill is relevant to the accuracy of the test, intrarater and test-retest reliability are essentially the same estimate (Portney and Watkins, 2000). In this study, the rater's skill was relevant to the accuracy of the test, and therefore test-retest reliability and intrarater reliability are the same estimate. In this study, the intrarater reliability was excellent (r=0.98) (Portney and Watkins, 2000), and indicates that the test shows good stability of data recorded by one individual across two trials (Portney and Watkins, 2000).

9.3.2.5 Reliability of the Infant Gross Motor Screening Test in Relation to other Screening Tests

The results are shown in relation to the reliability values of other commonly used screening tools in Table 9.3 below:

Table 9.3 Reliability of the Infant Gross Motor Screening Test Compared with other Commonly Used Screening Tools

Tool	Interrater reliability	Test-retest reliability
BINS	.79 to .96	.71 to .84
ASQ	94% (SEM = 0.12)	94%
Denver II	141 items had 100% agreement, 7 had 90 – 99% and one had 83%.	Excellent agreement for 59% of items ($k \geq 0.75$), 23% of items: fair to good agreement ($k \geq 0.40$).
AIMS	0.99.	0.98.
HINT	0.996 and above	0.86 to 0.99.3
Bayley III Screening Test	Not described	0.80-0.83
Infant Gross Motor Screening Test	r=1	r=0.98

The Infant Gross Motor Screening Test showed excellent results in comparison to other well-known screening tests. The results obtained are similar to those of the HINT and the AIMS, which have been shown to be reliable screening tests (Harris and Daniels, 2001; Coster, 1995).

9.4 Clinical Findings in Relation to the Sample

In the process of developing the Infant Gross Motor Screening Test, a lot of clinical data were collected, which have important clinical implications for the management of paediatric HIV in South Africa.

9.4.1 Developmental Outcomes of the Sample

HIV is a neurotropic disease, and has been shown to affect all aspects of

development (Potterton et al, 2009b; Baillieu and Potterton, 2008; Udgirkar et al, 2003; Blanchette et al, 2001; Pearson et al, 2000; Fragoso et al, 1999; Drotar et al, 1997; Wolters et al, 1997; Belman et al, 1996; Pollack et al, 1996; Chase et al, 1995; Nozyce et al, 1994; Belman, 1992; Belman et al, 1988; Epstein et al, 1986; Belman et al, 1985; Ultmann et al, 1985). Gay et al (1995) and Msellati et al (1993) reported an increase in prevalence of developmental delay with increasing age in the first 18 months of life, which has been demonstrated in this study. The extent of delay in language, cognitive and motor function will be discussed in the following section:

9.4.1.1 Language Development

HIV CNS disease in children is associated with deficits in both receptive and expressive language, although expressive language skills are more severely impaired (McClowry, 2000; Blanchette et al, 2001; Tardieu et al, 1995; Wolters et al, 1995; Nozyce et al, 1994; Papola et al, 1994; Pizzo et al, 1988; Epstein et al, 1986). Language is affected later in life than gross motor development and is thought to be a reflection of both the chronicity of the disease as well as direct CNS involvement (Bisiacchi et al, 2000; Msellati et al, 1993).

The mean Language composite score for the total group was 88 (±17.68), and individual age groups ranging from 79 (±14.19) in the 17-18 month age group, to 95 (±18.68) in the 6-8 month age group, and the Language composite scores deteriorated with increasing age in the first 18 months of life, which is similar to findings by Msellati et al (1993) and Gay et al (1995). Just over 41% of infants obtained Language composite scores in the Average range. Just over 17% obtained scores in the Low Average range, 16.1% in the Borderline range, and 15.2% in the Extremely Low range, which indicates that the prevalence of language delay in this sample is 49.1%. Just over eight percent of infants obtained scores in the High Average range, and just over one percent obtained scores in the Superior range. When analysing the scaled scores obtained in this study in order to differentiate between Expressive and Receptive Communication function, it was found that Expressive Communication was slightly more affected than Receptive Communication, which is in keeping with previous literature (Baillieu and Potterton, 2008; Blanchette et al, 2001; McClowry, 2000; Tardieu et al, 1995; Wolters et al, 1995; Nozyce et al, 1994; Papola, 1994; Pizzo et al, 1988; Epstein et al, 1986). In

comparison with the BSID III standardisation sample (Bayley, 2006b), where 50% of infants obtained scores in the Average range, 16.1% of infants obtained scores in the Low Average range, just over six percent obtained scores in the Borderline range, and just over two percent fell in the Extremely Low range (Bayley, 2006b), there were more infants in this sample who fell into the Low Average, Borderline and Extremely Low ranges. There were also far fewer in the High Average range. In comparison with a previous study done in South Africa, the prevalence of language problems is lower, as the previous study found that 82.5% of infants had language delay (Baillieu and Potterton, 2008). It is encouraging to note that the prevalence of language delay is lower in this study compared with previous studies (Baillieu and Potterton, 2008), which is most likely due to the fact that children now have access to HAART, which may result in slightly higher scores on developmental tests due to less rapid developmental deterioration.

Verbal expression is highly correlated to motor function, thus motor deficits, including those in muscle coordination and motor programming may affect oral-motor skills and contribute to feeding problems, articulation errors and speech difficulties (Wolters et al, 1997). The fact that motor developmental delay has a lower prevalence in this study compared with previous studies conducted on similar populations may account for the lower prevalence of language delays. Chronic otitis media is extremely common in children who are HIV positive and while this may explain language delay due to impaired hearing to some extent, it cannot account for all the deficits in speech and language experienced by these children (Layton and Scott, 2000). Children who experience an AIDS-defining illness early in life score significantly lower on verbal, perceptual performance, quantitative abilities, and memory subtests than infected children with no AIDS-defining illness or non-infected children (Abubakar et al, 2008a). This has implications for children reaching school-age, and the support they may need before reaching school, as well as extra support at school.

9.4.1.2 Cognitive Development

HIV infected infants have cognitive impairment, CNS dysfunction and deficits in neuropsychological functioning, and often present with microcephaly and mental retardation (McGrath et al, 2006a; Mazzoni et al, 2000; Fragoso et al, 1999; Belman et al, 1996; Henry et al, 1996; Fowler, 1994; Belman, 1992; Diamond et al, 1990;

Belman et al, 1988; Epstein et al, 1986; Belman et al, 1985; Ultmann et al, 1985). It has been found that cognitive deficits increase with increasing age (Drotar et al, 1997). HIV Infected children with brain abnormalities perform worse on measures of cognitive development (Brouwers et al, 1995). Deficits in visual scanning, academic achievement and psychomotor speed have been found (Cohen et al, 1991). HIV infected school-going children have memory and visuopraxic impairments, but in the early stages of infection only minimal impairment of executive functions has been noted (Bisiacchi et al, 2000). Children infected in utero appear to be the most severely affected (Brouwers et al, 1995; De Carli et al, 1993). The cognitive deficits described in HIV positive children exist independently of environmental and social risk factors (McGrath et al, 2006b; Chase et al, 2000; Knight et al, 2000; Henry et al, 1996).

The mean cognitive composite score for the total group (n=112) was 80 (±18.04), which falls into the Low Average range. Individual age groups ranged from 74 (± 15.22) (17-18 months) to 85 (±20.85) (6-8 months), and the cognitive composite scores deteriorated with increasing age, which is similar to findings by Msellati et al (1993) and Gay et al (1995). In a previous study by Potterton et al (2009b), the mean MDI was 65.61 (±21.93) (Severely Delayed), which was far lower than this study. The mean MDI obtained during the validation of the BSID II on a group of 35 HIV infected infants was 79.6 (±19.8), which falls into the Mildly Delayed category (Bayley, 1996). This is similar to the mean classification in this study. The difference between previous and current studies conducted in South Africa may be explained by the fact that the children in the validation sample in the United States all had access to antiretrovirals whereas only 14.5% of the children in Potterton et al's (2009b) study were on HAART, and 69% of infants in this study were on HAART.

Just over thirty percent of infants obtained Cognitive composite scores in the Average range. Just over 18% obtained scores in the Low Average range, 17% in the Borderline range, and 28.6% in the Extremely Low range, which indicates that the prevalence of language delay in this sample is 63.4%. Just over three percent of infants obtained scores in the High Average range, and less than one percent obtained scores in the Superior and Very Superior range.

In comparison with the BSID III standardisation sample, where 50% of infants obtained scores in the Average range, 16.1% of infants obtained scores in the Low Average range, just over six percent obtained scores in the Borderline range, and just over two percent fell in the Extremely Low range (Bayley, 2006b). There were far more infants in this sample who fell into the Low Average, Borderline and Extremely Low ranges. There were also far fewer in the High Average range. In comparison with previous studies done in South Africa, the prevalence of cognitive delay is lower. It was found that 90% of infants had cognitive delay, when no children had access to HAART (Baillieu and Potterton, 2008), and in a recent study, 78% of infants had cognitive delay, when 14. 5% had access to HAART (Potterton et al, 2009b). This high level of developmental delay is in keeping with findings from studies conducted in other parts of the world where prevalence rates of 50% or higher have been documented, but with more children having access to HAART (Pollack et al, 1996; Chase et al, 1995; Belman et al, 1986; Epstein et al, 1986).

9.4.1.3 Motor Development

Young children's motor development, coordination, muscle tone and reflexes are most consistently and strongly affected by HIV infection (Drotar, 1997). The delay is either a reflection of the chronicity of the disease or a direct expression of CNS involvement (Bisiacchi et al, 2000; Msellati et al, 1993). Motor function is compromised early in development in infants with HIV and it is mainly gross motor skills that are delayed (Chase et al, 1995; Nozyce et al, 1994). This may be because gross movements require muscle groups and some degree of physical effort, whereas fine movements are associated with more precise outputs but lower force. Therefore gross motor performance deficits may be related to an overall loss of strength (Parks et al, 1999).

The mean Motor composite score for the total group was 72 (±20.32), and individual age groups ranged from 68 (±15.59) (17-18 months) to 76 (±23.95) (6-8 months), and the Motor composite scores deteriorated with increasing age, which is similar to findings by Msellati et al (1993) and Gay et al (1995). In a previous study by Potterton et al (2009b), the mean PDI was 53.67 (±23.77) (Severely Delayed), which was far lower than this study. The mean PDI obtained by Bayley (1996) when validating the BSID II on a group of 35 HIV infected infants was 81.1 (±27.2), which falls into the

Mildly Delayed category. This is also higher than the mean in this study. Again, the difference may be explained by the differences in access to HAART.

Just over 19 percent of infants obtained motor composite scores in the Average range. Just under 17% obtained scores in the Low Average range, 18.8% in the Borderline range, and 42% in the Extremely Low range, which indicates that the prevalence of language delay in this sample is 77.7%. Just over two percent of infants obtained scores in the High Average range, and no infants obtained scores in the Superior or Very Superior range. When analysing the scaled scores obtained in this study, it was found that Fine Motor function was less affected than Gross Motor function, which is in keeping with previous studies (Baillieu and Potterton, 2008; Parks et al, 1999). Muscle weakness has been associated with poor motor function in children infected with HIV (Blanchette et al, 2002; Pearson et al, 2000) and could be a result of muscle atrophy secondary to malnutrition. Pearson et al (2000) found that decreased muscle bulk was an independent predictor of disease progression in HIV infected infants.

In comparison with the BSID III standardisation sample, where 50% of infants obtained scores in the Average range, 16.1% of infants obtained scores in the Low Average range, just over six percent obtained scores in the Borderline range, and just over two percent fell in the Extremely Low range (Bayley, 2006b), there were far more infants in this sample who fell into the Low Average, Borderline and Extremely Low ranges. There were also far fewer in the High Average range. In comparison with previous studies done in South Africa, the prevalence of Motor delay is lower. It was found that 87.5% of infants had motor delay, when no children had access to HAART (Baillieu and Potterton, 2008), and in a recent study, 87% of infants had motor delay, when 14.5% had access to HAART (Potterton et al, 2009b). This high level of developmental delay is similar to findings from studies conducted in other parts of the world where children have access to HAART (Pollack et al, 1996; Chase et al, 1995; Belman et al, 1986; Epstein et al, 1986).

It has been found that children with HIV infection who experienced an early AIDS-defining illness exhibit significant impairment in their overall cognitive and motor ability, relative to their HIV infected, but relatively healthy, counterparts and non-

infected peers (Abubakar et al, 2008a; Nozyce et al, 1994). Children in sub-Saharan Africa are more likely to experience an early AIDS-defining illness, due to lack of resources, late diagnosis and up until recently, lack of access to HAART. With early, appropriate, antiretroviral therapy, not only does the child stand a better chance of survival, but may also avoid significant developmental impairment due to decreased chances of an early AIDS defining illness. In the meantime, HIV infected children of school-going age may need additional educational and psychosocial resources in the classroom and regular re-evaluation of their progress and learning needs (Smith et al, 2006). The South African Department of Health's policy to provide ARV's to all children less than one year of age from 1st April 2010, will hopefully help this goal to be attained (DOH, 2010). Along with this, regular monitoring of development, and timeous referral to appropriate services is needed in order to detect developmental problems early on.

Previous studies have shown that home stimulation programmes may be effective in helping with the developmental delays seen in these children (Potterton et al, 2009a), and therefore due to overloaded therapy services, this may be a more effective means of management.

Research Implications

- Further research should be conducted to clarify the trends being seen in studies where expressive language is more affected than receptive language.
- Further research should be conducted to evaluate the cognitive problems of school-age children in sub-Saharan Africa.
- Imaging studies should be carried out in order to correlate areas of brain damage with clinical presentations.

Clinical Implications

- The above findings highlight the need for a screening tool which is able to detect developmental issues early on. Gross motor function is consistently most severely affected in HIV infected infants, and therefore a gross motor screening tool would be suitable for this setting, and may serve as a 'net' to identify children with developmental problems. Children who present with gross motor

problems on screening may then be referred for further assessment in the language and cognitive domains, as they are at risk for problems in all areas.

- Children who are HIV positive should be referred for routine hearing tests due to the prevalence of otitis media in this population, which may contribute to language problems.

- Simple home exercises and stimulatory activities should be given to all children in order to help combat some of the developmental delays seen. Due to the high case loads of HIV positive infants in the public sector this may be an effective management strategy, which would also help to empower caregivers.

9.4.2 Immunological Status of the Sample

A lower CD4 count and higher viral RNA load are associated with increased severity of disability, growth failure and slower attainment in milestones in infants and children with vertically transmitted HIV (Pearson et al, 2000; Belman et al, 1996; Pollack et al, 1996; Chase et al, 1995). Children with more advanced degrees of immune suppression early in life and those with high-plasma viral load in infancy also have higher rates of encephalopathy (Nozyce et al, 1994).

The immunologic status of the sample was poor, with the highest mean CD4 percentage being 28% (±0.45) for infants in the 17-18 month age group, and the total mean CD4 percentage being 23% (±9.9). Mean viral load was 1766.4 (±67.16), which is also an indication that the immunological status of the sample is poor, as with the NASBA assay which is used at the clinic, an ideal viral load of 25 is expected to indicate low viral load. This together with the low CD4 percentage of the sample indicates that these infants are at risk for increased severity of disability, growth failure and slower attainment in milestones as found by Pearson et al (2000), Belman et al (1996), Pollack et al (1996) and Chase et al (1995). This is despite the fact that 67.9% of the sample was on HAART, and indicates that their immunologic status and developmental status was very poor. The regimen used at Harriet Shezi is a high CNS penetrating regimen according to Patel et al (2009), and therefore less severe effects on development would be expected.

The children in this sample have a number of risk factors, including poor maternal education, low socioeconomic backgrounds, as well as being HIV infected in a developing country, which has been shown to be associated with more opportunistic infections. Many children in developing countries are exposed to multiple risk factors for poor development including poverty, poor health and nutrition. A conservative estimate is that more than 200 million children under five years of age in developing countries are not developing to their full potential due to poverty and the associated health, nutrition and social factors (Grantham-McGregor et al, 2007; Walker et al, 2007). This, together with the fact that up until April 2010, stringent guidelines for the initiation of ARV's have had to be followed in children under one year of age meant that timeous initiation of HAART did not occur. Adherence to taking medication is also problematic, with stigma, coupled with the need for refrigeration, poor palateability and regularity of doses being major factors (Davies et al, 2008). It has also been found that mothers/caregivers with secondary education were more likely to be successful in giving dosages correctly (Davies et al, 2008). All these factors may have had an influence on the sample in this study, which may partly account for the poor immunologic and virologic status of the sample, and in turn, their poor developmental outcomes.

Research Implications

- As children have better access to ARV's in South Africa, further research should be conducted to determine whether their immunological status improves further, and in turn, what effects this has on neurodevelopment.
- Further research into adherence issues and potential improvement of adherence should continue.

Clinical Implications

- The immunologic and virologic status of the infants still has a huge impact on their neurodevelopmental status, and therefore cannot be assessed in isolation.
- Education should be provided to caregivers as to the global effects of immunologic and virologic status on neurodevelopment, general health, and eventual school function, and the role of the caregiver in providing medication, which will help to facilitate improved development.

9.4.3 Growth

Growth parameters have previously been identified as being markers of disease progression (Bobat et al, 2001) as has developmental delay (Mintz, 1999). It has also been demonstrated that anthropometric status is predictive of developmental status in HIV infected children (Potterton et al, 2009b; Abubakar et al, 2009; Wiznia et al, 1996). There is an association between poor growth and neurodevelopmental delay in HIV positive children especially in children with advanced stages of the disease (Pollack et al, 1996). Infants with the most severe growth delay have significant cognitive and motor delay even if this only becomes apparent later (Pollack et al, 1996). Wiznia et al (1996) report a strong correlation between poor weight-for-age and decreased cognitive function and a low CD4 count. Missmer et al (2000) found that a decreased height-for-age z-score is a strong predictor of decreased functional status, suggesting that height may be used as an indicator of disease severity. HIV infected children in developing countries show a decline in length and weight within the first months of life, and eventually manifest a picture of chronic malnutrition (Tindyebwa et al, 2004). This may result from HIV infection itself, underlying disease such as tuberculosis, inadequate macro/micronutrient intake or a combination of any, or all of the above. A high viral load in children is associated with increased risk of failure to thrive, while infections such as pneumonia, diarrhoea and tuberculosis further exacerbate growth failure (Tindyebwa et al, 2004).

In this study, the mean height-for-age z-score (-2.21 ±1.7) was indicative of moderate stunting, and moderate malnutrition was present in the youngest age group. In a study by Potterton et al (2009b) weight-for-age was found to be positively related to BSID II Mental Developmental Index (MDI) scores. Weight-for-height and height-for-age were related to MDI scores but were not as important as weight for age. This has also been found by Wiznia et al 1996. Potterton et al (2009b) found that weight-for-age z-scores were strongly predictive of BSID II Psychomotor Developmental Index (PDI) scores. This has also been found by Abubakar et al (2009) who showed that weight-for-age and disease stage were predictive of psychomotor outcome.

Stunting is an indication of slow skeletal growth and is often associated with poor socio-economic conditions, chronic or repeated infections as well as inadequate nutrition. Stunting takes a long time to correct once adverse circumstances have been

corrected as skeletal growth is a slower process than growth in body mass (WHO, 1995). Children with growth failure may have decreased weight-for-height (wasting) and decreased height-for-age (stunting). Malnutrition initially causes wasting, while stunting is associated with prolonged malnutrition (Chantry and Moye, 2005). It has commonly been found in HIV positive children that stunting but not wasting is present (Tindyebwa et al, 2004; Nathan et al, 2003; Bobat et al, 2001).

In this study, head circumference z-scores were very poor, and the mean head circumference z-score was -3.01 (±2.2). Head circumference has previously been linked to the extent of developmental delay by a number of authors (Mitchell, 2001; Macmillan et al, 2001; Mintz, 1999). This could therefore be an indication as to why the cognitive developmental scores of the sample are low.

One of the primary targets of HIV-1 infection is brain tissue, which also commonly results in reduced brain growth (Macmillan et al, 2001). Acquired microcephally has been noted by a number of authors and is usually due to impaired brain growth and cerebral atrophy secondary to encephalopathy (Mitchell, 2001; Macmillan et al, 2001; Mintz, 1999). Cortical atrophy is associated with disease progression and neurodevelopmental delay (Llorente et al, 2003; Pearson et al, 2000). Overall, the size of the head predicts the size of the brain, which is associated with IQ (Wickett et al, 2000). Brain development during infancy and early childhood is more important than foetal brain growth in determining adult cognitive abilities (Gale et al, 2003). This has implications for the children in this study, as their head circumference z-scores were very low, which could indicate that cognitive abilities could be impaired, and therefore school difficulties may be encountered.

Regular monitoring of weight, height and head circumference is recommended in all HIV infected children (Chantry and Moye, 2005; Nachman et al, 2002; Coovadia and Meyers et al, 2000; Bobat et al, 1998). Chronic malnutrition has been associated with decreased intellectual capacity, and therefore decreased earning capacity in later life, which perpetuates the cycle of poverty (Bridge et al, 2006). In a child who is not meeting his growth milestones, vitamin and micronutrient deficiencies should be considered as well (Mintz, 1999).

Research Implications

- Little is known as to the effects of HIV on school function in developing countries. More longitudinal studies need to be conducted to analyse school outcomes in relation to early growth and developmental parameters.

Clinical Implications

- Regular monitoring of weight and height is recommended at every clinic visit in order to monitor growth, and to determine ART eligibility or response to ART (DOH, 2010).

- It should be noted that growth is predictive of developmental status, and therefore both should be monitored closely.

9.5 Challenges

There are many challenges associated with conducting research in a developing country such as South Africa. These will be outlined below and their impact on the results of this study will be discussed.

9.5.1 Standardisation Sample

The Infant Gross Motor Screening Test was developed for use on HIV positive infants between the ages of six and eighteen months. The tool was therefore standardised on this population, and its applicability to other populations is not yet known. Standardisation on normal children has not yet been conducted, but due to the pressing need for the tool, and the encouraging reliability and validity results on the HIV infected population, it should prove useful for these children. Further research should be directed towards standardisation on infants from a variety of backgrounds in order to ensure that the standardisation sample is representative of the South African population. Gross motor screening may be useful in other populations such as premature infants, and therefore before the tool can be used on these populations, further standardisation, reliability and validlity studies should be carried out.

9.5.2 Age Range

The age range for the Infant Gross Motor Screening Test is currently six to eighteen months, which may need to be extended in the future, especially downwards, so that infants with gross motor developmental problems may be identified even earlier. This age range was chosen due to the low number of children under six months of age who

attend the clinics, as they are usually only diagnosed with HIV once they present with clinical symptoms.

9.5.3 Reliability and Validity

For the purposes of this study, preliminary testing was carried out to determine the reliability and validity of the Infant Gross Motor Screening Test. Larger studies should be conducted in the future in order to expand on this. Depending on the future uses for the tool, predictive validity studies may need to be conducted.

9.5.4 Test Manual

Instructions and guidelines have been drawn up for the purposes of this study, but should the Infant Gross Motor Screening Test need to be widely distributed, a full test manual would need to be printed, containing the information required by Rothstein et al (1991) as outlined in the Standards for Tests and Measurements in Physical Therapy Practice.

9.5.5 Test Domains

The Infant Gross Motor Screening Test is limited in that it assesses gross motor development only, but in this population, this is the aspect of development, which is most severely affected, and the assessment of the other areas of development is problematic for a number of reasons. The assessment of cognitive development requires equipment, which would make the tool costly, and bulky, and in many other studies, cultural differences have been shown to affect outcomes, for example, puzzle building. Language assessment is problematic firstly due to the fact that there are eleven official languages in South Africa, and receptive language assessment requires picture recognition, which is often not culturally appropriate. Gross motor function has been shown to be the most universal aspect of development, and is the easiest aspect to assess, as no equipment is required.

9.5.6 Conclusion

This study has therefore resulted in the development of The Infant Gross Motor Screening Test, which may be used in clinics to assess the gross motor function of infants between six and 18 months of age. The study has also provided important anthropometric and developmental data which further highlight the issues that HIV infected children in South Africa face.

Chapter 10

CONCLUSIONS

The main conclusions drawn from this study are outlined below:

The Bayley-III Screening Test and the BSID III do not correlate well, and therefore the Bayley-III Screening Test is unsuitable for use in the context of paediatric HIV in South Africa. A new gross motor screening tool was therefore developed in order to detect gross motor delays in HIV positive children between the ages of six and 18 months.

Important clinical findings resulting from the study are as follows: HIV infected children are at high risk of developmental delay in all domains, most markedly gross motor development. Despite the ARV rollout in South Africa, developmental delay and HIV encephalopathy continue to be problematic, and the immunologic status of the children remains poor. The infants in the study population have many additional risk factors which have an impact on their developmental status including poor socioeconomic status, malnutrition and stunting. There is therefore a need for a screening tool which will help to identify children with developmental delays so that they may be referred for further more in-depth assessment in all areas, and intervention where necessary. Monitoring of development has also been outlined in the South African Department of Health Guidelines for Antiretroviral Treatment (2010) as an ARV initiation criterion, and a treatment response criterion.

The Infant Gross Motor Screening Test has been developed for use on an HIV positive paediatric South African population. The statistical properties of the Infant Gross Motor Screening Test are excellent, and the tool has been shown to be both reliable and valid. The statistical properties of the Infant Gross Motor Screening Test are within the ranges recommended by the American Psychological Association, and the tool has been developed in accordance with the Standards for Tests and Measurements in Physical Therapy Practice. The statistical properties of The Infant Gross Motor Screening Test after preliminary testing are presented in Table 10.1 below:

Table 10.1 Statistical Properties of The Infant Gross Motor Screening Test	
Content Validity	Assessed using the NGT
Concurrent Validity	K=0.85
Sensitivity	97.4%
Specificity	85.7%
Positive Predictive Value	92.7%
Negative Predictive Value	94.7%
Interrater Reliability	r=1
Intrarater Reliability	r=1
Test-retest Reliability	r=098

The final version of the Infant Gross Motor Screening Test is presented below:

The Infant Gross Motor Screening Test

Instructions for use

Purpose

The Infant Gross Motor Screening Test was developed to assess the gross motor function of HIV positive Infants between the ages of 6 and 18 months.

User Qualifications

The Infant Gross Motor Screening Test was designed to be used by those working in a paediatric HIV setting, and does not require profession specific training. Potential users should be trained in the administration of the items, and the observation of responses from the child.

Administration time

The Infant Gross Motor Screening Test takes 5 - 10 minutes to administer.

Administration procedure

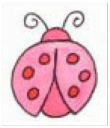
- The child's exact age in months should be calculated, and correction for prematurity should be made.
- The correct age group should be selected from the test score sheet
- Items should be administered in a quiet child-friendly environment. The child's caregiver may be used to place the child in the necessary position should the child be upset by the administrator handling him/her
- Items do not need to be administered in sequential order, but responses need to be observed in order to be credited, the administrator may not give credit based on parent-report.
- All items should be completed for the age group

Nicole Hilburn 2010 • nicolehilburn@gmail.com

Scoring

- For each item that the child achieves, a score of '1' should be given
- If the child does not achieve the item, a score of '0' should be given
- Once all the items have been completed, the child's total score should be obtained by adding up the '1's and '0's.
- The child's corresponding developmental category can be found at the bottom of the page.

Nicole Hilburn 2010 • nicolehilburn@gmail.com



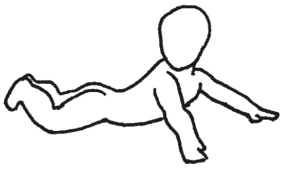
The Infant Gross Motor Screening Test 6-8 months

Date of Assessment____/____/_____

Child's name_____

Date of Birth____/____/_____

Age_____ months

		Score
	**Controls head at 90 degrees while lying on stomach** Child lies on stomach, and should be able to lift head up to 90 degrees for about 5 seconds	☐
	**Elevates chest whilst lying on stomach (with extended arms)** Child supports weight on both hands, whilst keeping arms straight, and lifting head at 90 degrees	☐
	**Plays with feet** Child should bring one or both feet to hands (above the hips), and be able to maintain this position whilst playing with feet	☐
	**Rolls from back to stomach** Child can do this to both or one side. This should be voluntary, and initiated by lifting the head	☐
	**Sits alone** The child should be able to sit alone for about 30 seconds without using arms for support	☐

6-7 months	8 months
☐ 4-5 Satisfactory	☐ 5 Satisfactory
☐ 0-3 At risk	☐ 0-4 At risk

Nicole Hilburn 2010 • nicolehilburn@gmail.com

The Infant Gross Motor Screening Test 9-12 months

Child's name_____

Date of Birth____/____/_____

Date of Assessment____/____/_____

Age_____ months

		Score
	**Turns body while seated** Child turns his or her trunk and reaches for object	☐
	**Makes stepping movements** Child makes at least two stepping movements that move him/herself forward whilst hands are held	☐
	**Moves from sitting to being on hands and knees** Child uses rotation to move from sitting to being on hands and knees	☐
	**Crawls/moves forward at least 1½ metres** Child uses crawling/moving on stomach and pulling with hands to move forwards at least 1 ½ metres	☐
	**Pulls up to standing position** Child uses an object such as a table in order to pull him/herself into a standing position, or pushes up on the floor without support	☐

9-10 months	11-12 months
☐ 4-5 Satisfactory	☐ 5 Satisfactory
☐ 0-3 At risk	☐ 0-4 At risk

Nicole Hilburn 2010 • nicolehilburn@gmail.com

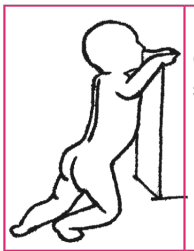
The Infant Gross Motor Screening Test 13-16 months

Child's name_____

Date of Birth____/____/_____

Date of Assessment____/____/_____

Age_____ months

		Score
	**Pulls up to standing position** Child uses an object such as a table in order to pull him/herself into a standing position, or pushes up on the floor without support	☐
	**Sits down from supported standing in a controlled manner** Child sits down from a standing position using good control e.g. by using a controlled squat	☐
	**Stands independently** Child is able to stand without support for at least 20 seconds	☐
	**Walks alone with coordination** Child should be able to walk a reasonable distance (such as 20 steps) with good control and coordination	☐
	**Squats without support** The child should be able to squat down with good control, and maintain this position for play. The child's bottom should not be resting on the floor	☐

13 months	14 months	15-16 months
☐ 3-5 Satisfactory	☐ 4-5 Satisfactory	☐ 5 Satisfactory
☐ 0-2 At risk	☐ 0-3 At risk	☐ 0-4 At risk

Nicole Hilburn 2010 • nicolehilburn@gmail.com

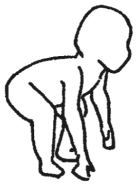
The Infant Gross Motor Screening Test 17-18 months

Date of Assessment____/____/_____

Child's name_____

Date of Birth____/____/_____

Age_____ months

		Score
	**Sits down from supported standing in a controlled manner** Child sits down from a standing position using good control e.g. by using a controlled squat	☐
	**Stands independently** Child is able to stand without support for at least 1 minute	☐
	**Stands up with no assistance** Child moves into a standing position without using an object to pull on. The child may push up on the floor in order to stand up	☐
	**Walks alone with coordination** Child should be able to walk a reasonable distance (such as 40 steps) with good control and coordination	☐
	**Squats without support** The child should be able to squat down with good control, and maintain this position for play. The child's bottom should not be resting on the floor	☐
	**Runs with coordination** The child should be able to run without falling over	☐

17-18 months

☐ 6 Satisfactory

☐ 0-5 At risk

Nicole Hilburn 2010 • nicolehilburn@gmail.com

REFERENCES

Abubakar A, Holding P, Newton C, van Baar A, van de Vijver F 2009 The Role of Weight for Age and Disease Stage in Poor Psychomotor Outcome of HIV-Infected Children in Kilifi, Kenya. Developmental Medicine and Child Neurology 51:968-973

Abubakar A, van Baar A, Van de Vivjer F, Holding P, Newton C 2008a Paediatric HIV and Neurodevelopment in Sub-Saharan Africa: A Systematic Review. Tropical Medicine and International Health 13 (7):1-8

Abubakar A, Van de Vijver F, Van Baar A, Mbonani L, Kalu R, Newton C, Holding P 2008b Socioeconomic Status, Anthropometric Status, and Psychomotor Development of Kenyan Children from Resource-Limited Settings: A Path-Analytic Study. Early Human Development 84:613-621

Abubakar A, Van de Vijver F, Mithwani S, Obiero E, Lewa N, Kenga S, Katana K, Holding P 2007 Assessing Developmental Outcomes in Children from Kilifi, Kenya, Following Prophylaxis for Seizures in Cerebral Malaria. Journal of Health Psychology 12(3):417-430

Agnarson A, Ericson J, Ekstrom A, Thorson A 2007 Antiretroviral Therapy: What About Food? AIDS 21:1225-26

Aina O and Morakinyo O 2005 Normative Data on Mental and Motor Development in Nigerian Children. The Journal of West African Medicine April-June: 24 (2):151-156

Aina O and Morakinyo O 2001 The Validation of Developmental Screening Inventory (DSI) on Nigerian children. Journal of Tropical Pediatrics 47:323-328

Altman D, Bland M 1994 Diagnostic Tests 2: Predictive Values. British Medical Journal 309:102

American Academy of Pediatrics Policy Statement 2006 Identifying Infants and Young Children With Developmental Disorders in the Medical Home: An Algorithm for Developmental Surveillance and Screening. Pediatrics 118(1):405-421

American Academy of Pediatrics: Committee on Children with Disabilities 2001 Developmental Surveillance and Screening of Infants and Young Children. Pediatrics 108(1):192-197

American Psychological Association 1985 Standards for Educational and Psychological Tests: Washington DC APA

ANNECCA 2005 Early Diagnosis of Pediatric HIV Infection in Sub-Saharan Africa. Advocacy Statement. www.annecca.org (accessed March 2008)

Armstrong F, Seidel J, Swales T 1993 Pediatric HIV Infection: A Neuropsychological and Educational Challenge. Journal of Learning Disabilities 26:92-103

Arpadi S 2000 Growth Failure in Children with HIV Infection. Journal of AIDS 25:S37-S42

Aylward G 2009 Developmental Screening and Assessment: What Are We Thinking? Journal of Developmental and Behavioural Pediatrics 30(2):169-173

Aylward G and Verhulst S 2000 Predictive Utility of the Bayley Infant Neurodevelopmental Screener (BINS) Risk Status Classifications: Clinical Interpretation and Application. Developmental Medicine and Child Neurology 42:25-51

Aylward G 1997 Conceptual Issues in Developmental Screening and Assessment. Developmental and Behavioural Pediatrics 18(5):340-349

Aylward G. Bayley Infant Neurodevelopmental Screener. New York, NY: Psychological Corporation; 1995

Bachou H, Tylleskar T, Downing R, Tumwine J 2006 Severe Malnutrition With and Without HIV-1 Infection in Hospitalised Children in Kampala, Uganda: Differences in Clinical Features: Haematological Findings and CD4+ Cell Count. Nutrition Journal 5(1):27

Baillieu N, Potterton J 2008 The Extent of Delay of Language, Motor and Cognitive Development in HIV Positive Infants. Journal of Neurologic Physical Therapy September 32 (3):118-121

Banks W, Ercal N, Price T 2006 The Blood-Brain Barrier in NeuroAIDS. Current HIV Research 4:259-266

Barbarin O and Richter L 2001 Mandela's Children. Growing Up in Post Apartheid South Africa. Routledge, New York, London

Barbarin, O and Khomo N 1997 Indicators of Economic Status and Social Capital in South African Townships. Childhood 4(2): 193 – 222

Bart O, Rosenberg L, Ratzon N, Jarus T 2010 Development and Initial Validation of the Performance Skills Questionnaire (PSQ). Research in Developmental Disabilities 31:46-56

Bartlett J, Shao J 2009 Successes, Challenges, and Limitations of Current Antiretroviral Therapy in Low-Income and Middle-Income Countries. Lancet 9:637-646

Bayley N 2006a Bayley Scales of Infant Development, Third Edition. Manual. San Antonio TX: The Psychological Corporation

Bayley N 2006b Bayley Scales of Infant Development, Third Edition. Technical Manual. San Antonio TX: The Psychological Corporation

Bayley N 2006c Bayley Scales of Infant Development, Third Edition. Screening Test Manual. San Antonio TX: The Psychological Corporation

Bayley, N 1996 Administration Manual for the Bayley Screening Test, San Antonio, TX: Psychological Corporation

Bayley N 1993 Bayley Scales of Infant Development, Second Edition. Manual. San Antonio TX: The Psychological Corporation

Beaglehole R, Bonita R, and Kjellstrom T 1993 Basic Epidemiology: World Health Organization Geneva.

Beck C, Gable R 2001 Ensuring Content Validity: An illustration of the Process. Journal of Nursing Measurement 9:201-215

Beckung E 2000 The Development and Validation of a Measure of Motor and Sensory Function in Children with Epilepsy. Pediatric Physical Therapy 12:24-35

Bellamy N, Anastassiades P, Watson Buchanan W, Davis P, Lee P, McCain G, Wells G, Campbell J 1991 Rheumatoid Arthritis Antirheumatic Drug Trials III: Setting the Delta for Clinical Trials of Antirheumatic Drugs - Results of a Consensus Development (Delphi) Exercise. Journal of Rheumatology 18:1908-1915

Belman A, Muenz L, Marcus J, Goedert J, Landesman S, Rubinstein A, Goodwin S, Durako S, Willoughby A 1996 Neurologic Status of Human Immunodeficiency Virus –1 Infected Infants and their Controls: A Prospective Study from Birth to 2 Years. Pediatrics 98 (6):1109-1118

Belman A 1994 HIV-1 Associated CNS Disease in Infants and Children. In: HIV, AIDS and the Brain (R.W. Price and S.W Perry, eds), Raven Press, New York pp 289-310

Belman A 1992 Acquired Immunodeficiency Syndrome and the Child's Central Nervous System. Pediatric Clinics of North America 39 (4):691-713

Belman A 1990 AIDS and Pediatric Neurology. Neurology Clinics 8 (3):571-603

Belman A, Diamond G, Dickson D, Horoupian D, Llena J, Lantos G, Rubinstein A 1988 Pediatric Acquired Immunodeficiency Syndrome: Neurologic Syndromes. American Journal of Disease in Childhood 142:29-35

Belman A, Ultmann H, Horoupian D, Novick B, Spiro A, Rubinstein A, Kurtzberg D, Cone-Wesson B 1985 Neurological Complications in Infants and Children with Acquired Immune Deficiency Syndrome. Annals of Neurology 18:560-566

Bennett D, Bendersky M, Lewis M 2002 Children's Intellectual and Emotional-Behavioral Adjustment at Four Years as a Function of Cocaine Exposure, Maternal Characteristics and Environmental Risk. Developmental Psychology 38:648-658

Bhargava A 1998 A Dynamic Model for the Cognitive Development of Kenyan Schoolchildren. Journal of Educational Psychology 90 (1): 162–166

Bissiachi P, Suppiej A, Laverda A 2000 Neuropsychological Evaluation of Neurologically Asymptomatic HIV-Infected Children. Brain and Cognition 43: 49 - 52

Black R, Morris S, Bryce J 2003 Where and Why Are 10 Million Children Dying Every Year? Lancet 361:2226-34

Blanche S, Tardieu M, Duliege A, Rouzioux C, Le Diest F, Fukunaga K, Caniglia M, Jacomet C, Messiah A, Griscelli C 1990 Longitudinal Study of 94 Symptomatic Infants with Perinatally Acquired Human Immunodeficiency Virus Infection. American Journal of the Diseases of Children 144:1210-1215

Blanchette N, Smith M, King S, Fernandes-Penney A, Read S 2002 Cognitive Development in School-age Children with Vertically Transmitted HIV Infection. Developmental Neuropsychology 21 (3):223-241

Blanchette N, Smith M, Fernandes-Penney A, King S, Read S 2001 Cognitive and Motor Development in Children with Vertically Transmitted HIV Infection. Brain and Cognition 46:46-49

Bobat R, Coovadia H, Moodley D, Coutsoudis A, Gouws E 2001 Growth in Early Childhood in a Cohort of Children Born to HIV-1 Infected Women from Durban, South Africa. Annals of Tropical Paediatrics 21:203–210

Bobat R, Moodley D, Coutsoudis A, Coovadia H, Gouws E 1998 The Early Natural History of Vertically Transmitted HIV-1 Infection in African Children from Durban, South Africa. Annals of Tropical Paediatrics 18:187-196

Bodkin A, Robinson C, Perales F 2003 Reliability and Validity of the Gross Motor Function Classification System for Cerebral Palsy. Pediatric Physical Therapy 15:247-252

Brackis-Cott E, Kang E, Dolezal C, Abrams E, Mellins C 2009 Brief Report: Language Ability and School Functioning of Youth Perinatally Infected With HIV. Journal of Pediatric Health Care 23(3):158-164

Bradley R, Corwyn R 2005 Caring for Children Around the World: a View From Home. International Journal of Behavioural Development 29:468-78

Bradley R, Corwyn R 2002 Socioeconomic Status and Child Development. Annual Review of Psychology 53:371-99

Bradley R, Corwyn R 1996 Life at Home: Same Time, Different Places. Early Development and Parenting 5:251-69

Bradley R, Caldwell B, Rock S, Barnard K, Gray C, Hammond M, Mitchell S, Siegel L, Ramey C, Gottfried A, Johnson D 1989 Home Environment and Cognitive Development in the First Three Years of Life: A Collaborative Study Involving Six Sites and Three Ethnic Groups in North America. Developmental Psychology 25:217-235

Brady M, McGrath N, Brouwers P, Gelber R, Glenn-Fowler M, Yogev R, Hutton N, Bryson Y, Mitchell C, Fikrig S, Borkowsky W, Jiminez E, McSherry G, Rubinstein A, Wilfert C, McIntosh K, Elkins M, Weintrub P and the Pediatric AIDS Clinical Trials Group 1995 Randomized Study of the Tolerance and Efficacy of High-versus- Low Dose Zidovudine in Human Immunodeficiency Virus-Infected Children with Mild to Moderate Symptoms. Journal of Infectious Disease 173:1097-1106

Brahmbhatt H, Kigozi G, Wabwire-Mangen F, Serwadda D, Lutalo T, Nalugoda F, Sewankambo N, Kiduggavu M, Wawer M, Gray R 2006 Mortality in HIV Infected and Uninfected Children of HIV-Infected and Uninfected Mothers in Rural Uganda. Journal of Acquired Immune Deficiency Syndromes. 41: 504-508

Bricker D, Squires J. Ages and Stages Questionnaires: A Parent-Completed, Child-Monitoring System, 2nd ed. Baltimore: Paul Brookes; 1999

Bricker D, Squires J, Mounts L. Ages and Stages Questionnaires: A Parent-Completed, Child-Monitoring System. Baltimore: Paul Brookes; 1995

Bridge A, Kipp W, Jhangri GS, Laing L, Konde-Lule J 2006 Nutritional Status Of Young Children In AIDS-Affected Households And Controls In Uganda. American Journal of Tropical Medicine and Hygiene 74:926-931

Bronfenbrenner U 1977 Toward an Experimental Ecology of Human Development. American Psychologist 32:513-31

Brooks-Gunn J, Klebanov K, Duncan G 1996 Ethnic Differences in Children's Intelligence Test Scores: Role of Economic Deprivation, Home Environment and Maternal Characteristics. Child Development 67:396-408

Brouwers P, DeCarli C, Civatello L, Moss H, Wolters P, Pizzo P 1995 Correlations between CT-Brain Scan Abnormalities and Neuropsychological Function in Children with Symptomatic HIV Disease. Archives of Neurology 52:39-44

Brown D 2009 Developmental Differences between Preterm and Full Term 18 Month Olds. Unpublished MSc Thesis. University of the Witwatersrand.

Buchacz K, Cervia J, Lindsey J, Hughes M, Seage G 3rd, Dankner W, Oleske J, Moye J 2001 Impact of Protease Inhibitor-Containing Combination Antiretroviral Therapies on Height and Weight Growth in HIV-Infected Children. Pediatrics 108(4) e72 (accessed 02/10)

Burgess S, Propper C, Rigg J and the ALSPAC Study Team 2004 The Impact of Low Income on Child Health: Evidence from a Birth Cohort Study. CASE paper 85 London School of Economics http://ssrn.com/abstract=1159316

Callens S, Kitetele F, Lusiama J, Shabani N, Edidi S, Colebunders R, Behets F, Van Rie A 2008 Computed CD4 Percentage as a Low-Cost Method for Determining Pediatric Antiretroviral Treatment Eligibility. BMC Infectious Diseases 8:31 http://www.biomedcentral.com/1471-2334/8/31

Carey V, Yong F, Frenkel L, McKinney R Jr 1998 Pediatric AIDS Prognosis Using Somatic Growth Velocity. AIDS 12:136-1369

Chantry C and Moye J 2005 Growth, Nutrition and Metabolism. In: Textbook of Pediatric HIV Care Edited by Zeichner S and Read J. Cambridge University Press

Chase C, Ware J, Hittelman J, Blasini I, Smith R, Llorente A, Anisfield E, Diaz C, Fowler M, Moye J, Kaligh L 2000 Early Cognitive and Motor Development Amongst Infants Born to Women Infected with Human Immunodeficiency Virus. Pediatrics; 106 (2): e25 (accessed 04/04)

Chase C, Vibbert M, Pelton S, Coulter D, Cabral H 1995 Early Neurodevelopmental Growth in Children with Vertically Transmitted Human Immunodeficiency Virus Infection. Archives of Pediatric and Adolescent Medicine 149:850-855

Chiappini E, Galli L, Tovo P, Gabiano C, Gattinara G, Guarino A, Badolato R, Giaquinto C, Lisi C, de Martino M 2006 Italian Register for HIV Infection in Children: Virologic, Immunologic, and Clinical Benefits From Early Combined Antiretroviral Therapy in Infants with Perinatal HIV-1 Infection. AIDS 20:207-15

Chouquet C, Richardson S, Burgard M, Blanche S, Mayaux M, Rouzioux C, Costagliola D 1999 Timing of Human Immunodeficiency Virus Type 1 (HIV-1) Transmission From Mother to Child: Bayesian Estimation Using a Mixture. Statistics in Medicine 18:815-33

Cocchi F, DeVico A, Garzino-Demo A, Arya S, Gallo R, Lusso P 1995 Identification of RANTES, MIP-1 a / MIP-1 b as the Major HIV-Suppressive Factors Produced by CD8/ T cells. Science 270: 1811–1815

Cohen S, Mundy T, Karassik B, Lieb L, Ludwig D, Ward J 1991 Neuropsychological Functioning in Human Immunodeficiency Virus Type 1 Seropositive Children Infected through Neonatal Blood Transfusions. Pediatrics 88:58-68

Connolly K, Grantham-McGregor S 1993 Key Issues in Generating a Psychological-Testing Protocol. American Journal of Clinical Nutrition 57:317-318

Cooper S, Lyall H, Walters S, Tudor-Williams G, Habibi G, de Munter C, Britto J, Nadel S 2004 Children with Human Immunodeficiency Virus Admitted to a Paediatric Intensive Care Unit in the United Kingdom Over a 10-year Period. Intensive Care Medicine 30 (1):113-118

Coovadia H, Coutsoudis A 2007 HIV Infant Feeding And Survival: Old Wine In New Bottles, But Brimming With Promise. AIDS 21:1837-1840

Coovadia H, Bland R 2007 Preserving Breastfeeding Practice Through the HIV Pandemic. Tropical Medicine and International Health 12:1116-1133

Coovadia A and Meyers T 2001 HIV/AIDS – Recognition and Management in the Absence of Antiretrovirals. Continuing Medical Education 19(7):460-467

Coplan J, Contello K, Cunningham C, Weiner L, Dye T, Roberge L, Wojtowycz M, Kirkwood K 1998 Early Language Development in Children Exposed to or Infected with Human Immunodeficiency Virus. Pediatrics 102:8-14

Coscia J, Christenson B, Henry R, Wallston K, Radcliffe J, Rutstein R 2001 Effects of Home Environment, Socio-Economic Status, and Health Status on Cognitive Functioning in Children with HIV-1 Infection. Journal of Pediatric Psychology 26 (6): 321-329

Coster W 1995 Critique of the Alberta Infant Motor Scale. Physical and Occupational Therapy in Pediatrics 15:53-64

Coutsoudis A, Dabis F, Fawzi W, Gaillard P, Haverkamp G, Harris D, Jackson J, Leroy V, Meda N, Msellati P, Newell M, Nsuati R, Read J, Wiktor S (breastfeeding and HIV International Transmission Study Group) 2004 Late Postnatal Transmission of HIV-1 in Breastfed Children: An Individual Patient Data Meta-Analysis. Journal of Infectious Diseases 189:2154-66

Cowburn C, Hatherill M, Eley B, Nuttall J, Hussey G, Reynolds L, Waggie Z, Vivian L, Argent A 2007 Short-term Mortality and Implementation of Antiretroviral Treatment for Critically Ill HIV-Infected Children in a Developing Country. Archives of Disease in Childhood 92:234-241

Cysique L, Maruff P, Brew B 2004 Antiretroviral Therapy in HIV Infection: Are Neurologically Active Drugs Important? Archives of Neurology 61:1699-1704

Dabis F, Ekpini E 2002 HIV-1/AIDS and Maternal and Child Health in Africa. Lancet 359:2097-2104

Dabis F, Leroy V, Castetbon K, Spira R, Newell M-L, Salamon R 2000 Preventing Mother-to-Child Transmission of HIV-1 in Africa in the Year 2000. AIDS 14:1017-1026

Darrah J, Piper M, Watt M 1998 Assessment of Gross Motor Skills of At-Risk Infants: Predictive Validity of the Alberta Infant Motor Scale. Developmental Medicine and Child Neurology 40:485-491

Davies M, Boulle A, Fakirt T, Nuttall J, Eley B 2008 Adherence to Antiretroviral Therapy in Young Children in Cape Town, South Africa, Measured by Medication Return and Caregiver Self-Report: A Prospective Cohort Study. BMC Pediatrics 8:34 http://www.biomedcentral.com/1471-2431/8/34 (Accessed 7th November 2008)

Davis L, Hjelle B, Miller V, Palmer D, Llewellyn A, Merlin T, Young S, Mills R, Wachsman W, Wiley C 1992 Early Viral Brain Invasion in Iatrogenic Human Immunodeficiency Virus Infection. Neurology 42:1736-9

Dearlove J, Kearney D 1990 How Good is General Practice Developmental Screening? British Medical Journal 300:1177-1180

DeCarli C, Civitello A, Brouwers P, Pizzo P 1993 The Prevalence of Computed Tomographic Abnormalities of the Cerebrum in 100 Consecutive Children Symptomatic with the Human Immune Deficiency Virus. Annals of Neurology 34**:**198-205

De Luca A, Ciancio B, Larussa D, Murri R, Cingolani A, Rizzo M, Giancola M, Ammassari M, Ortona L 2002. Correlates of Independent HIV-1 Replication in the CNS and of its Control by Antiretrovirals. Neurology 59:342-347

Deeks S, Barbour J, Martin J, Swanson M, Grant R 2000 Sustained CD4+ T Cell Response After Virologic Failure of Protease Inhibitor-Based Regimens in Patients with Human Immunodeficiency Virus Infection. Journal of Infectious Diseases 181: 946-53

De Villiers M, de Villiers P, Kent A 2005 The Delphi Technique in Health Sciences Education Research . Medical Teacher 27(7):639-643

Department of Health Republic of South Africa 2010 The South African Antiretroviral Treatment Guidelines. Internally Circulated Document

Department of Health Republic of South Africa 2008 Guidelines for the Treatment and Management of HIV www.doh.co.za (accessed 01/10)

Department of Health Republic of South Africa 2007 Progress Report on Declaration of Commitment on HIV and AIDS. www.doh.co.za (accessed 02/09)

Diamond G, Gurdin P, Wiznia A, Belman A, Rubinstein A, Cohen H 1990 Effects of Congenital HIV Infection on Neurodevelopmental Status of Babies in Foster Care. Developmental Medicine and Child Neurology 32:999-1005

Dibley M, Goldsby J, Staehling N, Trowbridge F 1987 Development of Normalized Curves for the International Growth Reference: Historical and Technical Considerations. Americal Journal of Clinical Nutrition 46(5):736-748.

Drotar D, Olness K, Wizniter M, Guay L, Marum L, Svilar G, Hom D, Fagan J, Ndugwa C, Kiziri-Mayengo R 1997 Neurodevelopmental Outcomes of Ugandan Infants with Human Immunodeficiency Virus Type 1 Infection. Pediatrics 100 (1): e5 http://www.pediatrics.org/cgi/content/full/100/1/e5 (accessed 06/04)

Duncan G, Brookes-Gunn J, Klebanov P 1994 Economic Deprivation and Early Childhood Development. Child Development 65:296-318

Dunfee R, Thomas E, Gabuzda D 2009 Enhanced Macrophage Tropism of HIV in Brain and Lymphoid Tissues is Associated with Sensitivity to the Broadly Neutralizing CD4 Binding Site Antibody B12. Retrovirology 6:69
http://www.retrovirology.com/content/6/1/69 (Accessed 09/10)

Dunn W, in Law M, Baum C and Dunn W (eds) Measuring Occupational Performance: Supporting Best Practice in Occupational Therapy. 2nd edition Slack Incorporated, USA 2005

Dunn D 2003 Short-term Risk of Disease Progression in HIV-1-Infected Children Receiving No Antiretroviral Therapy or Zidovudine Monotherapy: A Meta-Analysis. Lancet 362:1605-1611

Dworkin P 1993 Detection of Behavioral, Developmental, and Psychosocial Problems in Pediatric Primary Care Practice. Current Opinion in Pediatrics 5:531-536

Dyrbye L, Szydlo D, Downing S, Sloan J, Shanafelt T 2010 Development and Preliminary Psychometric Properties of a Well-being Index for Medical Students. BMC Medical Education 10: 8 http://www.biomedcentral.com/1472-6920/10/8/ (accessed 03/10)

Eickmann S, Lima A, Guerra M, Lima M, Lira P, Huttly S, Ashworth A 2003 Improved Cognitive and Motor Development in a Community-Based Intervention of Psychosocial Stimulation in Northeast Brazil. Developmental Medicine and Child Neurology 45:536-541

Ellis R, Calero P, Stockin M 2009 HIV Infection and the Central Nervous System: A Primer. Neuropsychology Reviews 19:144-151

Embree J, Bwayo J, Nagelkerke N, Njenga S, Nyange P, Ndinya-Achola J, Pamba H, Plummer F 2001 Lymphocyte Subsets in Human Immunodeficiency Virus Type 1-Infected and Uninfected Children in Nairobi. Pediatric Infectious Disease Journal 20(4):397-403.

Emond A, Blair P, Emmett P, Drewett P 2007 Weight Faltering in Infancy and IQ Levels at 8 Years in the Avon Longitudinal Study of Parents and Children. Pediatrics 120 (4) e1050-1058

Engle P, Black M, Behrman J, de Mello MC, Gertler P, Kipiriri L, Martorell R, Young M and the International Child Development Steering Group 2007 Strategies to Avoiding the Loss of Developmental Potential in More than 200 Million Children in the Developing Countries. Lancet 369:229-242

Epstein L, Gendelman H 1993 Human Immunodeficiency Virus Type 1 Infection of the Nervous System: Pathogenetic Mechanisms. Annals of Neurology 33:429-436

Epstein L, Goudsmit J, Paul D, Morrison S, Connor E, Oleske J, Holland B 1987 Expression of Human Immunodeficiency Virus in Cerebrospinal Fluid of Children with Progressive Encephalopathy. Annals of Neurology 21:397-401

Epstein L, Sharer L, Oleske J, Connor E, Goudsmit J, Bagdon L, Robert-Guroff M, Koeningsberger R 1986 Neurologic Manifestations of Human Immunodeficiency Virus Infection in Children. Pediatrics 78:678-687

Epstein L, Sharer L, Joshi V, Fojas M, Koenigsberger M, Oleske J 1985 Progressive Encephalopathy in Children with Acquired Immune Deficiency Syndrome. Annals of Neurology 17:488-496

European Collaborative Study 2003 Height, Weight and Growth in Children Born to Mothers with HIV-1 Infection in Europe. Pediatrics 111:e52-e60 (accessed 11/09)

European Collaborative Study 1999 Maternal Viral Load and Vertical Transmission of HIV-1: an Important Factor but not the Only One. AIDS 13:1377-85

European Collaborative Study 1996 Vertical Transmission of HIV-1: Maternal Immune Status and Obstetric Factors. AIDS 10:1675-81

European Collaborative Study 1995 Weight, Height and Human Immunodeficiency Virus Infection in Young Children of Infected Mothers. Pediatric Infectious Disease Journal 14:685-690

European Collaborative Study 1994 Natural History of Vertically Acquired Human Immunodeficiency Virus-1 Infection. Pediatrics 94 (6):815-19

European Mode of Delivery Collaboration 1999 Elective Caesarean Section versus Vaginal Delivery in Preventing Vertical HIV-1 Transmission: a Randomised Clinical Trial. Lancet 353:1035-9

Ferguson G, Jelsma J 2009 The Prevalence of Motor Delay Among HIV Infected Children Living in Cape Town, South Africa. International Journal of Rehabilitation Research 32:108-114

Feucht U, Kinzer M, Kruger M 2007 Reasons for Delay in Initiation of Antiretroviral Therapy in a Population of HIV-Infected South African Children. Journal of Tropical Pediatrics 53 (6): 398-402

Fiser D, Tilford J, Roberson P 2000 Relastionship of Illness Severity and Length of Stay to Functional Outcomes in the Pediatric Intensive Care Unit: A Multi- Institutional Study. Critical Care Medicine 28 (4):1173-1179

Flynn J 1999 Searching for Justice: The Discovery of IQ Gains Over Time. American Psychology 54:728-748

Fowler M 1994 Paediatric HIV Infection: Neurologic and Neuropsychological Findings. Acta Paediatra Supplementa 400:59-62

Fragoso Y, Andalat R, Adamo A, Lopes de Fonseca N, Moryiama M 1999 Neurologic Manifestations of AIDS in Children and Adolescents: A review of cases in Santos, Brazil. Medscape General Medicine 24:E3 www.medscape.com/viewarticles/408010 (accessed 09/4/04)

Frankenburg W, Dodds J, Archer P, Shapiro H, Bresnick B 1992. The Denver II: A Major Revision and Restandardization of the Denver Developmental Screening Test. Pediatrics 90:477-490

Fuller K, Owens J, Chambers T 1995 Macrophage Inflammatory Protein 1a and IL-8 Stimulate the Motility but Suppress the Resorption of Isolated Rat Osteoclasts. Journal of Immunology 154:6065-6071

Fuller R, Westmoreland S, Ratai E, Greco J, Kim J, Lentz M, He J, Prabhat S, Masliah E, Halpern E, Lackner A, Gonzalez G 2004 A Prospective Longitudinal In Vivo H MR Spectroscopy study of the SIV/Macaque Model of NeuroAIDS. BMC Neuroscience 5:10 www.biomedcentral.com/1471-2202/5/10 (accessed 7/9/04)

Gabuzda D, Hirsch M 1986 Neurologic Manifestations of Infection with Human Immunodeficiency Virus: Clinical Features and Pathogenesis. Annals of Internal Medicine 107:383-391

Gale C, O'Callaghan F, Godfrey K, Law C, Martyn C 2004 Critical Periods of Brain Growth and Cognitive Function in Children. Brain 127:321-329

Gale C, Walton S, Martyn C 2003 Foetal and Postnatal Head Growth and Risk of Cognitive Decline in Old Age. Brain 126:2273-2278

Gardner J, Powell C, Baker-Henningham H, Walker S, Cole T, Grantham-McGregor S 2005 Zinc Supplementation and Psychosocial Stimulation: Effects on the Development of Undernourished Jamaican Children. American Journal of Clinical Nutrition 82:399-405

Gardner J, Grantham-McGregor S, Himes J, Chang S 1999 Behaviour and Development of Stunted and Nonstunted Jamaican Children. Journal of Child Psychology and Psychiatry 40:819-27

Gay C, Armstrong D, Cohen D, Lai S, Hardy M, Swales T, Morrow C, Scott G 1995 The Effects of HIV on Cognitive and Motor Development in Children Born to HIV-Seropositive Women with no Reported Drug Use: Birth to 24 months. Pediatrics 96: 1078-1082

Gendelman H, Lipton S, Tardieu M, Bukrinsky M, Nottet H 1994 The Neuropathogenesis of HIV-1 Infection. Journal of Leukocyte Biology 56:389-398

Giagazoglou P, Tsiamaras V, Fatiadou E, Evaggelinous C, Tsikoulas J, Angelopoulou N 2005 Standardization of the Motor Scales of the Griffiths Test II on Children Aged 3 to 6 Years in Greece. Child Care Health and Development 31:321-30

Gladstone M, Lancaster G, Jones A, Maleta K, Mtitimila E, Ashorn P, Smyth R 2008 Can Western Developmental Screening Tools be Modified for Use in a Rural Malawian Setting?. Archives of Disease in Childhood 93:23-29

Glascoe F 2001 Are Overreferrals on Developmental Screening Tests Really a Problem? Archives of Pediatric and Adolescent Medicine155:54-59

Glascoe F 1997 Parent's Concerns About Children's Development: Prescreening Technique or Screening Test? Pediatrics 99:522-528

Glascoe F, Byrne K 1993 The Accuracy of Three Developmental Screening Tests. Journal of Early Intervention 17:368-379

Glascoe F, Byrne K, Ashford L, Johnson K, Chang B, Strickland B 1992 Accuracy of the Denver-II in Developmental Screening. Pediatrics 89:1221-1225

Glascoe F 1991 Developmental Screening: Rationale, Methods and Application. Infants and Young Children 4(1):1-10

Gorman K, Pollitt E 1996 Does Schooling Buffer the Effects of Early Risk? Child Development 67:314–26

Gortmaker S, Hughes M, Cervia J, Brady M, Johnson G, Seage G 2001 Effect of Combination Therapy Including Protease Inhibitors on Mortality Among Children and Adolescents Infected with HIV-1. New England Journal of Medicine 345:1522-8

Goudsmit J, de Wolf F, Paul D, Lange J, Speelman H, Van der Noordaa J, Van der Helm H, Epstein G, Krone W, Wolters E, Oleske J, Coutinho R 1986 Expression of Human Immunodeficiency Virus Antigen (HIV-ag) in Serum and Cerebrospinal Fluid during Acute and Chronic Infection. Lancet 2:177-180

Gowland C, Boyce W, Wright V, Russell D, Goldsmith C, Rosenbaum P 1995 Reliability of the Gross Motor Performance Measure. Physical Therapy 75:597-602

Grantham-McGregor S, Cheung Y, Cueto S, Glewwe P, Richter L, Strupp B and the International Child Development Steering Group 2007 Child Development in Developing Countries 1: Developmental Potential in the First 5 Years for Children in Developing Countries. Lancet 369:60-70

Gray L, Newell M, Thorne C European Collaborative Study 2001 Fluctuations in Symptoms in Human Immunodeficiency Virus- Infectedc Children: The First 10 Years of Life. Pediatrics 108:116-122

Gutman L, Sameroff A, Cole R 2003 Academic Growth Curve Trajectories From 1st Grade to 12th Grade: Effects of Multiple Social Risk Factors and Preschool Child Factors. Developmental Psychology 39:777-790

Haataja, L, McGready R, Arunjerdia R, Simpson J, Mercuri E, Nosten F, Dubowitz L 2002 A New Approach for Neurological Evaluation of Infants in Resource- Poor Settings. Annals of Tropical Pediatrics *22*:355-368

Harries A, Nyangulu D, Hargreaves N, Kaluwa O, Salaniponi F 2001 Preventing Antiretroviral Anarchy in Sub-Saharan Africa. Lancet 358 (9279): 410-414

Harris S, Megens A, Backman C, Hayes V 2005 Stability of the Bayley II Scales of Infant Development in a Sample of Low-Risk and High-Risk Infants. Developmental Medicine and Child Neurology 47:820-823

Harris S, Megens A, Backman C, Hayes V 2003 Development and Standardization of the Harris Infant Neuromotor Test. Infants and Young Children 16 (2):143-151

Harris S, Daniels L 2001 Reliability and Validity of the Harris Infant Neuromotor Test. Journal of Pediatrics 139:249-253

Harris S, Daniels L 1996 Content Validity of the Harris Infant Neuromotor Test. Physical Therapy 76:727-737

Harris S 1991 Development of an Infant Neuromotor Assessment Tool. Mary E, Switzer Research Fellowship final report. Washington, DC: National Institute on Disability and Rehabilitation Research, U.S. Department of Education.

Heller L 1997 Nutrition Support for Children with HIV/AIDS. Journal of the American Dietetic Association 97:473-474

Henry R, Christensen, B, Coscia J, Cohen F, Moore E 1996 Relationship Between Cognitive and Immune Functioning in Children Born to HIV-Seropositive Women. Developmental Neuropsychology 12(3): 283-298

Hess C, Papas M, Black M 2004 Use of the Bayley Neurodevelopmental Screener with an Environmental Risk Group. Journal of Pediatric Psychology 29(5): 321 – 330

Hickey W, Williams K 1999 Leukocyte Traffic in the Central Nervous System: The Participants and Their Roles. Seminars in Immunology 11:125-137

HIV Paediatric Prognostic Markers Collaborative Study 2005 Use of Total Lymphocyte Count for Informing When to Start Antiretroviral Therapy in HIV-infected Children: A Meta-analysis of Longitudinal Data. Lancet 366(9500):1868-1874

HIV Paediatric Prognostic Markers Collaborative Study Group 2003 Short-term Risk of Disease Progression in HIV-1-Infected Children Receiving No Antiretroviral Therapy or Zidovudine Monotherapy: Estimates According to CD4 Percent, Viral Load, and Age. Lancet 362:1605-1611

Ho D, Rota T, Schooley T, Kaplan J, Allan D, Groopman J, Resnick L, Felsenstein D, Andrews C, Hirsch M 1985 Isolation of HTLV-III from Cerebrospinal Fluid and Neural Tissues of Patients with Neurologic Syndromes Related to the Acquired Immunodeficiency Syndrome New England Journal of Medicine 313:1493-1497

Hochhauser C, Gaur S, Marone R and Lewis M 2008 The Impact Of Environmental Risk Factors On HIV-Associated Cognitive Decline In Children. AIDS Care 20 (6):692-699

Holding P, Kitsao-Wekulo P 2004 Describing the Burden of Malaria on Child Development: What Should We Be Measuring and How Should We Be Measuring It? American Journal of Tropical Medicine and Hygiene 71:71-79

Hyrkas K, Appelqvist-Schmidlechner K, Oksa L 2003 Validating an Instrument for Clinical Supervision Using an Expert Panel. International Journal of Nursing Studies 40(6):619–25

Ioannidis J, Tatsioni A, Abrams J 2004 Maternal Viral Load and Rate of Disease Progression Among Vertically HIV-1-Infected Children: An International Meta-Analysis. AIDS 18:99-108

Jeremy R, Kim S, Nozyce M, Nachman S, McIntosh K, Pelton S, Yogev R, Wiznia A, Johnson G, Krogstad P, Stanley K 2005 Neuropsychological Functioning and Viral Load in Stable Antiretroviral Therapy-Experienced HIV-Infected Children. Pediatrics 115:380-387

Jin X, Sun Y, Jiang F, Ma J, Morgan C, Shen X 2007 “Care for Development” Intervention in Rural China: A Prospective Follow-up Study. Journal of Developmental and Behavioural Pediatrics 28:213-218

Jones J, Brown E, Volicer L 2000 Target Outcomes for Long-Term Oral Health Care in Dementia: A Delphi Approach. Journal of Public Health Dentistry 60:330-334

Jones J, Hunter D 1995 Consensus Methods For Medical and Health Services Research. British Medical Journal, 311:376-380

Jones J, Sanderso C, Black N 1992 What Will Happen To The Quality Of Care With Fewer Junior Doctors? A Delphi Study Of Consultant Physicians’ Views. Journal of the Royal College of Physicians of London 26:36-40

Kaul M, Lipton S 2006 Mechanisms of Neuronal Injury and Death in HIV-1 Associated Dementia. Current HIV Research 4:307-318

Kelly Y, Sacker A, Schoon I, Nazroo J 2006 Ethnic Differences in Achievement of Developmental Milestones by 9 Months of Age: The Millenium Cohort Study. Developmental Medicine and Child Neurology 48:825-830

Kim J, Gilks C 2005 Scaling Up Treatment - Why We Can't Wait. New England Journal of Medicine 353(22):2392-2394

Kirby A, Edwards L, Sugden D, Rosenblum S 2010 The Development and Standardisation of the Adult Developmental Co-ordination Disorders/Dyspraxia Checklist (ADC). Research in Developmental Disabilities 31:131-139

Knight W, Mellins C, Levenson R, Arpadi S, Kairam R 2000 Brief Report: Effects of Pediatric HIV Infection on Mental and Psychomotor Development. Journal of Pediatric Psychology 25(8):583-587

Kolson D 2002 Neuropathogenesis of Central Nervous System HIV-1 Infection. Clinical Laboratory Medicine 22:703-717

Korenman S, Miller J, Sjaastad J 1995 Longterm Poverty and Child Development in the US: Results from the NLSY. Children and Youth Services Review 17:127-155

Kuhn L, Steketee RW, Weedon J 1999 Distinct Risk Factors for Intrauterine and Intrapartum Human Immunodeficiency Virus Transmission and Consequences For Disease Progression in Infected Children. Perinatal AIDS Collaborative Transmission Study. The Journal of Infectious Diseases 179:52-58

Landesman S, Kalish L, Burns D, Minkoff H, Fox H, Zorrilla C, Garcia P, Fowler M, Mofenson L, Tuomala R for the Women and Infants Transmission Study 1996 Obstetrical Factors and the Transmission of Human Immunodeficiency Virus Type 1 from Mother-to-Child. New England Journal of Medicine 334 (25):1617-23

Landis J, and Koch G 1977 The Measurement of Observer Agreement for Categorical Data. Biometrics 33:159-174

Laue L, Pizzo P, Butler K, Cutler G 1990 Growth and Neuroendocrine Dysfunction in Children with Acquired Immunodeficiency Syndrome. Journal of Pediatrics 117:541-545

Laufer M, Scott G 2000 Medical Management of HIV Disease in Children. Pediatric Clinics of North America 47 (1):127-153

Layton T, Scott G 2000 Language Development and Assessment in Children with Human Immunodeficiency Virus: 3 to 6 years. Seminars in Speech and Language 21(1): 37-46

Lepage P, Van de Perre P, Van Vliet G, Nsengumuremyi F, Van Goethem C, Kestelyn P, Msellati P, Hitimana D 1991 Clinical and Endocrinologic Manifestations in Perinatally Human Immunodeficiency Virus type-1 Infected Children aged 5 years or older. American Journal of Diseases in Childhood 145:1248-1251

Letendre S, Marquie-Beck J, Capparelli E, Best B, Clifford D, Collier AC, Gelman B, McArthur J, McCutchan A, Morgello S, Simpson D, Grant I, Ellis R, for the CHARTER Group 2008 Validation of the CNS Penetration-Effectiveness Rank for Quantifying Antiretroviral Penetration into the Central Nervous System. Archives of Neurology 65: 65-70

Li T, Tubiana R, Katlama C, Calvez V, Ait Mohand H, Autran B 1998 Long-Lasting Recovery of CD4 T-Cell Function and Viral-Load Reduction After Highly Active Antiretroviral Therapy in Advanced HIV-1 Disease. Lancet 351:1682-86

Lindsey J, Malee K, Brouwers P, Hughes M for the PACTG 219C Study Team 2007 Neurodevelopmental Functioning In HIV-Infected Infants and Young Children Before and After The Introduction Of Protease-Inhibitor-Based Highly Active Antiretroviral Therapy. Pediatrics 119:e681-693 (accessed 12th August 2009)

Lindsey J, Hughes M, McKinney R, Cowles M, Englund J, Baker C, Burchett S, Kline M, Kovacs A, Moye J 2000 Treatment-Mediated Changes In Human Immunodeficiency Virus (HIV) Type 1 RNA and CD4 Cell Counts as Predictors of Weight Growth Failure, Cognitive Decline, and Survival In HIV-Infected Children. Journal of Infectious Diseases182:1385-1393

Linver M, Brooks-Gunn J 2002 Family Processes as Pathways From Income to Young Children's Development. Developmental Psychology 38 (5):719-734

Lipton S, Gendelman H 1995 Dementia Associated with the Acquired Immunodeficiency Syndrome. New England Journal of Medicine 332:934-940

Little K, Thorne C, Luo C, Bunders M, Ngongo N, McDermott P, Newell M 2007 Disease Progression in Children with Vertically-Acquired HIV Infection in Sub-Saharan Africa: Reviewing the Need for HIV Treatment. Current HIV Research: 5:139-153

Llorente A, Brouwers P, Magder L, Mellins C, Ware J, Hittleman J, Mofenson L, Velez-Borras J, Adeniyi-Jones S 2003 Early Neurodevelopmental Markers Predictive of Mortality in Infants Infected with HIV-1 Developmental Medicine and Child Neurology 45:76-84

Lugada ES, Mermin J, Kaharuza F, Ulvestad E, Were W, Langeland N, Asjo B, Malamba S, Downing R 2004 Population-Based Hematologic and Immunologic Reference Values for a Healthy Ugandan Population. Clinical and Diagnostic Laboratory Immunology 11(1):29-34

Luzuriaga K and Sullivan J 2002 Pediatric HIV-1 Infection: Advances and Remaining Challenges. AIDS Review 4:21-26

Lyman W, Kress Y, Kure K, Rashbaum W, Rubinstein A, Soeiro R 1990 Detection of HIV in Fetal Central Nervous System Tissue. AIDS 4:917-920

Macmillan C, Magder L, Brouwers P, Chase C, Hittelman J, Lasky T, Malee K, Mellins C, Velez-Borras J for the Women and Infants Transmission Study 2001 Head Growth and Neurodevelopment of Infants Born to HIV-1 Infected Drug-Using Women. Neurology 57:1402-1411

Mandelbrot L, Le Chenadec J, Berrebi A, Bongain A, Benifla J, Delfraissy J, Blanche S, Mayaux M for the French Perinatal Cohort 1998 Perinatal HIV-1 Transmission-Interaction between Zidovudine Prophylaxis and Mode of Delivery in the French Perinatal Cohort. Journal of the American Medical Association 280 (1):55-60

Marcotte T, Deutsch R, McCutchan J, Moore D, Letendre S, Ellis R, Wallace M, Heaton R, Grant I 2003 Prediction of Incident Neurocognitive Impairment by Plasma HIV RNA and CD4 levels Early after HIV Seroconversion. Archives of Neurology 60: 1406-12

Marks K, Glascoe F, Aylward G, Shevell M, Lipkin P, Squires J 2008 The Thorny Nature of Predictive Validity Studies on Screening Tests for Developmental-Behavioural Problems. Pediatrics 122:866-868

Martin K, Hoover D, Wagoner E, Wingler T, Evans T, O'Brien J, Zeunik J 2009 Development and Reliability of an Observational Gait Analysis Tool for Children with Down Syndrome. Pediatric Physical Therapy 21:261-268

Martin S, Wolters P, Toledo-Tamula M, Zeichner S, Hazra R, Civitello L 2006 Cognitive Functioning In School-Aged Children with Vertically Acquired HIV Infection Being Treated With Highly Active Antiretroviral Therapy (HAART). Developmental Neuropsychology 30:633-657

Martorell R, Horta B, Adair L, Stein A, Richter L, Fall C, Bhargava S, Boswas S, Perez L, Barros F, Victora C and the Consortium on Health Orientated Research in Transitional Societies Group 2010 Weight Gain in the First Two Years of Life Is an Important Predictor of Schooling Outcomes in Pooled Analyses from Five Birth Cohorts from Low- and Middle-Income Countries. The Journal of Nutrition 140:348-354

Masliah E, Deteresa R, Mallory M, Hansen L 2000 Changes in Pathological Findings at Autopsy in AIDS Cases for the last 15 Years. AIDS 14:69-74

Masliah E, Heaton R, Marcotte T, Ellis R, Wiley C, Mallory M, Achim C, McCutchan J, Nelson J, Atkinson J, Grant I and the HNRC Group 1997 Dendritic Injury is a Pathological Substrate for Human Immunodeficiency Virus-related Cognitive Disorders. HNRC Group. Annals of Neurology 42:963-972

Mason E 2006 Positioning Paediatric HIV in the Child Survival Agenda. Presentation to UNICEF−WHO Consultation. January 11−13 New York, NY, UNICEF

Mastaglia B, Toye C, Kristjanson L 2003 Ensuring Content Validity in Instrument Development: Challenges and Innovative Approaches. Contemporary Nurse 14:281-291

Matarazzo P, Palomba E, Lala R, Ciuti E, Altare F, deSanctis L, Tovo P 1994 Growth Impairment, IGF-1 Hyposecretion and Thyroid Dysfunction in Children with Perinatal HIV-1 Infection. Acta Paediatrica 83:1029-1034

Mazzoni P, Chiriboga C, Millar W, Rogers A 2000 Intracerebral Aneurysms in Human Immuno-deficiency Virus Infection: Case Report and Literature Review. Pediatric Neurology 23:252-255

McClowry, D 2000 Development and Assessment of School-Age and Adolescent Children with Human Immunodeficiency Virus. Seminars in Speech and Language, 21:49-60

McEwan I, Arnold S, Hansen L, Johnson D 2003 Interrater Reliability and Content Validity of a Minimal Data Set to Measure Outcomes of Students Receiving School-Based Occupational Therapy and Physical Therapy. Physical and Occupational Therapy in Pediatrics 23(3):77-96

McFarland E 2005 The Immunology of Pediatric HIV Disease. In: Zeichner S, Read J 2005 Text Book of Pediatric HIV Care. Cambridge

McGrath N, Fawzi W, Bellinger D, Robins J, Msamanga G, Manjii K, Tronick E 2006a The Timing of Mother-to-Child Transmission of Human Immunodeficiency Virus Infection and the Neurodevelopment of Children in Tanzania. The Pediatric Infectious Disease Journal 25:47-52

McGrath N, Bellinger D, Robins J, Msamanga G, Tronick E, Fawzi W 2006b Effect of Maternal Multivitamin Supplementation on the Mental and Psychomotor Development of Children who are born to HIV-1-Infected Mothers in Tanzania. Pediatrics 117(2): e216 – 225 (accessed august 2009)

McKinney R, Johnson G, Stanley K, Yong F, Keller A, O'Donnell K, Brouwers P, Mitchell W, Yogev R, Wara D, Wiznia A, Mofenson L, McNamara J, Spector S 1998 A Randomized Study of Combined Zidovudine-Lamivudine versus Didanosine Monotherapy in Children with Symptomatic Therapy-Naive HIV-1 Infection. Journal of Pediatrics 133:500-508

McLoyd V 1998 Socioeconomic Disadvantage and Child Development. American Psychologist 53:185-204

McPhillips M, Jordan-Black J 2007 The Effect of Social Disadvantage on Motor Development in Young Children: A Comparative Study. Journal of Child Psychology and Psychiatry 48 (12):1214-1222

Meisels S 1989 Can Developmental Screening Tests Identify Children who are Develpmentally at Risk? Pediatrics 83:578-585

Mekmullica J, Brouwers P, Charurat M, Paul M, Shearer W, Mendez H, Diaz C, Read J, Mondal P, Smith R, McIntosh K 2009 Early Immunological Predictors of Neurodevelopmental Outcomes in HIV-Infected Children. Clinical Infectious Diseases 48:338-46

Meyers T, Moultrie H, Naidoo K, Cotton M, Eley B, Sherman G 2007 Challenges to Pediatric HIV Care and Treatment in South Africa. The Journal of Infectious Diseases 196:S474-81

Meyers T, Pettifor J, Gray G, Crewe-Brown H, Galpin J 2000 Pediatric Admissions with Human Immunodeficiency Virus Infection at a Regional Hospital in Soweto, South Africa. Journal of Tropical Pediatrics 46 (4):224-230

Mintz M 1999 Clinical Features and Treatment Interventions for Human Immunodeficiency Virus- Associated Neurologic Disease in Children. Seminars in Neurology 19(2):165-176

Missmer S, Spiegelman D, Gorbach S, Miller T 2000 Predictors of Change in the Functional Status of Children With Human immunodeficiency Virus Infection. Pediatrics 106:e24 (Accessed 08/04)

Mitchell W 2001 Neurological and Developmental Effects of HIV and AIDS in Children and Adolescents. Mental Retardation and Developmental Disabily Research Reviews 7:211-216.

Msellati P, Lepage P, Hihama D, Van Goetham C, Van de Perre P, Dabis F 1993 Neurodevelopmental Testing of Children born to Human Immunodeficiency Virus Type 1 Seropositive and Seronegative Mothers: A Prospective Cohort in Kigali, Rwanda. Pediatrics 92 (6):843-848

Nachman S, Lindsay J, Pelton S, Mofenson L, McIntosh K, Wiznia A, Stanley K Yogev R 2002 Growth in Human Immunodeficiency Virus-Infected Children Receiving Ritonavir-Containing Antiretroviral Therapy. Archives of Pediatric and Adolescent Medicine 156:497-503

Nachman S, Stanley K, Yogev R, Pelton S, Wiznia A, Lee S, Mofenson L, Fiscus S, Rathore M, Jimenez E, Borkowsky W, Pitt J, Smith M, Wells B, McIntosh K 2000 Nucleoside Analogs plus Ritonavir in Stable Antiretroviral Therapy-Experienced HIV-Infected Children: A Randomized Controlled Trial: Pediatric AIDS Clinical Trials Group 338 Study Team. Journal of the American Medical Association 283:492-498

Nadal D, Steiner F, Cheseaux JJ, Lazarevitch C, Aebi C, Kind C, Rudin C and the Pediatric AIDS Group of Switzerland 2000 Long-term Responses to Treatment Including Ritonavir or Nelfinavir in HIV-1-Infected Children: Pediatric AIDS Group of Switzerland. Infection 28:287-296

Nannan N, Norman R, Hendricks M, Dhansay MA, Bradshaw D, Group SACRAC 2007 Estimating the Burden of Disease Attributable to Childhood and Maternal Undernutrition in South Africa in 2000. South African Medical Journal 97:733-739

Nathan L, Nerlander L, Dixon J, Ripley R, Barnabas R, Wholeben B, Musoke R, Palakudy T, D'Agostino A, Chakraboty 2003 Growth, Morbidity and Mortality in a Cohort of Institutionalised HIV-1-Infected African Children. Journal of Acquired Immune Deficiency Syndrome 34(2):237-241

National Strategic Plan for HIV 2007 – 2011. www.doh.gov.za/docs/misc/stratplan

Navia B, Cho E, Petito C, Price R 1986 The AIDS Dementia Complex: II. Neuropathology. Annals of Neurology 19:525-535

Nduati R, Mbori-Ngacha D, Richardson B, Panteleeff D, Mwatha A, Overbaugh J, Bwayo J, Ndinya-Achola J, Kreiss J 2001 Correlates of Mother-to-Child Human Immunodeficiency Virus Type 1 (HIV-1) Transmission: Association with Maternal Plasma HIV-1 RNA Load, Genital HIV-1 DNA Shedding, and Breast Infections. Journal of Infectious Diseases 183:206-212

Nduati R, John G, Mbori-Ngacha D, Richardson B, Overbaugh J, Mwatha A, Ndinya-Achola J, Bwayo J, Onyango F, Hughes J, Kreiss J 2000 Effect of Breastfeeding and Formula Feeding on Transmission of HIV-1: A Randomized Clinical Trial. Journal of the American Medical Association 283:1167-74

Newell M, Coovadia H, Cortina-Borja M, Rollins N, Gaillard P, Dabis F, Ghent International AIDS Society (IAS) Working Group on HIV Infection in Women and Children 2004 Mortality of Infected and Uninfected Infants Born to HIV-Infected Mothers in Africa: A Pooled Analysis. Lancet 364:1236-43

Newell, M 1998 Mechanisms and Timing of Mother-to-Child Transmission of HIV-1. AIDS 12:831-7

Niccols A, Latchman A 2002 Stability of the Bayley Mental Scale of Infant Development with High Risk Infants. The British Journal of Developmental Disabilities 48 (94):3-13

Nozyce M, Hitteman J, Muenz L, Durako C, Fischer M, Willoughby A 1994 Effect of Perinatally Acquired Human Immunodeficiency Virus Infection on Neurodevelopment in children during the first two years of life. Pediatrics 94 (6):883-91

Nunnally, J and Bernstein I. Psychometric Theory New York: McGraw Hill, 3rd ed 1994

Obimbo E, Mbori-Ngacha D, Ochieng J, Richardson B, Otieno P, Bosire R, Farquhar C, Overbaugh J, John-Stewart G 2004 Predictors of Early Mortality in a Cohort of Human Immunodeficiency Virus Type 1-Infected African Children. Pediatric Infectious Disease Journal. 23:536-543

Olness K 2003 Effects on Brain Development Leading to Cognitive Impairments: A Worldwide Pandemic. Journal of Developmental and Behavioural Pediatrics 24:120-30

Palisano R, Rosenbaum P, Bartlett D, Livingston M 2008 Content Validity of the Expanded and Revised Gross Motor Function Classification System. Developmental Medicine and Child Neurology 50:744-750

Papola P, Alvarez M, Cohen H 1994 Developmental and Service Needs of School-Age Children with Human Immunodeficiency Virus Infection: A Descriptive Study. Pediatrics 94(6):914-918

Parks R, Cintas H, Chaffin M, Gerber L 2007 Brief Assessment of Motor Function: Content Validity and Reliability of the Fine Motor Scale. Pediatric Physical Therapy 19: 315-325

Parks R, Danoff J 1999 Motor Performance Changes in Children Testing Positive for HIV over 2 Years. American Journal of Occupational Therapy 53:524-528

Patel K, Ming X, Williams P, Robertson K, Oleske J, Seage G 2009 Impact of HAART and CNS-Penetrating Antiretroviral Regimens on HIV Encephalopathy Among Perinatally Infected Children and Adolescents. AIDS 23:1-9

Pearson D, McGrath N, Nozyce M, Nichols S, Raskino C, Brouwers P, Lifschitz M, Baker C, Englund J 2000 Predicting HIV Disease Progression in Children Using Measures of Neuropsychological And Neurological Functioning. Pediatrics 106: e76 http://www.pediatrics.org/cgi/content/full/106/6/e76 (accessed April 2004)

Peckham C, Gibb D 1995 Mother-to-Child Transmission of Human Immunodeficiency Virus. New England Journal of Medicine 333: 298 - 303

Pelto G, Dickin K, Engel P 1999 A Critical Link, Interventions for Physical Growth and Psychological Development: A Review. World Health Organization; 1999 www.who.int (accessed 02/07)

Peluso R, Haase A, Stowring L, Edwards M, Ventrura P 1985 A Trojan Horse Mechanism for Spread of Visna Virus in Monocytes. Virology 147: 231 – 236

Peterson NJ, Drotar D, Olness K, Guay L, Kiziri-Mayengo R 2001 The Relationship Of Maternal And Child HIV Infection To Security Of Attachment Among Ugandan infants. Child Psychiatry and Human Development 32:3-17

Piper M, Darrah J. Motor Assessment of the Developing Infant. Philadelphia: WB Saunders; 1994

Piper M, Pinnell L, Darrah J, Maguire T, Byrne P 1992 Construction and Validation of the Alberta Infant Motor Scale. Canadian Journal of Public Health. 83(Suppl 2):S46-S50

Pizzo P, Eddy J, Falloon J, Balis F, Murphy R, Moss H, Wolters P, Brouwers P, Jarosinski P, Rubin M, Broder S, Yarchoan R, Brunetti A, Maha M, Nusinoff-Lehrman S, Poplack D 1988 Effect of Continuous Intravenous Infusion of Zidovudine (AZT) in Children with Symptomatic HIV Infection. New England Journal of Medicine 319 (14):889–896

Pollack H, Kuckuk A, Cowan L, Hacimamutoglu S, Glasberg H, David R, Krasinski K, Borkowsky W, Oberfield S 1996 Neurodevelopment, Growth and Viral Load in HIV-1 Infected Infants. Brain, Behaviour and Immunity 10:298-312

Polit D, Beck C 2006 The Content Validity Index: Are You Sure You Know What's Being Reported? Critique and Recommendations. Research in Nursing and Health 29:489-497

Pollitt E, Watkins W, Husaini M 1997 Three Months Nutritional Supplementation in Indonesian Infants and Toddlers Benefit Memory Function 8 Years Later. American Journal of Clinical Nutrition 54:799-804

Portney L, Watkins M Statistical Measures of Reliability. In: Portney LG, Watkins MP, eds. Foundations of Clinical Research: Applications to Practice, 2nd ed. Upper Saddle River, NJ: Prentice Hall Health; 2000:557-586

Potter M, Gordon S, Hamer P 2004 The Nominal Group Technique: A Useful Consensus Methodology in Physiotherapy Research. New Zealand Journal of Physiotherapy 32(3): 126-130

Potterton J, Stewart A, Cooper P, Becker P 2009a The Effect of a Home Stimulation Programme on the Development of Young Children Infected with HIV. Developmental Medicine and Child Neurology http://dx.doi.org/10.1111/j.1469-8749.2009.03534.x (accessed April 2010)

Potterton J, Stewart A, Cooper P, Baillieu N, Gajdosik C 2009b Neurodevelopmental Delay in Children Infected with HIV in Soweto, South Africa. Vulnerable Children and Youth Studies March 4 (1):48-57

Potterton J, Stewart A, Cooper P 2007 Parenting Stress of Caregivers of Young Children who are HIV Positive. African Journal of Psychiatry 10:210-214

Potterton J, Eales C 2001 Prevalence of Developmental Delay in Infants who are HIV Positive. South African Journal of Physiotherapy 57(3):11-15

Powell C, Baker-Henningham H, Walker S, Gernay J, Grantham-McGregor S 2004 Feasibility of Integrating Early Stimulation into Primary Care for Undernourished Jamaican Children: Cluster Randomised Controlled Trial. Britsh Medical Journal 329: 89-93

Pratt R 2003 Retroviruses, the cause of AIDS. In HIV and AIDS: A Foundation for Nursing and Healthcare Practise. 5th Edition pg 25, 27. Arnold, London

Prendergast A, Tudor-Williams G, Jeena P, Burchett S, Goulder P 2007 International Perspectives, Progress, and Future Challenges of Paediatric HIV Infection. Lancet 370:68-80

Price R, Brew B, Sidtis J, Rosenblum M, Scheck A, Cleary P 1988 The Brain in AIDS: Central Nervous System HIV-1 Infection and AIDS Dementia Complex. Science 239: 586-592

Pulliam L, Herndier B, Tang N, McGrath M 1991 Human Immunodeficiency Virus-Infected Macrophages Produce Soluble Factors that Cause Histological and Neurochemical Alterations in Cultured Human Brains. Journal of Clinical Investigation 87:503-512

Rademeyer V 2010 A Study to Evaluate the Performance of Black South African Urban Infants on the Bayley Scales of Infant Development III. Unpublished Mmed thesis. University of the Witwatersrand

Raskino C, Pearson D, Baker C, Lifschitz M, O'Donnell T, Nozyce M, Brouwers P, McKinney R, Jiminez E, Englund J 1999 Neurologic, Neurocognitive and Brain Growth Outcomes in Human Immuno-Deficiency Virus-Infected Children Receiving Different Nucleoside Antiretroviral Regimens. Pediatrics 104:e32 http://www.pediatrics.org/cgi/content/full/104/3/e32 (Accessed 08/04)

Rausch D, Stover E 2001 Neuroscience Research in AIDS. Progress in Neuro-Psychopharmacology and Biological Psychiatry 25:231-257

Reddi A, Leeper S, Grobler A, Geddes R, France H, Dorse G, Vlok W, Mntambo M, Nixon K, Holst H, Karim A, Rollins N, Coovadia H, Giddy J 2007 Preliminary Outcomes of a Paediatric Highly Active Antiretroviral Therapy Cohort from KwaZulu-Natal, South Africa. BMC Pediatrics 7:13 http://www.biomedcentral.com/1471-2431/7/13 (accessed July 2009)

Reid D, Rigby P 1997 Critique of the Bayley Infant Neurodevelopmental Screener (BINS). Physical and Occupational Therapy in Pediatrics 17(1):59-74

Resino S, Resino R, Micheloud D, Gutie D, Antonio J, Ramos J, Ciria L, de Jose I, Mellado J, Mun A 2006 Long-Term Effects of Highly Active Antiretroviral Therapy in Pretreated, Vertically HIV Type 1– Infected Children: 6 Years of Follow-Up. Clinical Infectious Diseases 42:862-9

Resnick L, Veronese F, Schupbach J, Tourtellotte W, Ho D, Muller F, Shashak P, Vogt M, Groopman J, Markham P, Gallo R 1985 Intra-Blood-Brain-Barrier Synthesis of HTLV-III-Specific IgG in Patients with Neurologic Symptoms associated with AIDS or AIDS-related complex. New England Journal of Medicine 313:1498-1504

Richter L, Griesel R, Rose C 1992 The Bayley Scales of Infant Development-A South African Standardisation. The South African Journal of Occupational Therapy May:14-25

Richter LM, Griesel RD 1988 BSID Norms for Interpreting the Performance of Black South African Infants. Pretoria, University of South Africa.

Robertson K, Kopnisky K, Hakim J, Merry C, Nakasujja N, Hall C, Traore M, Sacktor N, Clifford D, Newton C, Van Rie A, Holding P, Clements J, Zink C, Mielke J, Hosseinipour M, Lalloo U, Amod F, Marra C, Evans S, Liner J 2008 Second Assessment of NeuroAIDS in Africa. Journal of Neurovirology 14:89-101

Rosenbaum P, Missiuna C, Echeverria D, Knox D 2009 Proposed Motor Development Assessment Protocol for Epidemiological Studies in Children. Journal of Epidemiology and Community Health 63 (Suppl I): i27-i36 (Accessed August 2009)

Rosenbaum P 1998 Screening Tests and Standardized Assessments Used to Identify and Characterize Developmental Delays. Seminars in Pediatric Neurology 5(1):27-32

Rosenbaum P, Russell D, Cadman D, Gowland C, Jarvis S, Hardy S 1990 Issues in Measuring Change in Motor Function in Children with Cerebral Palsy: A Special Communication. Physical Therapy 70:125-131

Rothstein J, Campbell S, Echternach J, Jette A, Knecht H, Rose S 1991 Standards for Tests and Measurements in Physical Therapy Practice. Physical Therapy 71:598-622

Russell D, Rosenbaum P, Cadman D, Gowland C, Hardy S, Jarvis S 1989 The Gross Motor Function Measure: A Means to Evaluate the Effects of Physical Therapy. Developmental Medicine and Child Neurology 31:341-352

Ryder R, Nsa W, Hassig S, Behets F, Rayfield M, Ekungola B, Nelson A, Mulenda U, Francis H, Mwandagalirwa K, Davachi F, Rogers M, Nzilambi N, Greenberg A, Mann J, Quinn T, Piot P, Curran J 1989 Perinatal Transmission of the Human Immunodeficiency Virus Type 1 to Infants of Seropositive Women in Zaire. New England Journal of Medicine 320 (25):1637-42

Rydz D, Srour M, Oskoui M, Marget N, Shiller M, Birnbaum R, Majnemer A, Shevell M 2006 Screenng for Developmental Delay in the Setting of a Community Pediatric Clinic: A Prospective Assessment of Parent-Report Questionnaires. Pediatrics 118: 1178-1186

Saigal S, Rosenbaum P, Stoskopf B, Hoult L, Furlong W, Feeny D, Hagan R 2005 Development, Reliability and Validity of a New Measure of Overall Health for Pre-school Children. Quality of Life Research 14:243-257

Sanchez-Ramon, Resino S, Bellon J, Ramos J, Gurbindo D, Munoz-Fernandez A 2003 Neuroprotectiveeffects of Early Antiretrovirals in Vertical HIV-infection. Pediatric Neurology 29:218-221

Sand N, Silverstein M, Galscoe F, Gupta V, Tonniges T, O'Connor K 2005 Pediatricians' Reported Practices Regarding Developmental Screening: Do Guidelines Work? Do They Help? Pediatrics 116:174-179

Schor E, Billingsley M, Golden A, McMillan J, Meloy L, Pendarvis B 2003 Family Pediatrics: Report of the Task Force on the Family. Pediatrics 111(6):1541-1571

Scott V, Votova K, Scanlan A, Close J 2007 Multifactorial and Functional Mobility Assessment Tools for Fall Risk Among Older Adults in Community, Home-Support, Long-Term and Acute Care Settings. Age and Ageing 36:130-139

Sei S, Saito K, Stewart S, Crowley J, Brouwers P, Kleiner D, Katz D, Pizzo P, Heyes M 1995 Increased Human Immunodeficiency Virus (HIV) Type 1 DNA Content and Quinlinic Acid Concentration in Brain Tissues from Patients with HIV Encephalopathy. The Journal of Infectious Diseases 172:638-647

Sharer L, Epstein L, Cho E, Joshi V, Meyenhofer M, Rankin L, Petito C 1986 Pathologic Features of AIDS Encephalopathy in Children: Evidence of LAV/HTLV-III Infection of Brain. Human Pathology 17:271-284

Sharland M, di Zub GC, Ramos JT, Blanche S, Gibb DM 2002 PENTA Guidelines for the Use of Antiretroviral Therapy in Paediatric HIV Infection. HIV Medicine 3:215-26

Shea B, Grimshaw J, Wells G, Boers M, Andersson N, Hamel C, Porter A, Tugwell P, Moher D, Bouter L 2007 Development of AMSTAR: A Measurement Tool to Assess the Methodological Quality of Systematic Reviews. BMC Medical Research Methodology 7:10

Siegel E, Stoltzfus R, Kariger P, Katz J, Khatry S, LeClerq S, Pollitt E, Tielsch J 2005 Growth Indices, Anemia and Diet Predict Motor Milestones Acquisition of Infants in South Central Nepal. Journal of Nutrition135:2840-4

Silva A, Metha Z, O'Callaghan F 2005 The Relative Effect Size at Birth, Postnatal Growth and Social Factors on Cognitive Function in Late Childhood. Annals of Epidemiology16:469-76

Sim J, Arnell P 1993 Measurement Validity in Physical Therapy Research. Physical Therapy 73:102-115

Simpson D, Berger J 1996 Neurologic Manifestations of HIV Infection. Medical Clinics of North America 80 (6):1363-94

Smith M, Danoff J, Parks R 2002 Motor Skill Development of Children with HIV Infection Measured with the Peabody Developmental Motor Scales. Pediatric Physical Therapy 14:74-84

Smith R, Malee K, Leighty R, Brouwers P, Mellins C, Hittelman J, Chase C, Blasini I, 2006 Effects of Perinatal HIV Infection and Associated Risk Factors on Cognitive Development Among Young Children. Pediatrics 117:851-862

Smith R, Malee K, Charurat Magder L, Mellins C, Macmillan C, Hittleman J, Lasky T, Llorente A, Moye J 2000 Timing of Perinatal Human Immunodeficiency Virus Type 1 Infection and Rate of Neurodevelopment. The Pediatric Infectious Disease Journal 19 (9):862-871

Smits-Engelsman B, Fiers M, Henderson S, Henderson L 2008 Interrater Reliability of the Movement Assessment Battery for Children. Physical Therapy 88:286-294

Soh C, Oleske J, Brady M, Spector S, Borkowsky W, S Burchett, Foca M, Handelsman E, Jiménez E, Dankner W, Hughes M, for the Pediatric AIDS Clinical Trials Group 2003 Long-term Effects of Protease-Inhibitor-Based Combination Therapy on CD4 T-cell Recovery in HIV-1-Infected Children and Adolescents. Lancet 362:2045-2051

Spector S, Hsia K, Pratt D, Lathey J, McCutchan J, Alcaraz J, Atkinson J, Gulevich S, Wallace M, Grant I 1993 Virologic Markers of Human Immunodeficiency Virus Type 1 in Cerebrospinal Fluid. The HIV Neurobehavioral Research Centre Group. Journal of Infectious Diseases 168:68-74

Spira S, Lepage P, Msellati P, Van de Perre P, Leroy V, Simonon A, Karita E, Dabis F for the Mother-to –Child HIV-1 transmission study group. 1999 Natural History of Human Immunodeficiency Virus Type1 Infection in Children: A Five Year Prospective Study in Rwanda. Pediatrics 104(5):e56

Spurr G 1983 Nutritional Status and Physical Work Capacity. Yearbook of Physical Anthropology 1–35

Squires J, Bricker D, Potter L 1997 Revision of a Parent-Completed Developmental Screening Tool: Ages and Stages Questionnaires. Journal of Pediatric Psychology 22: 313-328

Sternberg, R, Nokes C, Geissler P, Prince R, Okatcha F, Bundy D, Grigorenko E 2001 The Relationship Between Academic and Practical Intelligence: A Case Study in Kenya. Intelligence 29:401-418

Storm D, Boland M, Gortmaker S, He Y, Skurnick J, Howland L, Oleske J for the Pediatric AIDS Clinical Trials Group 2005 Protease Inhibitor Combination Therapy, Severity of Illness, and Quality of Life. Pediatrics 115:173-182

Sutcliffe C, van Dijk J, Bolton C, Persaud D, Moss W 2008 Effectiveness of Antiretroviral Therapy Among HIV-Infected Children in Sub-Saharan Africa. Lancet (8):477-489

Taha T, Graham S, Kumwenda N, Broadhead R, Hoover D, Markarkis D, van der Hoever L, Liomba G, Chiphangwi J, Miotti P 2000 Morbidity Among Human Immunodeficiency Virus-1-Infected and Uninfected African Children. Pediatrics 106 (6): e77 (accessed August 2009)

Takahashi K, Wesselingh S, Griffin D, McArthur J, Johnson R, Glass J 1996 Localisation of HIV-1 in Human Brain using Polymerase Chain Reaction/in situ Hybridization and Immunocytochemistry. Annals of Neurology 39:705-711

Tardieu M, Le Chenadec M, Persoz A, Meyer L, Blanche S, Mayaux M 2000 HIV-1 related Encephalopathy in Infants Compared With Children and Adults. Neurology 54: 1089-1095

Tardieu M, Mayaux M, Seibel N, Funck-Brentano I, Straub E, Teglas J, Blanche S 1995 Cognitive Assessment of School-age Children Infected with Maternally Transmitted Human Immunodeficiency Virus Type-1. Journal of Pediatrics 126: 375-9

Tardieu M, Hery C, Peudenier S, Boespflug O, Montagnier L 1992 Human Immunodeficiency Virus Type-1-Infected Monocytic Cells can Destroy Human Neural Cells After Cell-to-Cell Adhesion. Annals of Neurology 32:11-17

Thomas-Stonell N, Oddson B, Robertson B, Rosenbaum P 2010 Development of the FOCUS (Focus on the Outcomes of Communication Under Six), a Communication Measure for Preschool Children. Developmental Medicine and Child Neurology 52:47-53

Thorne C and Newell N 2000 Epidemiology of HIV Infection in the Newborn. Early Human Development 58:1-16

Tieman B, Palisano R, Sutlive A 2005 Assessment of Motor Development and Function in Preschool Children. Mental Retardation and Developmental Disabilities Research Reviews 11:189-196

Tindyebwa D, Kayita J, Musoke P, Eley B, Nduati R, Coovadia H, Bobart R, Mbori-Ngacha D, Kieffer M 2004 Handbook on Paediatric AIDS in Africa by the African Network for the Care of Children Affected by AIDS. www.rcqhc.org (Accessed July 2009)

Tornatore C, Chandra R, Berger J, Major E 1994 HIV-1 Infection of Subcortical Astrocytes in the Pediatric Central Nervous System. Neurology 44:481-487

Tovo P, de Martino M, Gabiano C, Galli L, Cappello N, Ruga E 1996 Mode of Delivery and Gestational Age influence Perinatal HIV-1 Transmission. Journal of Acquired Immune Deficiency Syndrome and Human Retrovirology 11:88-94

Tovo P, de Martino M, Gabiano C, Cappello N, D'Elia R, Loy A, Plebani A, Zuccotti G, Dallacasa P, Ferraris G, Caselli D, Fundaro C, D'Argenio P, Galli L, Principi N, Stegagno M, Ruga E, Palombia E 1992 Prognostic Factors and Survival in Children with Perinatal HIV-1 Infection. The Italian Register for HIV Infections in Children. Lancet 339:1249-1253

Trujillo R, Rogers R, Molina R, Dangonds F, McLane M, Essex M, Brain J 2007 Noninfectious Entry of HIV-1 into Peripheral and Brain Macrophages Mediated by the Mannose Receptor. Proceedings of the National Academy of Sciences 104(12): 5097-5102

Udagawa N, Takahashi N, Katagiri T, Tamura T, Wada S, Findlay D, Martin T, Hirota H, Taga T, Kishimoto T, Suda T 1995 Interleukin (IL)-6 Induction of Osteoclast Differentiation Depends on IL-6 Receptors Expressed on Osteoblastic cells but not on Osteoclast Progenitors. Journal of Experimental Medicine 182:146-1468

Udgirkar V, Tullu M, Bavdekar S, Shaharao V, Kamat J, Hira P 2003 Neurological Manifestations of HIV Infection. Indian Pediatrics 40:230-234

Ultmann M, Belman A, Ruff H 1985 Developmental Abnormalities in Infants and Children with Acquired Immune Deficiency Syndrome (AIDS) and AIDS-related Complex. Developmental Medicine and Child Neurology 27:563- 571

UNAIDS 2009 AIDS Epidemic Update. www.unaids.org (accessed 02/10)

UNAIDS 2008 Epidemiological Fact Sheet on HIV and AIDS Core Data on Epidemiology and Response: South Africa. www.unaids.org (accessed 08/09)

UNAIDS Overview of the global AIDS epidemic: 2006 Report on the Global AIDS Epidemic. www.UNAIDS.org. (Accessed June 2007)

UNAIDS Update 2005 http://www.unaids.org (Accessed September 2006)

Van de Vijver F, Tanzer N 2004 Bias and Equivalence in Cross-Cultural Assessment: An Overview. European Review of Applied Psychology 54:119-135

Van de Vijver F 1997 Meta-Analysis of Cross-Cultural Comparisons of Cognitive Test Performance. Journal of Cross Cultural Psychology 28:678-709

Van den Broek J, Willie D, Younger N 2009 The World Health Organisation Child Growth Standards: Expected Implications for Clinical and Epidemiological Research. European Journal of Pediatrics 168:247-251

Van der Camp K, Vernooij-Dassen M, Grol R, Bottema B 2006 Professionalism in General Practice: Development of an Instrument to Assess Professional Behaviour in General Practitioner Trainees. Medical Education 40:43-50

Van der Loeff S, Hansmann A, Awasana A, Ota M, O'Donovan D, Sarge-Njie, R, Ariyoshi K, Milligan P, Whittle H 2003 Survival of HIV-1 and HIV-2 Perinatally Infected Children in The Gambia. AIDS 17: 2389–94

Van Dyke R, Wang L, Williams P 2008 Toxicities Associated with Dual Nucleoside Reverse Transcriptase Inhibitor Regimens in HIV-infected Children. Journal of Infectious Diseases 198:1599-1608

Van Dyke R, Korber B, Popek E, Macken C, Widmayer A, Bardeguez A, Hanson C, Wiznia A, Luzuriaga K, Viscarello R, Wolinsky S and the Ariel Core Investigators 1999 The Ariel Project: A Prospective Cohort Study of Maternal–Child Transmission of Human Immunodeficiency Virus Type-1 in the era of Maternal Antiretroviral Therapy. Journal of Infectious Diseases 179:319-28

Van Rie A, Dow A, Mapula A, Stewart P 2009 Neurodevelopmental Trajectory of HIV-Infected Children Accessing Care in Kinshasa, Democratic Republic of the Congo. Journal of the Acquired Immune Deficiency Syndrome 00:000

Van Rie A, Mupuala A and Anna Dow 2008 Impact of the HIV/AIDS Epidemic on the Neurodevelopment of Preschool-Aged Children in Kinshasa, Democratic Republic of the Congo. Pediatrics 122; e123-e128 (accessed March 2009)

Van Rie A, Harrington P, Dow A, Robertson K 2007 Neurologic and Neurodevelopmental Manifestations of Pediatric HIV/AIDS: A Global Perspective. European Journal of Paediatric Neurology 11:1-9

Verweel G, Rossum A, Hartwig N, Wolfs T, Scherpbier H, de Groot R 2002 Treatment with Highly Active Antiretroviral Therapy in Human Immunodeficiency Virus Type 1-Infected Children Is Associated With a Sustained Effect on Growth. Pediatrics 109: e25 (accessed august 2009)

Victora C, Adair L, Fall C, Hallal P, Martorell R, Richter L, Sachdev H, (for the Maternal and Child Undernutrition Study Group) 2008 Maternal and Child Undernutrition: Consequences for Adult Health and Human Capital. Lancet. 371:340-57

Violari A, Paed F, Cotton M, Gibb D, Babiker A, Steyn J, Madhi S, Paed J, Patrick J, McIntyre J, (for the CHER Study Team) 2008 Early Antiretroviral Therapy and Mortality Among HIV-1 Infected Infants. New England Journal of Medicine 359:2233-44

Walka H, Pollitt E 2000 A Preliminary Test of a Developmental Model for the Study of Undernourished Children in Indonesia. European Journal of Clinical Nutrition 54:S21-7

Walensky R, Wolf L, Wood R, Fofana M, Freedberg K, Martinson N, Paltiel D, Anglaret X, Weinstein M, Losina E for the CEPAC (Cost-Effectiveness of Preventing AIDS Complications)-International Investigators 2009 When to Start Antiretroviral Therapy in Resource-Limited Settings. Annals of Internal Medicine 151:157-166

Walker S, Wachs T, Gardner J, Lazoff B, Wasserman G, Pollitt E, Carter J and the International Child Development Steering Group 2007 Child Development: Risk Factors for Adverse Outcomes in Developing Countries. Lancet 369: 145–57

Walker S, Mulenga V, Sinyinza F, Lishimpi K, Nunn A, Chintu C, Gibb D, and the CHAP trial team 2006 Determinants After Survival Without Anturetroviral Therapy After Infancy in HIV-1-Infected Zambian Children in the CHAP Trial. Journal of Acquired Immune Deficiency Syndrome 42:637-645

Walker S, Chang S, Powell C, Grantham-McGregor S 2005 Effects of Early Childhood Psychosocial Stimulation and Nutritional Supplementation on Cognition and Education in Growth-Stunted Jamaican Children: Prospective Cohort Study. Lancet 366:1804-7

Walker S, Grantham-McGregor S, Powell C, Chang S 2000 Effects of Growth Restriction in Early Childhood on Growth, IQ and Cognition at Age 11 to 12 Years and the Benefits of Nutritional Supplementation and Psychosocial Stimulation. Journal of Pediatrics 137:36-41

Wamalwa D, Farquhar C, Mbori-Ngacha D, Richardson B, Overbaugh J, Emery S, Wariua G, Gichuhi C, Bosire R, John-Stewart G 2007 Early Response to Highly Active Antiretroviral Therapy in HIV-1–Infected Kenyan Children. Journal of Acquired Immune Deficiency Syndrome 45:311-317

Wickett J, Vernon P, Lee D 2000 Relationships Between Factors of Intelligence and Brain Volume. Personality and Individual Differences 29:1095-1122

Wiley C, Schrier R, Nelson J, Lampert P, Oldstone M 1986 Cellular Localization of Human Immunodeficiency Virus Infection within the Brains of Acquired Immune Deficiency Syndrome Patients. Proceedings of the National Academy of Science U.S.A 83:7089-7093

Willen E 2006 Neurocognitive Outcomes in Pediatric HIV. Mental Retardation and Developmental Disabilities Research Reviews 12:223-228

Williams A, Duong T, McNally L 2001 Pneumocystis Carinii Pneumonia and Cytomegalovirus Infection in Children with Vertically Acquired HIV Infection. AIDS 15:335-39

Winter H, Miller T 1994 Gastrointestinal and Nutritional Problems in Pediatric HIV Disease in Pediatric AIDS: the challenge of HIV Infection in Infants, Children and adolescents by Pizzo P, Wilfert C (eds). 2nd Edition pg 513 - 515. Williams and Wilkins. Baltimore, Maryland

Wiznia A, Stanley K, Krogstad P, Johnson G, Lee S, McNamara J, Moye J, Jackson JB, Mendez H, Aguayo R, Dieudonne A, Kovacs A, Bamji M, Abrams E, Rana S, Sever J, Nachman S 2000 Combination Nucleoside Analog Reverse Transcriptase Inhibitor(s) Plus Nevirapine, Nelfinavir, or Ritonavir in Stable Antiretroviral Therapy-Experienced HIV-Infected Children: Week 24 Results of a Randomized Controlled Trial. AIDS Research and Human Retroviruses 16:1113-21

Wiznia A, Lambert G, Pavlakis S 1996 Pediatric HIV Infection. Medical Clinics of North America 80(6):1309-1336

Wolters P, Brouwers P, Civitello L, Moss H 1997 Receptive and Expressive Language Function of Children with Symptomatic HIV Infection and Relationship with Disease Parameters: a Longitudinal 24-month Follow-up Study. AIDS 11:1135-1144

Wolters P, Brouwers P, Moss H, Pizzo P 1995 Differential Receptive and Expressive Language Functioning of Children with Symptomatic HIV Disease and Relation to CT brain abnormalities. Pediatrics 95 (1): 112-119

Working Group on Antiretroviral Therapy and Medical Management of HIV-Infected Children 2009 Guidelines for the Use of Antiretroviral Agents in Pediatric HIV Infection. February: 1-139. http://aidsinfo.nih.gov/ContentFiles/PediatricGuidelines.pdf**.** Accessed (12th August 2009)

World Health Organisation 2007 WHO Case Definitions for Surveillance of HIV and Revised Clinical Staging and Immunological Classification of HIV-Related Disease in Adults and children. www.who.int (accessed 02/10)

World Health Organisation Multicentre Growth Reference Study Group 2006a WHO Child Growth Standards: Length/Height-for-Age, Weight-for-Age, Weight-for-Length, Weight-for-Height and Body Mass Index-for-Age: Methods and Development. Geneva: World Health Organization (312 pages) www.who.int (accessed 04/01/2010)

World Health Organisation 2006b Antiretroviral Therapy of HIV Infection in Infants and Children in Resource-Limited Settings: Towards Universal Access. Recommendations for a Public Health Approach. [http://www.who.int/hiv/pub/guidelines/paediatric020907.pdf] (Accessed July 2009)

World Health Organisation 2006c Interim WHO Clinical Staging of HIV/AIDS and HIV/AIDS Case Definitions for Surveillance. African Region. http://www.who.int/hiv/pub/guidelines/clinicalstaging.pdf (accessed Feb 2010)

World Health Organisation 2005 Interim WHO Clinical Staging of HIV/AIDS and HIV/AIDS Case Definitions for Surveillance. www.who.int (Accessed 03/10)

World Health Organisation 1997 WHO Global Database on Child Growth and Malnutrition. Geneva: WHO http://www.who.int (accessed 03/10)

World Health Organisation Expert Committee 1995 Physical status: The Use and Interpretation of Anthropometry. Geneva: WHO www.who.int (accessed March 2010)

Wynd C, Schmidt B, Schaefer M 2003 Two Quantitative Approaches for Estimating Content Validity. Western Journal of Nursing Research 25(5) 508 - 518

Zeichner S, Read J 2005 Text Book of Pediatric HIV Care. Cambridge

Zink C, Laast V, Helke K, Brice A, Barber S, Clements J, Mankowski J 2006 From Mice to Macaques – Animal Models of HIV Nervous System Disease. Current HIV Research 4:293-305

Appendix 1

Ethical Clearance Certificate

UNIVERSITY OF THE WITWATERSRAND, JOHANNESBURG

Division of the Deputy Registrar (Research)

HUMAN RESEARCH ETHICS COMMITTEE (MEDICAL)
R14/49 Baillieu

CLEARANCE CERTIFICATE	**PROTOCOL NUMBER M070817**
PROJECT	The Evaluation of a Screening Tool to Assess Neurodevelopment in HIV Positive Infants
INVESTIGATORS	Ms N Baillieu
DEPARTMENT	Department of Physiotherapy
DATE CONSIDERED	07.08.31
DECISION OF THE COMMITTEE*	APPROVED UNCONDITIONALLY

Unless otherwise specified this ethical clearance is valid for 5 years and may be renewed upon application.

UNIVERSITY OF THE WITWATERSRAND
PROF AJ WOODIWISS
Co - Chair
2007 -10- 03
HREC (Medical)
JOHANNESBURG

DATE 07.09.27 **CHAIRPERSON** .....
(Professors PE Cleaton-Jones, A Dhai, M Vorster, C Feldman, A Woodiwiss)

*Guidelines for written 'informed consent' attached where applicable

cc: Supervisor : J Potterton

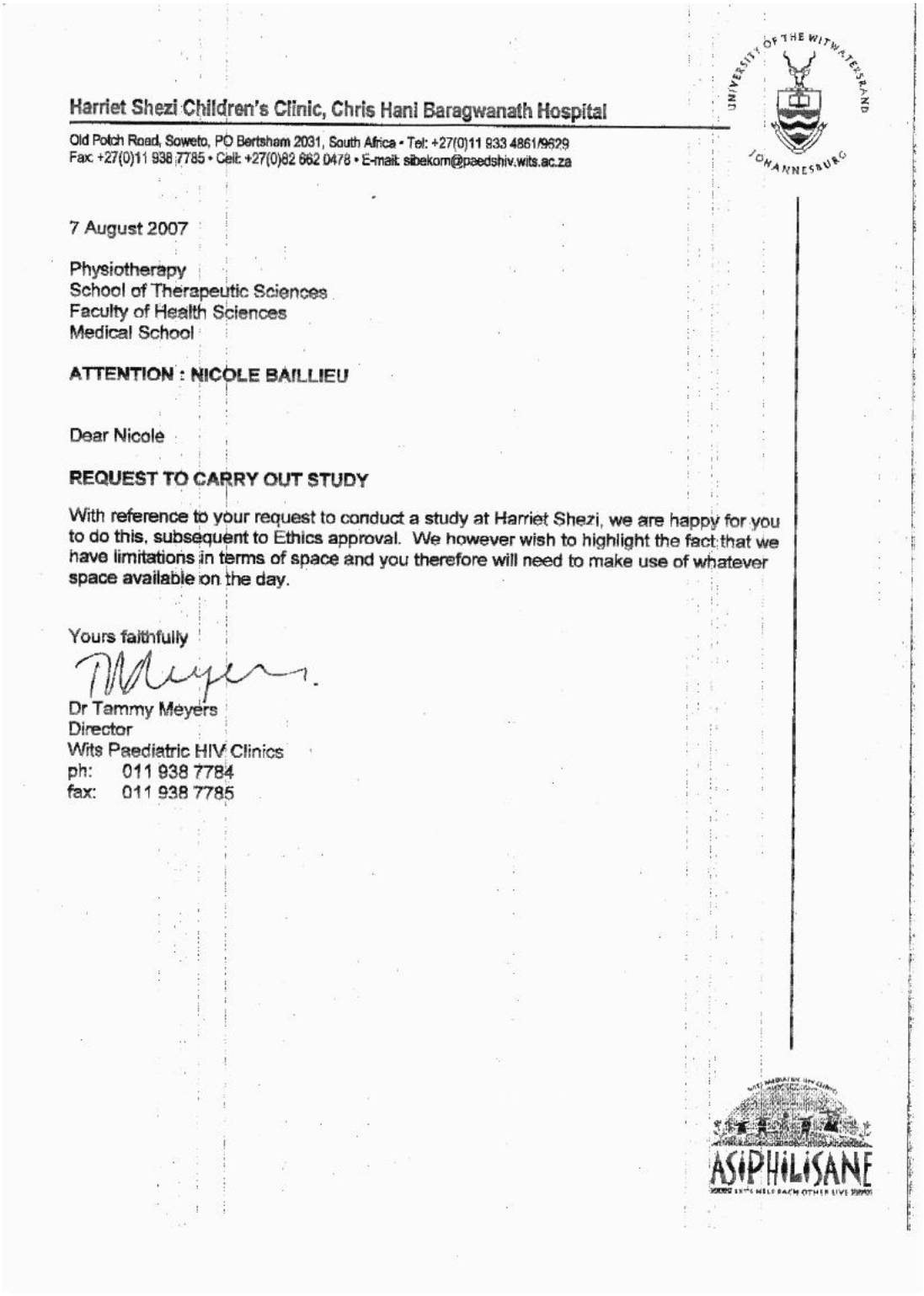
DECLARATION OF INVESTIGATOR(S)

To be completed in duplicate and **ONE COPY** returned to the Secretary at Room 10005, 10th Floor, Senate House, University.
I/We fully understand the conditions under which I am/we are authorized to carry out the abovementioned research and I/we guarantee to ensure compliance with these conditions. Should any departure to be contemplated from the research procedure as approved I/we undertake to resubmit the protocol to the Committee. **I agree to a completion of a yearly progress report.**

PLEASE QUOTE THE PROTOCOL NUMBER IN ALL ENQUIRIES

Appendix II

Consent from Harriet Shezi Clinic



Harriet Shezi Children's Clinic, Chris Hani Baragwanath Hospital

Old Potch Road, Soweto, PO Bertsham 2031, South Africa • Tel: +27(0)11 933 4861/9629
Fax: +27(0)11 938 7785 • Cell: +27(0)82 662 0478 • E-mail: sibekom@paedshiv.wits.ac.za

7 August 2007

Physiotherapy
School of Therapeutic Sciences
Faculty of Health Sciences
Medical School

ATTENTION : NICOLE BAILLIEU

Dear Nicole

REQUEST TO CARRY OUT STUDY

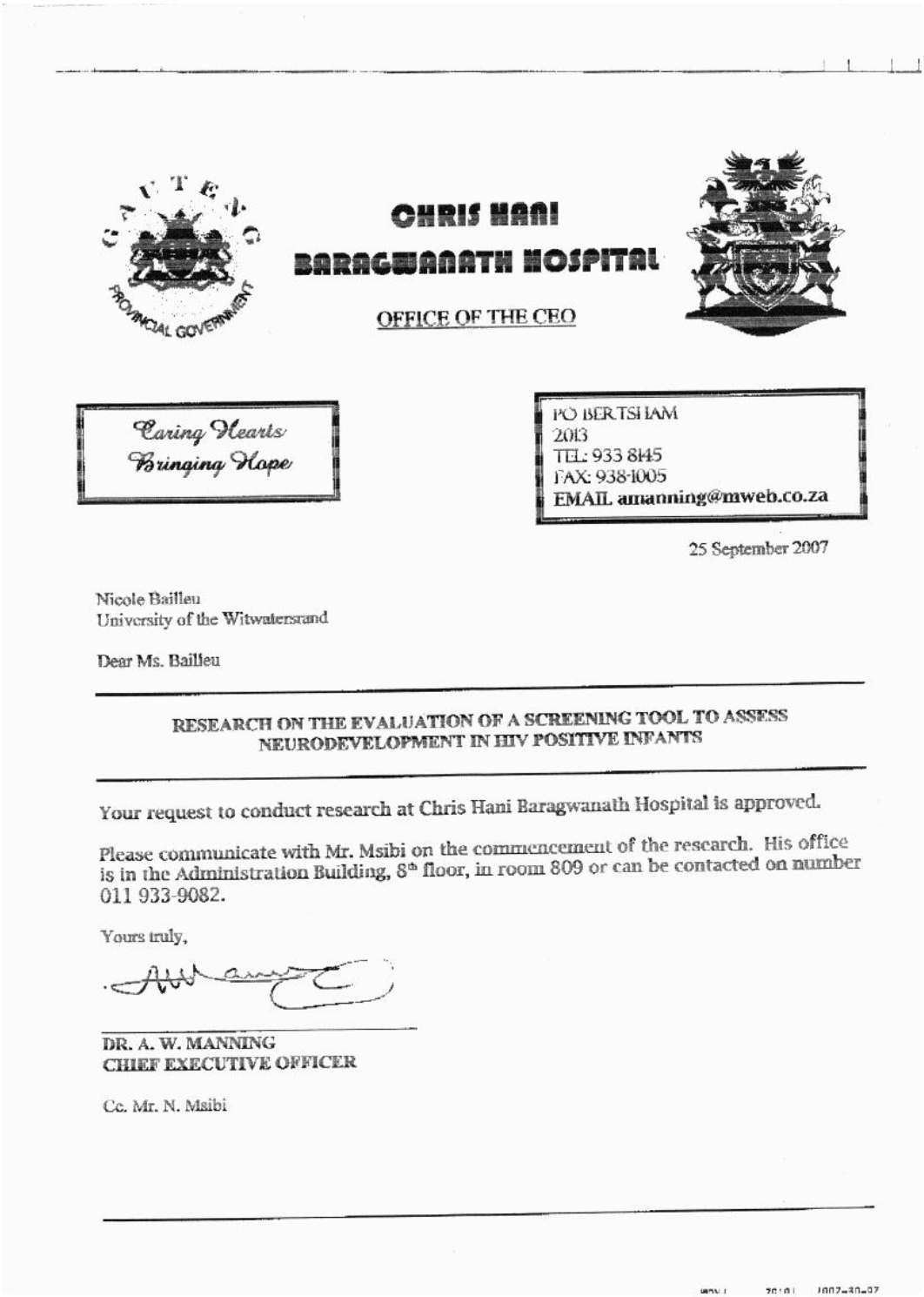
With reference to your request to conduct a study at Harriet Shezi, we are happy for you to do this, subsequent to Ethics approval. We however wish to highlight the fact that we have limitations in terms of space and you therefore will need to make use of whatever space available on the day.

Yours faithfully

Dr Tammy Meyers
Director
Wits Paediatric HIV Clinics
ph: 011 938 7784
fax: 011 938 7785

Appendix III

Consent from CEO of Baragwanath Hospital



GAUTENG PROVINCIAL GOVERNMENT

CHRIS HANI BARAGWANATH HOSPITAL

OFFICE OF THE CEO

Caring Hearts Bringing Hope

PO BERTSHAM
2013
TEL: 933 8145
FAX: 938-1005
EMAIL amanning@mweb.co.za

25 September 2007

Nicole Bailleu
University of the Witwatersrand

Dear Ms. Bailleu

RESEARCH ON THE EVALUATION OF A SCREENING TOOL TO ASSESS NEURODEVELOPMENT IN HIV POSITIVE INFANTS

Your request to conduct research at Chris Hani Baragwanath Hospital is approved.

Please communicate with Mr. Msibi on the commencement of the research. His office is in the Administration Building, 8th floor, in room 809 or can be contacted on number 011 933-9082.

Yours truly,

DR. A. W. MANNING
CHIEF EXECUTIVE OFFICER

Cc. Mr. N. Msibi

Appendix IV

Information sheet and informed consent

Good Morning

My Name is Nicole Hilburn. I am a Physiotherapist from the University of the Witwatersrand. I am doing a study to look at how we can test your child's development at your visits to the clinic so that the doctors can include this in their services to you.

HIV can affect a child's development, and make your child slower to develop. If we can test your child's development, and he/she needs some help to catch up, then we will be able to send you to the correct place for help. We would like to try this test of development on children so that we can see if it is able to tell us whether the child is developing on time. The test uses toys and is fun for the child. We want to try and make the test as short as possible so that it does not take up too much of your time.

I would like to test your child's development this morning, using a test that takes about an hour. You will not lose your place in the queue, as you will be called when it is your turn. The test is not painful, and involves toys and games, so your child will enjoy it. I am also seeking your permission to look at your child's medical file, so that we can find out about his/her height, weight and CD4 count to see what may cause slow development. Your information will be kept confidential at all times and your child's name will not be used. You may withdraw from the study at any time without a reason; this will not impact on your treatment the clinic in any way.

Nicole Hilburn
Physiotherapist Telephone: 011 7173702

Informed Consent

If you would like to take part in this study, please sign below

I_____(your name) and my child_____(child's name) agree to take part in this study. I give permission for Nicole Hilburn to obtain information from my child's file. I understand that I may withdraw from the study at any time.

Signed: _____ Date: _____

Appendix V

Example of data capture for item analysis

Key

8 month olds	7 month olds	6 month olds

maingross1	maingross1	maingross1	maingross1	maingross1	maingross1	maingross2	maingross2	maingross2	
2	2	2	2	2	2	2	2	2	
0	0	0	2	2	2	2	2	2	
2	2	2	2	2	2	2	2	2	
0	0	0	0	0	2	2	2	2	
2	2	2	2	2	2	2	2	2	
0	0	0	2	2	2	2	2	2	
2	2	2	2	2	2	2	2	2	
1	1	1	1	1	1	0	0	1	
2	2	2	2	2	2	2	2	2	
2	2	2	2	2	2	2	2	2	
1	1	1	1	0	0	0	0	0	
0	0	0	0	0	2	2	2	2	
0	0	0	0	0	2	2	2	2	
1	1	1	0	0	0	0	0	2	
1	0	0	0	0	0	2	2	2	
1	1	1	1	1	1	0	0	1	
5	4	4	3	2	2	0	0	2	
31%	25%	25%	19%	13%	13%	0%	0%	13%	
1	1	1	1	1	1	0	0	1	At-Risk
1	1	1	1	1	1	1	1	1	
1	1	1	1	1	1	0	1	1	
1	1	1	1	1	1	1	1	1	
1	1	1	1	1	1	1	1	0	
1	1	1	1	1	1	0	0	1	
1	1	1	1	1	1	1	0	1	
1	1	1	1	0	1	0	0	1	
8	8	8	8	7	8	4	4	7	
100%	100%	100%	100%	88%	100%	50%	50%	88%	
1	1	1	1	1	1	1	1	1	Emerging
1	1	1	1	1	1	0	0	1	
1	1	1	1	1	1	0	1	1	
1	1	1	1	1	1	1	0	1	
1	1	1	1	1	1	1	0	1	
1	1	1	1	1	1	1	1	1	
1	1	1	1	1	1	1	1	1	
1	1	1	1	1	1	1	1	1	
1	1	1	1	1	1	1	1	1	
1	1	1	1	1	1	0	0	1	
1	1	1	1	1	1	1	1	1	
1	1	1	1	0	0	1	1	1	
1	1	1	1	1	1	1	1	1	
1	1	1	1	1	1	1	1	1	
14	14	14	14	13	13	11	10	14	
100%	100%	100%	100%	93%	93%	79%	71%	100%	Competent

Appendix VI

Information sheet and informed consent for study five

Good Morning

My Name is Nicole Hilburn. I am a Physiotherapist from the University of the Witwatersrand. I am doing a study to look at how we can test your child's development at your visits to the clinic so that the doctors can include this in their services to you.

HIV can affect a child's development, and make your child slower to develop. If we can test your child's development, and he/she needs some help to catch up, then we will be able to send you to the correct place for help. We would like to try this test of development on children so that we can see if it is able to tell us whether the child is developing on time. We want to try and make the test as short as possible so that it does not take up too much of your time.

I would like to test your child's development this morning, using two tests that will take about half an hour. You will not lose your place in the queue, as you will be called when it is your turn. The test is not painful, and involves toys and games, so your child will enjoy it. If your child is found to be delayed, information will be given to you that will help you to make your child stronger.

.

Your information will be kept confidential at all times and your child's name will not be used. You may withdraw from the study at any time without a reason; this will not impact on your treatment the clinic in any way.

Nicole Hilburn
Physiotherapist Telephone: 011 7173702

Informed Consent

If you would like to take part in this study, please sign below

I_____(your name) and my child_____(child's name) agree to take part in this study. I understand that I may withdraw from the study at any time.

Signed: _____ Date: _____

Appendix VII

Consent from CEO of Rahima Moosa Hospital

University of the Witwatersrand

Faculty of Health Sciences

Department of Physiotherapy

011 7173702

18th March 2010

Dear Mrs Jordaan

Request for permission to carry out research at Rahima Moosa Mother and Child Hospital

I am a paediatric physiotherapist, and lecturer in the Department of Physiotherapy at the University of the Witwatersrand. I am currently carrying out a PhD study entitled The Development of a Screening Tool to Evaluate Gross Motor Development in Infants who are HIV positive.

The screening tool is complete, and now needs to undergo reliability testing, which I would like to request to undertake in the wards at your hospital. The screening tool evaluates gross motor function of infants between the ages of 6 and 18 months, is non-invasive, and takes 5 – 10 minutes to administer, so it is in no way painful or upsetting to the child.

I would need to assess 15 infants twice, and would request permission from their parent/caregiver before conducting the assessment.

Ethical clearance for the study has been obtained from the university Ethics committee.

Please could you advise as to whether this would be possible

Kind regards

Nicole Hilburn

Paediatric Physiotherapist

0827718443

nicole.hilburn@wits.ac.za

nicolehilburn@gmail.com

Approved. Dr. H. J. Rossi

Appendix VIII

Information sheet and informed consent for test-retest reliability

Good Morning

My Name is Nicole Hilburn. I am a Physiotherapist from the University of the Witwatersrand. I would like to use a test of development to see if it is able to tell us whether your child is developing on time, and whether the score is the same over two tests.

The test that takes about ten minutes, and is not painful for your child. If you agree to take part in the study, your child will be assessed twice, with the second test being done two days after the first. Your information will be kept confidential at all times and your child's name will not be used. If your child is found to be delayed, information will be given to you which will help your child to get stronger. You may withdraw from the study at any time without a reason; this will not impact on your treatment in the hospital in any way.

Nicole Hilburn
Physiotherapist
011 7173702

Informed Consent

If you would like to take part in this study, please sign below

I_____(your name) and my child_____(child's name) agree to take part in this study. I understand that I may withdraw from the study at any time.

Signed: _____ Date: _____

Appendix IX

Information sheet and informed consent for inter-rater reliability

Good Morning

My Name is Nicole Hilburn. I am a Physiotherapist from the University of the Witwatersrand. I would like to use a test of development to see whether two people are able to get the same score for your child.

The test that takes about ten minutes, and is not painful for your child. If you agree to take part in the study, your child will be assessed by myself, or Joanne Potterton, and the other therapist will watch the assessment and will score it at the same time. Your information will be kept confidential at all times and your child's name will not be used. If your child is found to be delayed, information will be given to you which will help your child to get stronger. You may withdraw from the study at any time without a reason; this will not impact on your treatment in the hospital in any way.

Nicole Hilburn
Physiotherapist
011 7173702

Informed Consent

If you would like to take part in this study, please sign below

I_____(your name) and my child_____(child's name) agree to take part in this study. I understand that I may withdraw from the study at any time.

Signed: _____ Date: _____