# A GLMM analysis of data from the Sinovuyo Caring Families Program (SCFP)

MSc Statistics Thesis (STA5004W)

Raymond T Nhapi (NHPRAY001)
nhpray001@myuct.ac.za

Supervised by: A/Prof F. Little & Dr R. Kassanjee

Data owned and collected by: A/Prof C. Ward & I. Wessels



17th September 2018

# Contents

# Chapter 1

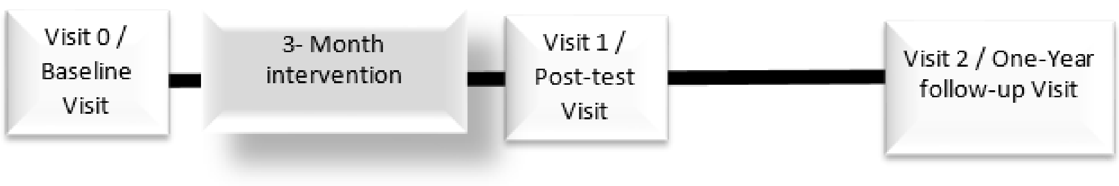
# Introduction

The aim of this project is to analyse the effects of the Sinovuyo Caring Families Program (SCFP) - a group-based parent skills training intervention for primary caregivers of children aged 2 to 9 years. A randomized control trial (RCT) was implemented whereby 296 eligible child-caregiver dyads were randomized into one of two study arms (the Intervention Arm and the Control Arm). The eligibility criteria included the following:

1. the caregiver had to be 18 years or older;
2. the child was aged between 2 and 9 years;
3. the caregiver was living in the same house as the child for at least 4 nights per week;
4. If the primary caregiver had provided informed consent to participate in the full study including the intervention (for the intervention group) and at all data collection points; and
5. Reporting 15 or more problem behaviours on the Eyeberg Child Behaviour Inventory (ECBI) problem scale, this condition ensured that the participants actually needed the intervention.

Block randomization was applied to ensure a balanced study with respect to both age category (i.e. 2-5 and 6-9 years old) and sex (boys and girls) of the children. Each participant in the intervention arm was further allocated to one of eleven groups of size 11-17 individuals per group. The allocation was based on location (proximity to other participants) and availability of the participants for programme sessions. Each of these groups had a facilitator pair (one of two pairs) assigned to it throughout the programme.

Figure 1.1: Outline of SCFP Study Visits




This RCT was conducted in two independent waves, the first being in Khayelitsha between May and August 2014 and the second one was in Nyanga spanning from September to December 2014. The data was collected at three time points: (1) at the beginning of the study (visit

0 / baseline visit), (2) immediately after the completion of the 3-month intervention (visit 1 /post-test visit) and (3) 12 months after visit 1 (visit 2 / one year follow-up visit). Figure 1.1 shows an outline of the 3 study visits where the data was collected. This data collection process was mainly done through interviewer-administered questionnaires (self-reported measures) as well as 'observational' measures where child-caregiver interactions were videoed and scored. The intervention was primarily made up of 12 group sessions. If the participants missed the group sessions, they were offered home visits (depending on reasons).

The effects of the programme are measured by several endpoints (multiple responses) which could be classified as either primary or secondary endpoints. The following lists the sets of endpoints that were considered and further lists variables used to measure these endpoints:

**Primary endpoints**:

(a) Reported child behaviour problems

1. Eyeberg Child Behaviour Inventory (ECBI) Intensity
2. Eyeberg Child Behaviour Inventory (ECBI) Problem

(b) Reported positive parenting

3. Supporting Positive Behaviour Frequency (subscore)
4. Setting Limits Frequency (subscore)
5. Positive Parenting Frequency (sum of the above two)
6. Supporting Positive Behaviour Problem (subscore)
7. Setting Limits Problem (subscore)
8. Positive Parenting Problem (sum of the above two)

(c) Reported harsh parenting

9. Non-Violent Discipline
10. Physical Discipline
11. Severe Physical Discipline
12. Psychological Discipline
13. Neglect

(d) Observed parent and child behaviour (done via video coding by researchers)

14. Parent Positive Behaviour
15. Parent Negative Behaviour
16. Child Positive Behaviour
17. Child Negative Behaviour

**Secondary endpoints**:

(e) Monitoring and supervision

18. Poor Monitoring and Supervision

(f) Depression

19. Beck Depression Inventory (BDI) score

(g) Parenting stress

20. Parental Distress (subscore)

21. Parent Child Dysfunctional Interaction (subscore)
22. Difficult Child (subscore)
23. Parenting Stress (sum of the above three)

(h) Social support

24. Social support score

The main aim of the study is to evaluate whether the SCFP is effective in improving the level of childcare and child behaviour within families. Additionally, the study also aims to evaluate whether the effects of the programme differ between the different wave - intimate partner violence (ipv) strata/combinations. Reasons for including wave include the fact that the two communities, Khayelitsha (wave 1) and Nyanga (wave 2) have different socio-economic profiles and also to account for the possibly different implementations of the program (as they happened one after the other, a lot could have differed). Amongst other reasons, caregiver's experience of IPV is also considered because of the suspicion that it would affect the way those affected would respond to the program (the ipv-exposed might have greater responses to the intervention). This is evaluated through fitting generalized linear mixed-effects models (GLMMs) which accommodate the repeated-measures component of the study design as well as the different distributions of the study outcomes. Furthermore, since there were some missing data, multiple imputation using chained equations was implemented and the GLMMs were fitted on five imputed datasets and results were pooled using Rubin's rules. The same models were also fitted on the complete cases from the unimputed datasets and the results were compared to the imputation model results.

A detailed description of the data including the relevant summaries of the variables, reliability analyses and an a brief description of the missingness pattens is documented in Chapter 2. Details of the imputation models used are explained in Chapter 3. This chapter also documents how the data was imputed by assuming missingness at random (MAR) and using the chained equations approach to do the imputations. Chapter 4 firstly describes the generalized linear mixed models (GLMMs) in detail and goes on to explain how they were subsequently fitted and also describes the model results. Lastly, the general conclusions and limitations of the analysis are documented in Chapter 5.

The analysis plan as set out in Appendix B was developed independently of the analysis presented in this dissertation. The brief for the research underlining this dissertation was to implement the analysis as set out in the plan.

# Chapter 2

# Data

The data comes in two main forms: parental self-reports and observational video scores. The parental self-reports are those responses obtained via questionnaires and they constitute the majority of the data at hand. As for the observational video scores, trained research assistants would score the participants behaviours based on the videos recorded at the assessment visits. Data from the questionnaires were mostly indicators of **presence or absence** (binary) of behavioural characteristics and assessments of the **degree** of behaviour (mainly Likert scales). Observational video scores mainly consisted of (numeric) count data resulting from observers counting the number of occurrences of certain kinds of child-caregiver interactions of interest that were recorded. The data is longitudinal and was collected over 3 time points: baseline (visit 0; before the intervention), post-test (visit 1; 3 months after baseline) and one year follow-up (visit 2; 12 months after post-test visit). There are hundreds of these longitudinal data items captured at each of the three time points and each of them was used to construct the primary and secondary outcomes (often referred to as the outcomes of interest in this analysis) that would then be used to evaluate the program. A small subset of the data items obtained at baseline capture the demographics of the child and caregivers such as age, gender, HIV status and caregiver's history of maltreatment.

The following sections present an exploratory analysis of the data by mainly focusing on the outcomes of interest that were included in the final analysis. Section 2.1 shows how the outcomes were constructed from the data collected through the questionnaires and video-coding. Since all these scores/ outcomes of interest are created by summing a number of items from the questionnaires and video-coding output, a reliability analysis of the scores is presented in Section 2.2. Some summary statistics of the baseline demographics and the outcomes of interest (at the baseline visit) are presented in Section 2.3, this is followed by a brief look into the distribution of the outcomes across the study visits in Section 2.4. Lastly exploratory analyses into the missingness of data as well as the missing data patterns are presented in sections 2.5 and 2.6 respectively.

## 2.1 Score construction

Each outcome of interest is a composite score constructed as a sum of several related items/ responses from either the questionnaires or the video scores. The score construction rules also maintained that data items would only be added together if they had the same data class (e.g binary responses aren't summed with Likert scales and so on). By way of example, the Eyeberg Child Behaviour Inventory (ECBI) Intensity score was constructed by adding up 36 items/responses from the ECBI questionnaire as completed by the primary caregiver. These items were all scored on a 7-point Likert scale with 1,2, ... ,7 representing 'Never', 'Very rarely', 'Rarely', 'Sometimes', 'Often', 'Very often' and 'Always' respectively. Higher scores on the ECBI Intensity indicate less desirable child behaviour and the score ranges from 36 to 252. Similarly, the ECBI Problem score was created by adding up 36 binary items from the same questionnaire. Higher scores imply less desirable behaviour and the score ranges from 0 to 36. Observed Parenting and Child behaviour outcomes like the Positive Child Score were obtained by adding up occurrences of certain behavioural items as observed (by trained research assistants) on the videos that were recorded. These scores would range from 0 (indicating no such item observed) and would not have a specified upper limit.

By default, observational video scores were also not added up with any self reported items during score construction. Appendix A of the Analysis Plan document gives a more detailed account of how these outcomes were constructed. In short, each outcome would be a sum of related data items. All the scores that were created by summing up Likert-type data were treated as continuous throughout the analysis and those created from summing up binary outcomes were modelled as count data except for a few whose support/ range was wide enough to allow for modelling as continuous data. Adding up these individual Likert scales (or even the binary items) essentially means that the differences between the different frequency assessments were considered to be constant i.e. "Very rarely" - "Never" $\equiv$ "Rarely" - "Very rarely" $\equiv \cdots \equiv$ "Always" - "Very often" for the 7-point Likert scale described above.

The outcomes were divided into two broad categories: primary and secondary endpoints. The primary outcomes/ endpoints were further subdivided into four groups, both focussing on parent and child behaviour. These included (a) **reported child behaviour problems**, (b) **reported positive parenting**, (c) **reported harsh parenting**, and (d) **observed parent & child behaviour**. The secondary endpoints included subcategories like (e) **monitoring & supervision**, (f) **depression**, (g) **parenting stress** and (h) **social support**.

## 2.2 Reliability analysis

In general, reliability analyses is done to check whether some scale of interest produces consistent (and therefore reliable) results. Any assessment tool/scale is required to be consistent especially with regards three main aspects namely (a) inter rater consistency, (b) test-retest reliability and (c) internal consistency. Inter-rater consistency is achieved if for any participant, the tool produces similar results regardless of who conducted the assessment. To test for this, one would require data to be collected on the same set of participants by a number of different interviewers/ assessors. Test-retest reliability is achieved if similar results are achieved when the assessment/ test is done on the same participant on two reasonably similar time-points. Simple tests of association like t-tests or analysis of variance or their non-parametric equivalents can be used to assess any of the above mentioned aspects of reliability. These could not be evaluated on the SCFP RCT data as it did not meet the requirements i.e. by design, the three data collection points of the RCT would not meet the test-retest criteria and at each time point each dyad would only have one assessment.

Internal consistency evaluates the reliability of the tool by focussing on the individual items that are summed together to create the composite score. The assessment aims to determine whether the different items being added up to form a unique scale are hugely positive;y correlated with one another i.e. moving in the same direction.For example, the ECBI Intensity Score discussed in Section 2.1 above would be considered to be internally consistent if all the 36 individual Likert scale items that are used are consistent with each other, i.e. if they all give higher scores for children with more behaviour problems. The most commonly used measure for internal consistency is the Cronbach's alpha (Cronbach, 1951) which is generally defined as:

$$\alpha = \frac{n}{n-1}\left(1 - \frac{\sum_{i=1}^{n} V_i}{V_t}\right) \tag{2.1}$$

where $n$ is the number of items being summed to obtain the test (or composite) score, $V_i$ is the sample variance of the $i^{th}$ item; in case of a binary (0/1) item $V_i = p_i(1 - p_i)$ where $p_i$ is the proportion (in the sample) receiving an outcome of zero on the item. $V_t$ is the (sample) variance of the test/ composite score itself. If the sample is such that all items have one unique value (for example if everyone ticked "yes" for all items in a questionnaire with "yes/ no" options) then $\alpha$ will be have a value of 1. The formula in also reduces to 1 when the items (on the same scale) are perfectly correlated, i.e. when they're moving in the perfectly same direction. Less consistent items will result in smaller alpha values and one can even achieve an alpha value below zero if there are some negatively related items. An alpha value of zero is achieved if there is a zero correlation, a special case of this which occurred frequently in the data is when only one item has some non-zero variance and the rest are constant.

Though the data was based on validated questionnaires, an evaluation of internal consistency was still necessary to show which composite scores/ outcomes of interest were reliable as there is a chance that some of the items in the questionnaires were (a) misunderstood or (b) irrelevant to the (African) population at hand. For each outcome at each of the time points, an alpha value was computed and then recalculated with each of the items omitted from the set that was used to construct it. This was done for all the items within the aforementioned sets. Tables 2.1, 2.2 and 2.3 presents four main results from this exercise: (1) the alpha value based on the whole set of summands, (2 & 3) the minimum and maximum alphas obtained after omitting one of the summands iteratively, (4) the item/ summand that gives the maximum alpha if omitted. The aim is to flag the items that aren't consistent with the rest of the summands.

From the three tables, it seems that **Severe Physical Discipline** and **Neglect** have poor reliability across the three time points. These two outcomes were ultimately removed from the analysis because of the low reliability and also because of low variability. Outcomes like **Supporting Positive Behaviour Frequency** (SPB Frequency), **Poor Monitoring and Supervision** (Poor M & S) and all outcomes making up the **Harsh Parenting** group seem to have low reliability (alpha below 0.7, as a rule of thumb) across all time points. In addition, scores making up the **Caregiver experience of intimate partner violence** at baseline also seem to have low reliability. Besides the outcomes flagged above, there are no other notable ones with poor reliability or which would improve if one of their summands is omitted. Apart from not including **severe physical discipline** and **neglect**, no adjustments were made to the scores in spite of these low reliability assessments.

Table 2.1: **Cronbach's Alpha: Baseline Scores**

| Score | Actual Alpha | Min Alpha | Max Alpha | Item Name |
|---|---|---|---|---|
| **Child Behaviour Problems** | | | | |
| ECBI intensity | 0.7984 | 0.7877 | 0.8034 | Wets bed prequency |
| ECBI problem | 0.7340 | 0.7222 | 0.7351 | Wets bed problem |
| **Positive Parenting** | | | | |
| Supporting Positive BehaviourFrequency | 0.5884 | 0.5056 | 0.6302 | Family meal together |
| Setting Limits Frequency | 0.7616 | 0.6967 | 0.8128 | Speak calmly with child |
| Positive Parenting Frequency | 0.7575 | 0.7211 | 0.7668 | Family meal together |
| Supporting Positive BehaviourProblem | 0.6923 | 0.6044 | 0.7017 | Problem solving behaviour |
| Setting Limits Problem | 0.7966 | 0.7473 | 0.7945 | Stick to rules behaviour |
| Positive Parenting Problem | 0.8412 | 0.8185 | 0.8446 | Family meal together behaviour |
| **Harsh Parenting** | | | | |
| Non-Violent Discipline | 0.2897 | 0.0608 | 0.4213 | ICAST5 |
| Physical Discipline | 0.5808 | 0.4894 | 0.5998 | ICAST8 |
| Severe Physical Discipline | 0.1422 | 0.0061 | 0.2347 | ICAST24 |
| Psychological Discipline | 0.6321 | 0.5680 | 0.6287 | ICAST21 |
| Neglect | -0.0091 | -0.0122 | 0.0000 | ICAST26 |
| **Monitoring and Supervision** | | | | |
| Poor Monitoring and Supervision | 0.4632 | 0.3033 | 0.5672 | Lets you know where going |
| **Depression** | | | | |
| Beck Depression Inventory | 0.8992 | 0.8909 | 0.9019 | Loss of interest in sex |
| **Parenting Stress** | | | | |
| Parental Distress | 0.8194 | 0.7919 | 0.8212 | Meet child needs |
| Parent Child Dysfunctional Interaction | 0.8574 | 0.8353 | 0.8664 | Self assessment of parent |
| Difficult Child | 0.7773 | 0.7480 | 0.8013 | Getting child to stop |
| Parenting Stress | 0.9093 | 0.9040 | 0.9128 | Getting child to stop |
| **Social Support** | | | | |
| Social Support | 0.8515 | 0.8255 | 0.8439 | Someone to listen |
| **Caregiver's experience of intimate partner violence (IPV)** | | | | |
| IPV Chronicity | 0.8466 | 0.8043 | 0.8495 | My partner insulted or shouted or yelled |
| **Caregiver's experience of maltreatment** | | | | |
| Physical Abuse Experience | 0.3566 | 0.0004 | 0.4089 | Hit, punch or kick you very hard |
| Emotional Abuse Experience | 0.5752 | 0.4294 | 0.6115 | Insult and criticise you to make you feel that you were bad |
| Sexual Abuse Experience | 0.6495 | 0.4393 | 0.6637 | Made you touch their private parts genitals |

Table 2.2: **Cronbach's Alpha: Post-test Scores**

| Score | Actual Alpha | Min Alpha | Max Alpha | Item Name |
|---|---|---|---|---|
| **Child Behaviour Problems** | | | | |
| ECBI intensity | 0.8872 | 0.8810 | 0.8902 | Whines frequency |
| ECBI problem | 0.9004 | 0.8954 | 0.9007 | Wets bed problem |
| **Positive Parenting** | | | | |
| Supporting Positive BehaviourFrequency | 0.6069 | 0.5198 | 0.6532 | Family meal together |
| Setting Limits Frequency | 0.7899 | 0.7397 | 0.8100 | Speak calmly with child |
| Positive Parenting Frequency | 0.7843 | 0.7582 | 0.7969 | Family meal together |
| Supporting Positive BehaviourProblem | 0.7400 | 0.6682 | 0.7504 | Family meal together behaviour |
| Setting Limits Problem | 0.8439 | 0.8114 | 0.8375 | Stick to rules behaviour |
| Positive Parenting Problem | 0.8778 | 0.8641 | 0.8818 | Family meal together behaviour |
| **Harsh Parenting** | | | | |
| Non-Violent Discipline | 0.3335 | 0.0397 | 0.3892 | ICAST19 |
| Physical Discipline | 0.6255 | 0.5291 | 0.6433 | ICAST8 |
| Severe Physical Discipline (*) | 0.0000 | 0.0000 | 1.0000 | ICAST24 |
| Psychological Discipline | 0.6537 | 0.5881 | 0.6586 | ICAST11 |
| Neglect | -0.0085 | -0.0113 | 0.0000 | ICAST26 |
| **Monitoring and Supervision** | | | | |
| Poor Monitoring and Supervision | 0.4137 | 0.2550 | 0.5070 | Check home time |
| **Depression** | | | | |
| Beck Depression Inventory | 0.9056 | 0.8975 | 0.9073 | Loss of interest in sex |
| **Parenting Stress** | | | | |
| Parental Distress | 0.8289 | 0.8035 | 0.8343 | Meet child needs |
| Parent Child Dysfunctional Interaction | 0.8664 | 0.8468 | 0.8689 | Self assessment of parent |
| Difficult Child | 0.8543 | 0.8342 | 0.8700 | Getting child to stop |
| Parenting Stress | 0.9294 | 0.9260 | 0.9320 | Getting child to stop |
| **Social Support** | | | | |
| Social Support | 0.8819 | 0.8608 | 0.8782 | Someone understands problems |
| **Caregiver's experience of IPV** | | | | |
| IPV Chronicity | 0.7182 | 0.6332 | 0.7620 | My partner insulted or shouted or yelled |

(*) Outcome was made up of 3 items (ICAST16 ICAST23 and ICAST24), the first two were just constant in the dataset with all values being 0.

Table 2.3: **Cronbach's Alpha: Follow-up Scores**

| Score | Actual Alpha | Min Alpha | Max Alpha | Item Name |
|---|---|---|---|---|
| **Child Behaviour Problems** | | | | |
| ECBI intensity | 0.8975 | 0.8921 | 0.8987 | Whines frequency |
| ECBI problem | 0.918 | 0.9142 | 0.9176 | Hits parents problem |
| **Positive Parenting** | | | | |
| Supporting Positive BehaviourFrequency | 0.6315 | 0.5295 | 0.686 | Family meal together |
| Setting Limits Frequency | 0.8475 | 0.8083 | 0.8711 | Speak calmly with child |
| Positive Parenting Frequency | 0.7904 | 0.7611 | 0.7959 | Family meal together |
| Supporting Positive BehaviourProblem | 0.7402 | 0.6333 | 0.7484 | Praise child behaviour |
| Setting Limits Problem | 0.7442 | 0.6719 | 0.7487 | Expected behaviour from child behaviour |
| Positive Parenting Problem | 0.8361 | 0.8096 | 0.8366 | Family meal together behaviour |
| **Harsh Parenting** | | | | |
| Non-Violent Discipline | 0.4935 | 0.2217 | 0.4927 | ICAST19 |
| Physical Discipline | 0.5073 | 0.3871 | 0.5024 | ICAST15 |
| Severe Physical Discipline | -0.004 | -0.0053 | 0.0000 | ICAST16 |
| Psychological Discipline | 0.5498 | 0.4531 | 0.5555 | ICAST21 |
| Neglect | -0.007 | -0.0094 | 0.0000 | ICAST27 |
| **Monitoring and Supervision** | | | | |
| Poor Monitoring and Supervision | 0.3352 | 0.1311 | 0.4881 | Lets you know where going |
| **Depression** | | | | |
| Beck Depression Inventory | 0.9344 | 0.9286 | 0.9402 | Loss of interest in sex |
| **Parenting Stress** | | | | |
| Parental Distress | 0.8263 | 0.7974 | 0.8398 | Meet child needs |
| Parent Child Dysfunctional Interaction | 0.7251 | 0.6883 | 0.7467 | Self assessment of parent |
| Difficult Child | 0.8183 | 0.7846 | 0.8443 | Number of things child bothers |
| Parenting Stress | 0.8844 | 0.8771 | 0.8894 | Getting child to stop |
| **Social Support** | | | | |
| Social Support | 0.914 | 0.8982 | 0.9154 | Someone to listen |
| **Caregiver's experience of IPV** | | | | |
| IPV Chronicity | 0.7776 | 0.7103 | 0.8042 | Sex without a condom |

## 2.3 Baseline Comparisons

This section explores the data at baseline with the aim to: (1) show whether the randomization of participant dyads into the two arms was a success and to describe the cohort and (2) show whether there was any biased drop out from the study (loss to follow-up). If the distributions of the data between the two arms is roughly the same, it can be concluded that randomization into the arms worked. Similarly for the **loss to follow-up** construct, it will be deemed unbiased if the distributions between the two groups are roughly the same.

### 2.3.1 Summaries by Arm

Table 2.4 summarizes the baseline characteristics of the caregivers and children. A count and percentage (of the number in the arm) is presented for all categorical variables. The number of non-missing responses, minimum, maximum, lower & upper quartiles, median and mean are recorded for the numeric variables. If the overall number of zeros recorded for the numeric variables is above 50%, the summary statistics are then split in such a way that the zeros are shown separately from the non-zero responses. The above strategy helps illustrate the extent of **zero-inflation** in some of the characteristics. A similar reporting style is used in tables 2.5, 2.6 and 2.7.

The most notable characteristics illustrated in Table 2.4 are as follows. Only a few caregivers have an unknown HIV status but a bigger proportion of children have unknown HIV statuses across the arms. The main reasons for the unknown status could be that the participant wasn't tested or that they chose not to disclose. Of those that know (and disclosed) their HIV status, it seems that there are more positive cases among the caregivers than the children and that those in the intervention arm seem to have a marginally higher percentage of positive cases as well (about 11% higher). The majority of caregivers in the study were female in addition, the child age and gender were balanced by design. Caregiver age also seems to be fairly balances between the arms. It also seems that there are similar proportions across the arms when comparing the presence of the biological mother or father in the household though it is worth noting that the population shows very low proportions of households where the mother (around 12%) and in about half of the households have the biological father present. The data also shows high unemployment among the caregivers (above 80% in both arms) as well as high levels of reported alcohol use (around a third in both arms). A substantial proportion (25% in the control arm and 34% in the intervention arm) of the caregivers reported some experience of intimate partner violence (IPV). There is a very high proportion of caregiver's experience of maltreatment in the whole sample mainly contributed to by physical abuse experience (above 50% prevalence). Overall, it doesn't seem that there are any substantial differences between the two arms with respect to the aforementioned baseline characteristics.

Table 2.4: **Baseline characteristics of child/caregiver/household by Arm**

| | **Control Arm** ( n = 148) | **Intervention Arm** ( n = 148) |
|---|---|---|
| | **n** (min, Q1, med, **mean**, Q3, max) | **n** (min, Q1, med, **mean**, Q3, max) |
| **Age of caregiver** | 148 (18.00, 27.75, 31.00, **34.43**, 39.25, 75.00) | 148 (20.00, 27.00, 32.50, **33.61**, 38.25, 62.00) |
| **Age of child** | 148 (2.00, 3.00, 5.00, **5.20**, 7.00, 9.00) | 148 (2.00, 3.00, 5.00, **5.26**, 7.00, 9.00) |
| | | |
| | n (**% of Control Arm**) | n (**% of Intervention Arm**) |
| **HIV status of caregiver** | | |
| Positive | 30 (**20.27**) | 46 (**31.08**) |
| Negative | 110 (**74.32**) | 97 (**65.54**) |
| Unknown | 8 (**5.41**) | 5 (**3.38**) |
| **HIV status of child** | | |
| Positive | 1 (**0.68**) | 2 (**1.35**) |
| Negative | 77 (**52.03**) | 89 (**60.14**) |
| Unknown | 70 (**47.30**) | 57 (**38.51**) |
| **Family Structure** | | |
| Biological mother present in household | 19 (**12.84**) | 18 (**12.16**) |
| Biological father present in household | 65 (**43.92**) | 74 (**50.00**) |
| **Childs HIV orphanhood status** | | |
| Single orphan | 1 (**0.68**) | 2 (**1.35**) |
| Double Orphan | 0 (**0.00**) | 0 (**0.00**) |
| **Gender of caregiver** | | |
| Male | 1 (**0.68**) | 0 (**0.00**) |
| Female | 147 (**99.32**) | 148 (**100.00**) |
| **Gender of child** | | |
| Male | 79 (**53.38**) | 79 (**53.38**) |
| Female | 69 (**46.62**) | 69 (**46.62**) |
| **Caregiver's employment status** | | |
| Employed | 25 (**16.89**) | 18 (**12.16**) |
| Unemployed | 123 (**83.11**) | 130 (**87.84**) |
| **Caregiver substance use** | | |
| Alcohol | 44 (**29.73**) | 49 (**33.11**) |
| Drugs | 8 (**5.41**) | 11 (**7.43**) |
| **Caregiver experience of intimate partner violence** | | |
| IPV incidence | 37 (**25.17**) | 50 (**34.25**) |
| **Caregiver experience of maltreatment** | | |
| Physical abuse [1] | | |
| Overall | 129 (0.00, 0.00, 1.00, **0.74**, 1.00, 4.00) | 138 (0.00, 0.00, 1.00, **0.75**, 1.00, 3.00) |
| Zeros | 62 (**48.06**) | 65 (**47.10**) |
| Non-zeros | 67 (1.00, 1.00, 1.00, **1.43**, 2.00, 4.00) | 73 (1.00, 1.00, 1.00, **1.43**, 2.00, 3.00) |
| Emotional abuse [2] | | |
| Overall | 127 (0.00, 0.00, 0.00, **0.59**, 1.00, 4.00) | 135 (0.00, 0.00, 0.00, **0.54**, 1.00, 5.00) |
| Zeros | 80 (**62.99**) | 84 (**62.22**) |
| Non-zeros | 47 ( 1.0, 1.0, 1.0, **1.6**, 2.0, 4.0) | 51 (1.00, 1.00, 1.00, **1.43**, 2.00, 5.00) |
| Sexual abuse [2] | | |
| Overall | 140 (0.00, 0.00, 0.00, **0.09**, 0.00, 3.00) | 141 (0.00, 0.00, 0.00, **0.14**, 0.00, 4.00) |
| Zeros | 132 (**94.29**) | 130 (**92.20**) |
| Non-zeros | 9 (1.00, 1.00, 1.00, **1.33**, 1.00, 3.00) | 10 (1.00, 1.25, 2.00, **2.00**, 2.00, 4.00) |

[1] Sum of 4 dichotomous responses, higher score being less desirable.

[2] Sum of 5 dichotomous responses, higher score being less desirable.

Table 2.5 summarizes the baseline outcomes across the two arms. There is an indication of zero inflation in the **harsh parenting** outcomes: severe physical discipline and neglect (these were omitted from the final analysis for lack of reliability, see Section 2.2) as well as **child negative behaviour** which is an observed outcome. The means and medians are comparable across the arms and there is significant overlap in both the interquartile range and overall ranges. Possible exceptions are the **observed parent positive and negative behaviour** assessments where the maximum count for the intervention arm seem to be much higher than for the control arm and the maximum level for child positive behaviour is notably higher in the control arm. These values could be possible outliers however, in this analysis, they are considered as normal values expected from long tailed distributions. The **ECBI problem** score was used as a screening tool so it is to be expected that the baseline scores are left-censored at 15 in both arms.

Table 2.5 also shows the minimum and maximum possible scores attainable for each of the outcomes of interest in the study. This helps with hypothesising about **the extent of the problem** within the cohort at baseline by comparing the means and/ or medians with the aforementioned maximum and minimum values. For example, one can clearly see that the whole sample has high child behaviour problems as indicated by the average **ECBI problem** score (25.16 in the control arm and 24.61 in the intervention arm) being closer to the score's theoretical maximum value (36). Again, this is to be expected since it was used as a screening tool in order to be able to give the intervention to those who actually needed it.

Parents scored on average in the middle of the possible ranges for the **positive parenting** scores. The prevalence of reported **severe physical discipline** was low (9.46% among the controls and 16.22% in the intervention arm), an even lower prevalence also occurred for neglect (0.68% in the control arm and 4.73% in the intervention arm). The prevalence of observed **child negative behaviour** was high (around 50% in both arms). The **positive parenting frequency score** show that parents in the whole sample were doing slightly well as they scored a mean of just above the 50% of the maximum but the problem score showed parents doing worse as they scored closer to the minimum values. Harsh parenting scores were generally closer to the minimum showing better initial outcomes for the sample but one could argue that parents just didn't report (they were less inclined to tell if they were harsh). The data also shows high levels of self-reported **monitoring and supervision** (the **poor monitoring and supervision score** had a lower average), lower **depression** scores and relatively higher levels of **parenting stress**. Lastly, the observed **positive behaviours** seemed to be higher than the observed **negative behaviours** in general for both parents and children.

Table 2.5: **Baseline Scores/ Outcomes by Arm**

| | **Control Arm** (148) | **Intervention Arm** (148) |
|---|---|---|
| **Outcome** [range] | **n** (min, Q1, med, **mean**, Q3, max) | **n** (min, Q1, med, **mean**, Q3, max) |
| **Child Behaviour Problems** | | |
| ECBI intensity [36 - 252] | 148 (89,126,142.5,**143**,157.2,225) | 148 (89,125.8,142,**141.2**,155.2,201) |
| ECBI problem [0 - 36] | 148 (15,21,26,**25.16**,28,36) | 148 (15,20,25,**24.61**,29,36) |
| **Positive Parenting** | | |
| Supporting Positive Behaviour Frequency (subscore) [0 - 48] | 148 (6,23,27,**26.57**,30,42) | 148 (12,23,26,**25.76**,29,38) |
| Setting Limits Frequency (subscore) [0 - 42] | 148 (0,19,24,**22.66**,28,36) | 148 (4,18,22,**21.79**,25.25,38) |
| Positive Parenting Frequency [0 - 90] | 148 (4,40,46,**44.42**,51,68) | 148 (15,38,43,**42.8**,48.25,69) |
| Supporting Positive Behaviour Problem (subscore) [0 - 8] | 147 (0,1,2,**2.16**,3,7) | 148 (0,1,1,**1.9**,3,7) |
| Setting Limits Problem (subscore) [0 - 7] | 148 (0,0,2,**2.62**,4,7) | 146 (0,1,2,**2.5**,4,7) |
| Positive Parenting Problem [0 - 15] | 147 (0,1,4,**4.69**,8,13) | 146 (0,2,4,**4.29**,6,13) |
| **Harsh Parenting** | | |
| Non-Violent Discipline [0 - 20] | 148 (0,4,6,**6.34**,9,14) | 147 (0,5,6,**6.42**,8.5,14) |
| Physical Discipline [0 - 30] | 148 (0,3,6,**6.29**,9,24) | 147 (0,4,5,**6.03**,8,17) |
| Severe Physical Discipline [1] [0 - 15] | | |
| Overall | 148 (0,0,0,**0.2**,0,4) | 148 (0,0,0,**0.27**,0,4) |
| Zeros | 134 (**90.54**) | 124 (**83.78**) |
| Non-zeros | 14 ( 1.00, 1.25, 2.00, **2.14**, 2.75, 4.00) | 24 (1.00, 1.00, 1.00, **1.67**, 2.00, 4.00) |
| Psychological Discipline [0 - 50] | 148 (0,3,5,**6.91**,10,30) | 148 (0,3,6,**6.89**,9.25,22) |
| Neglect [1] [0 - 3] | | |
| Overall | 148 (0,0,0,**0.01**,0,1) | 148 (0,0,0,**0.05**,0,1) |
| Zeros | 147 (**99.32**) | 141 (**95.27**) |
| Non-zeros | 1 ( 1, 1, 1, **1**, 1, 1) | 7 ( 1, 1, 1, **1**, 1, 1) |
| **Observed Parenting and Child Behaviour** | | |
| Parent Positive Behaviour [0 - $\infty$) | 145 (0, 5, 10, **12.63**, 16, 52) | 147 (0, 5, 11, **15.27**, 22, 125) |
| Parent Negative Behaviour [0 - $\infty$) | 145 (0, 0, 2, **2.86**, 4, 17) | 147 (0, 1, 2, **3.14**, 4, 36) |
| Child Positive Behaviour [0 - $\infty$) | 145 (0, 9, 20, **27.94**, 47, 101) | 147 (0, 10, 22, **26.76**, 41, 87) |
| Child Negative Behaviour [0 - $\infty$) | | |
| Overall | 145 (0, 0, 0, **1.75**, 2, 17) | 147 (0, 0, 0, **1.71**, 2, 36) |
| Zeros | 73 (**50.34**) | 75 (**51.02**) |
| Non-zeros | 72 (1, 1, 2, **3.53**, 4, 17) | 72 (1, 1, 2, **3.49**, 4.25, 36) |
| **Monitoring and Supervision** | | |
| Poor Monitoring and Supervision [9 - 45] | 145 (9,16,19,**19.32**,23,34) | 146 (9,15.25,19,**19.79**,23.75,33) |
| **Depression** | | |
| Beck Depression Inventory[0 - 63] | 132 (0,6,12.5,**15.39**,23.25,46) | 140 (0,6,14.5,**15.74**,24,47) |
| **Parenting Stress** | | |
| Parental Distress (subscore) [12 - 60] | 141 (12,28,35,**33.7**,41,51) | 145 (16,28,34,**34.14**,41,49) |
| Parent Child Dysfunctional Interaction (subscore) [12 - 60] | 144 (12,37,43.5,**41.61**,48,57) | 146 (17,39.25,45,**42.91**,48,56) |
| Difficult Child (subscore) [12 - 60] | 144 (13,32,38,**36.82**,43,52) | 145 (17,34,38,**37.37**,42,53) |
| Parenting Stress [36 - 180] | 138 (44,100.2,116,**111.9**,126,153) | 143 (56,103.5,115,**114.6**,126,157) |
| **Social Support** | | |
| Social Support [8 - 40] | 146 (8,16,21,**20.58**,24,36) | 148 (8,17,22,**21.03**,24,34) |

[1] Removed from the analysis for very low reliability, see Section 2.2.

### 2.3.2 Summaries by loss to follow-up

During the course of the study, some participants were not available for the assessment either at post-test or one year follow-up for a variety of reasons. These participants were considered lost to follow-up in the study. There was a total of 30 participants that were categorized as being lost to follow-up during the study time period. Nine were lost at post-test, four of which were also lost to follow-up at the one year visit and an additional twenty one were lost to follow-up only at the one year mark.

Tables 2.6 and 2.7 compare the baseline characteristics and outcomes of participants that completed the study to those lost to follow-up, which for simplicity was defined as being lost to the study on at least one of the time points. There are generally very small differences between the group that was lost to follow-up and those that were not lost to follow-up except for a few. It seems that a higher proportion of caregivers that had drug problems were lost to follow-up (13% vs about 6%), this group also had higher incidences of reported maltreatment with physical abuse exhibiting the bigger difference between the two. The incidence of child negative behaviour also seemed higher in the lost to follow-up group (about 58%, here 'incidence' is calculated as (1 - proportion of zeros)100% ). There is an indication of higher unemployment levels in the lost to follow-up group (90%) compared to those that attended all three assessments (84.96%). Those who were lost to follow-up were also comparable to those who completed the study with respect to baseline outcome scores with the possible exception of a slightly higher incidence of severe physical discipline, lower child positive behaviour scores, higher parent negative behaviour scores and a higher incidence of child negative behaviour among the lost to follow-up group. These differences were all small but there is an indication that the lost to follow-up group had a slightly more negative profile than those who completed the study.

Table 2.6: **Baseline characteristics of child/ caregiver/ household for those who attended all 3 assessments versus those who were lost to follow-up**

| | **Not lost to follow-up** (266) | **Lost to follow-up** (30) |
|---|---|---|
| | **n** (min, Q1, med, **mean**, Q3, max) | **n** (min, Q1, med,**mean**, Q3, max) |
| **Age of caregiver** | 266 (18.00, 27.00, 32.00, **34.11**, 39.00, 75.00) | 30 (20.00, 25.00, 31.50, **33.20**, 37.00, 65.00) |
| **Age of child** | 266 (2.00, 4.00, 5.00, **5.30**, 7.00, 9.00) | 30 (2.00, 3.00, 4.00, **4.63**, 6.75, 9.00) |
| | n (**% of not lost to follow-up**) | n (**% of lost to follow-up**) |
| **HIV status of caregiver** | | |
| Positive | 68 (**25.56**) | 8 (**26.67**) |
| Negative | 187 (**70.30**) | 20 (**66.67**) |
| Unknown | 11 (**4.14**) | 2 (**6.66**) |
| **HIV status of child** | | |
| Positive | 3 (**1.13**) | 0 (**0.00**) |
| Negative | 152 (**57.14**) | 14 (**46.67**) |
| Unknown | 111 (**41.73**) | 16 (**53.33**) |
| **Family Structure** | | |
| Biological mother present in household | 33 (**12.41**) | 4 (**13.33**) |
| Biological father present in household | 121 (**45.49**) | 18 (**60.00**) |
| **Childs HIV orphanhood status** | | |
| Single orphan | 3 (**1.05**) | 0 (**0.00**) |
| Double Orphan | 0 (**0.00**) | 0 (**0.00**) |
| **Gender of caregiver** | | |
| Male | 1 (**0.38**) | 0 (**0.00**) |
| Female | 265 (**99.62**) | 30 (**100.00**) |
| **Gender of child** | | |
| Male | 141 (**53.01**) | 17 (**56.67**) |
| Female | 125 (**46.99**) | 13 (**43.33**) |
| **Caregiver's employment status** | | |
| Employed | 40 (**15.04**) | 3 (**10.00**) |
| Unemployed | 226 (**84.96**) | 27 (**90.00**) |
| **Caregiver substance use** | | |
| Alcohol | 82 (**30.83**) | 11 (**36.67**) |
| Drugs | 15 (**5.64**) | 4 (**13.33**) |
| **Caregiver experience of intimate partner violence** | | |
| IPV Incidence | 81 (**30.45**) | 8 (**26.67**) |
| **Caregiver experience of maltreatment** | | |
| Physical abuse | | |
| Overall | 243 (0.00, 0.00, 1.00, **0.75**, 1.00, 4.00) | 24 (0.00, 0.00, 1.00, **0.71**, 1.00, 2.00) |
| Zeros | 117 (**48.15**) | 10 (**41.67**) |
| Non-zeros | 126 (1.00, 1.00, 1.00, **1.45**, 2.00, 4.00) | 14 (1.00, 1.00, 1.00, **1.21**, 1.00, 2.00) |
| Emotional abuse | | |
| Overall | 235 (0.00, 0.00, 0.00, **0.58**, 1.00, 5.00) | 27 (0.00, 0.00, 0.00, **0.44**, 1.00, 2.00) |
| Zeros | 148 (**62.98**) | 16 (**59.26**) |
| Non-zeros | 87 (1.00, 1.00, 1.00, **1.56**, 2.00, 5.00) | 11 (1.00, 1.00, 1.00, **1.09**, 1.00, 2.00) |
| Sexual abuse | | |
| Overall | 252 (0.00, 0.00, 0.00, **0.11**, 0.00, 4.00) | 29 (0.00, 0.00, 0.00, **0.14**, 0.00, 2.00) |
| Zeros | 236 (**93.65**) | 26 (**89.66**) |
| Non-zeros | 16 (1.00, 1.00, 1.50, **1.75**, 2.00, 4.00) | 3 (1.00, 1.00, 1.00, **1.33**, 1.50, 2.00) |

Table 2.7: **Baseline scores by loss to follow-up (after baseline assessment)**

| | **Not lost to follow-up** (n = 266) | **Lost to follow-up** (n = 30) |
|---|---|---|
| | **n** (min, Q1, med, **mean**, Q3, max) | **n** (min, Q1, med, **mean**, Q3, max) |
| **Child Behaviour Problems** | | |
| ECBI Intensity | 266 ( 89 , 126 , 142.5 , **142.6** , 157 , 225 ) | 30 ( 96 , 120.2 , 134.5 , **137.3** , 152 , 208) |
| ECBI Problem | 266 ( 15 , 21 , 26 , **25.05** , 29 , 36 ) | 30 ( 15 , 18.25 , 24 , **23.4** , 27 , 34) |
| **Positive Parenting** | | |
| Supporting Positive Behaviour Frequency (subscore) | 266 ( 6 , 23 , 27 , **26.19** , 29 , 42 ) | 30 ( 8 , 23 , 26.5 , **26** , 30 , 34) |
| Setting Limits Frequency (subscore) | 266 ( 0 , 19 , 23 , **22.45** , 27 , 38 ) | 30 ( 4 , 16.25 , 21 , **20.23** , 26 , 34) |
| Positive Parenting Frequency | 266 ( 10 , 43 , 49 , **48.64** , 55 , 74 ) | 30 ( 15 , 41.25 , 48 , **46.23** , 53 , 64) |
| Supporting Positive Behaviour Problem (subscore) | 265 ( 0 , 1 , 1 , **2.05** , 3 , 7 ) | 30 ( 0 , 1 , 1 , **1.87** , 3 , 5) |
| Setting Limits Problem (subscore) | 264 ( 0 , 1 , 2 , **2.54** , 4 , 7 ) | 30 ( 0 , 1 , 2 , **2.67** , 4 , 7) |
| Positive Parenting Problem | 263 ( 0 , 1 , 4 , **4.59** , 7 , 14 ) | 30 ( 0 , 2 , 4 , **4.53** , 6 , 12) |
| **Harsh Parenting** | | |
| Non-Violent Discipline | 265 ( 0 , 4 , 6 , **6.44** , 9 , 14 ) | 30 ( 0 , 4.25 , 5.5 , **5.9** , 8.75 , 13) |
| Physical Discipline | 265 ( 0 , 2 , 4 , **4.45** , 6 , 22 ) | 30 ( 0 , 2 , 5 , **4.4** , 6 , 10) |
| Severe Physical Discipline | | |
| Overall | 266 ( 0 , 0 , 0 , **0.22** , 0 , 4 ) | 30 ( 0 , 0 , 0 , **0.4** , 0.75 , 3) |
| Zeros | 236 (**88.72**) | 22 (**73.33**) |
| Non-zeros | 30 (1, 1, 2, **1.93**, 2, 4) | 8 (1, 1, 1, **1.5**, 2, 3 ) |
| Psychological Discipline | 266 ( 0 , 3 , 5 , **6.7** , 9 , 30 ) | 30 ( 0 , 3 , 5.5 , **6.67** , 9 , 20) |
| Neglect | | |
| Overall | 266 ( 0 , 0 , 0 , **0.03** , 0 , 1 ) | 30 ( 0 , 0 , 0 , **0.03** , 0 , 1) |
| Zeros | 259 (**97.37**) | 29 (**96.67**) |
| Non-zeros | 7 (1, 1, 1, **1**, 1, 1) | 1 (1, 1, 1, **1**, 1, 1) |
| **Observed Parenting and Child Behaviour** | | |
| Parent Positive Behaviour | 263 ( 0 , 5 , 10 , **13.64** , 19 , 59 ) | 29 ( 1 , 5 , 10 , **16.9** , 18 , 125) |
| Child Positive Behaviour | 263 ( 0 , 9 , 22 , **27.9** , 45.5 , 101 ) | 29 ( 0 , 9 , 20 , **22.28** , 31 , 70) |
| Parent Negative Behaviour | 263 ( 0 , 1 , 2 , **2.92** , 4 , 25 ) | 29 ( 0 , 1 , 2 , **3.72** , 4 , 36) |
| Child Negative Behaviour | | |
| Overall | 263 ( 0 , 0 , 0 , **1.74** , 2 , 36 ) | 29 ( 0 , 0 , 1 , **1.62** , 1 , 11) |
| Zeros | 136 (**51.71**) | 12 (**41.38**) |
| Non-zeros | 127 (1, 1, 2, **3.61**, 4.5, 36) | 17 ( 1, 1, 1, **2.77**, 4, 11) |
| **Monitoring and Supervision** | | |
| Poor Monitoring and Supervision | 261 ( 9 , 19 , 22 , **22.18** , 25 , 37 ) | 30 ( 15 , 17 , 19 , **20.57** , 22.75 , 29) |
| **Depression** | | |
| Beck Depression Inventory | 244 ( 0 , 6 , 13 , **15.48** , 24 , 47 ) | 28 ( 0 , 6.75 , 14.5 , **16.32** , 24.25 , 41) |
| **Parenting Stress** | | |
| Parental Distress (subscore) | 257 ( 12 , 28 , 34 , **33.96** , 41 , 51 ) | 29 ( 21 , 28 , 33 , **33.59** , 39 , 49) |
| Parent-Child Dysfunctional Interaction (subscore) | 261 ( 12 , 38 , 44 , **42.16** , 48 , 57 ) | 29 ( 25 , 39 , 46 , **43.24** , 49 , 55) |
| Difficult Child (subscore) | 260 ( 13 , 33 , 38 , **36.95** , 42 , 52 ) | 29 ( 18 , 32 , 36 , **38.38** , 46 , 53) |
| Parenting Stress | 252 ( 44 , 101 , 116 , **113.1** , 126 , 153 ) | 29 ( 75 , 106 , 113 , **115.2** , 132 , 157) |
| **Social Support** | | |
| Social Support | 265 ( 8 , 16 , 21 , **20.91** , 24 , 36 ) | 29 ( 8 , 17 , 21 , **19.9** , 23 , 35) |

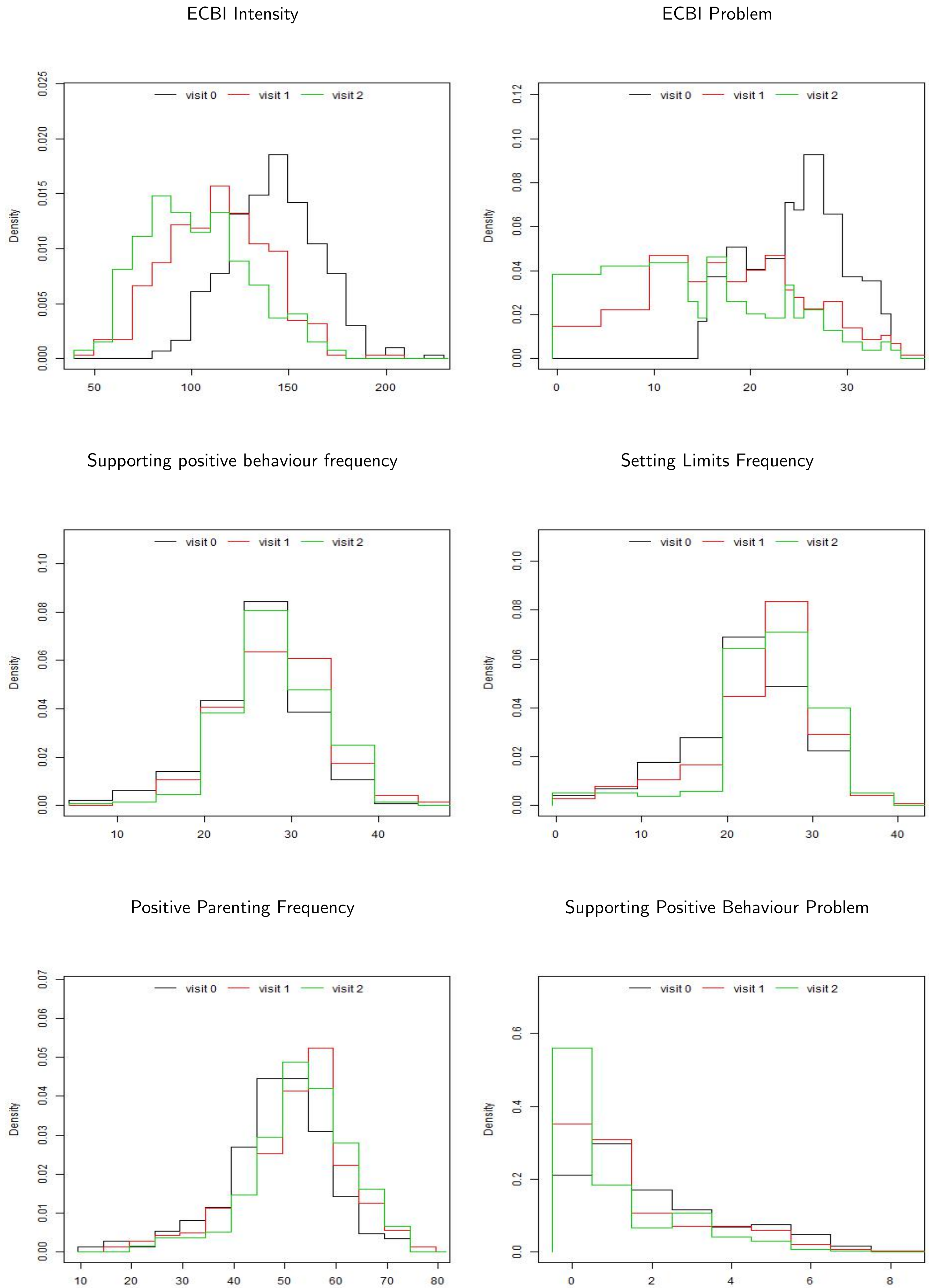
## 2.4 Distribution of the outcomes over time

The tables and discussion in Section 2.3 give a slight indication of the possible distributions of the outcomes of interest at baseline. Responses with means higher than medians would have right skewed distributions whilst those with means lower than medians would have left skewed distributions. The histograms in Figure 2.1 below show the empirical distributions of the outcomes of interest at the three time points. Ideally, if the intervention program had worked, one should observe the histograms moving to the left with time for all negative outcomes (those where higher scores are less desirable e.g. **child behaviour problems** and **harsh parenting**) and vice versa for positive outcomes.

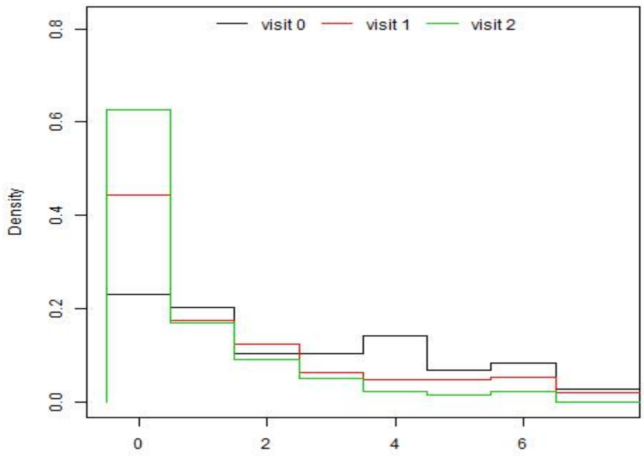
The distributions of the **child behaviour problem** scores seem to be roughly symmetrical and also moved in the desired direction over the three time points. This is to be expected since this was the screening tool and therefore left-censored at baseline. The **physical discipline** score, all the problem scores on **positive parenting**, **psychological discipline**, all the **observed scores** and the **Beck Depression Inventory** have tails to the right. The distributions/ histograms of **neglect** and **severe physical discipline** indicate an excess of zeros and a very few non-zero values. These outcomes were not analysed due to their lack in reliability.

Among the **harsh parenting** scores, only **non-violent discipline** and **psychological discipline** seem to be moving in the right direction over time since the histograms show that the peak of the observed distributions moved to lower scores for the later time points. All the scores in this group look right skewed and severe physical discipline and neglect look zero inflated at all time points as well. All **observational** scores in Figure 2.1 look right skewed and its difficult to tell from the histograms whether the program had an impact based on these (the distributions did not shift much over time). **Poor monitoring and supervision** and **social support** scores look roughly symmetric, the **Beck Depression Inventory (BDI)** looks right skewed and all **parenting stress** outcomes look left skewed over the three time points. **BDI** and the **parenting stress** outcomes seem to be moving in the right direction, the others aren't clearly showing any noticeable changes.

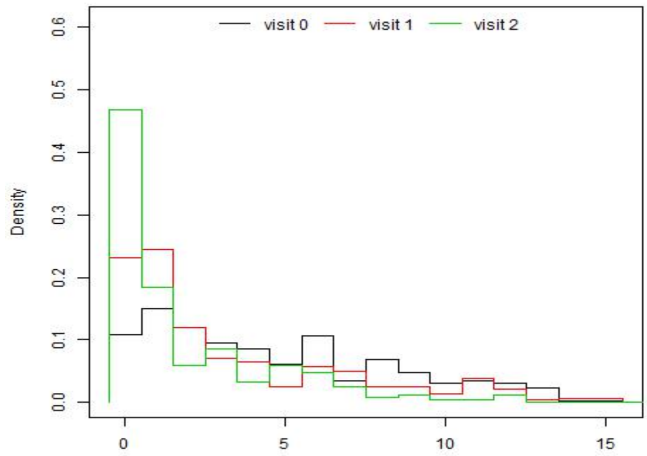
Figure 2.1: **Empirical distributions of outcomes over time**



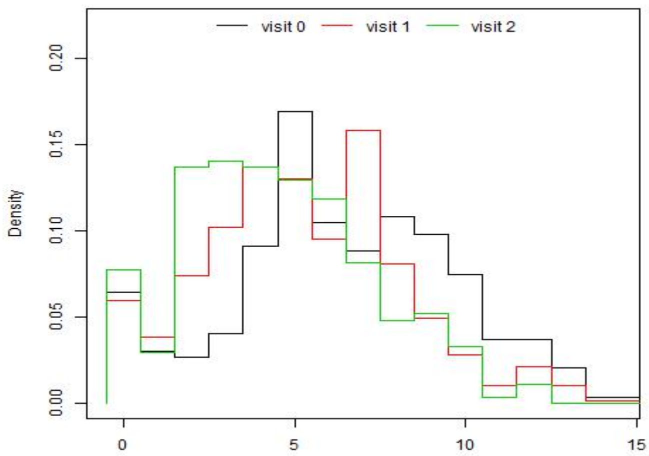
Setting Limits Problem



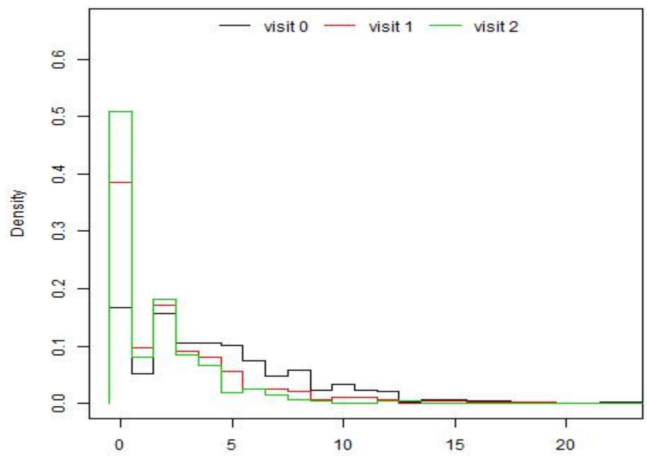

Positive Parenting Problem



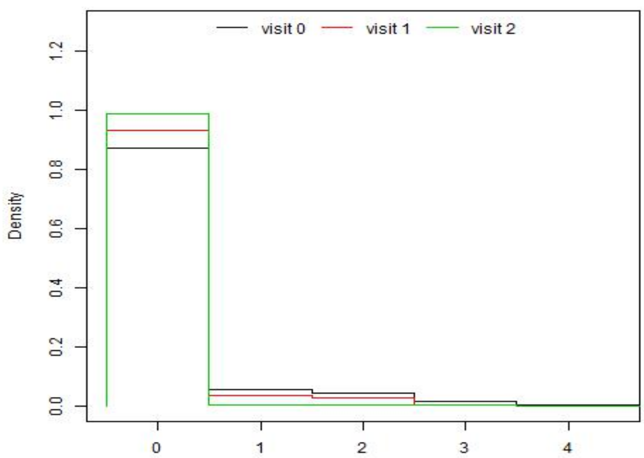

Non-Violent Discipline



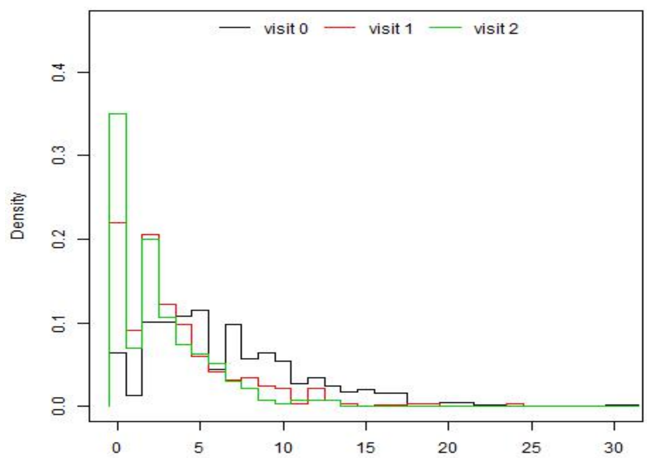

Physical Discipline



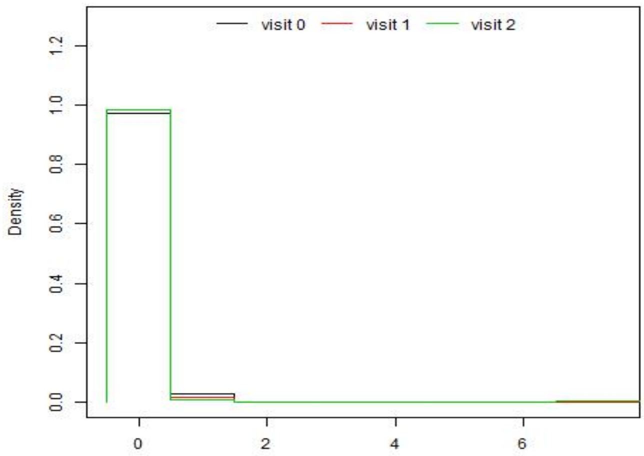

Severe Physical Discipline



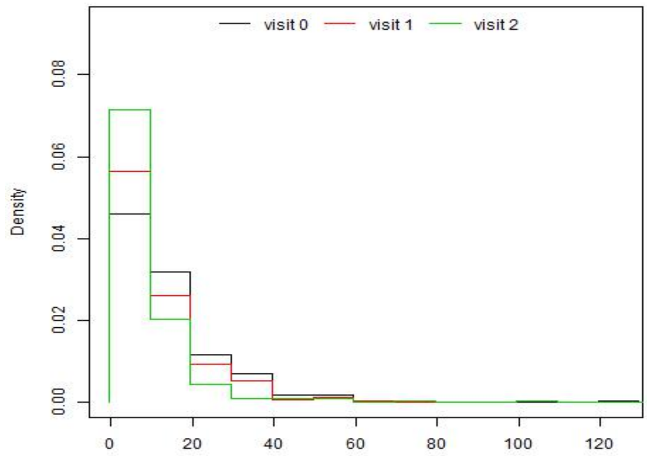

Psychological Discipline



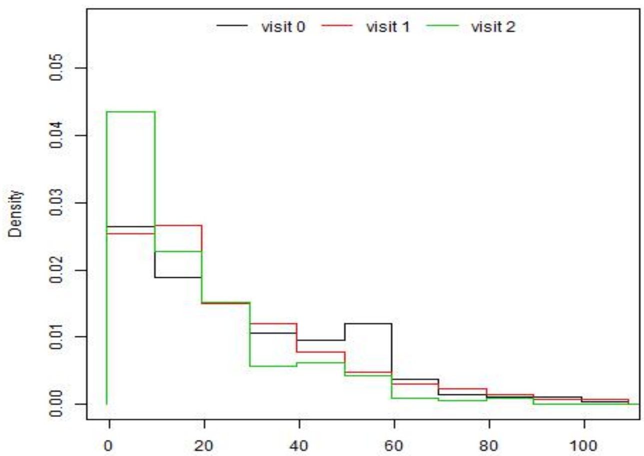
Neglect



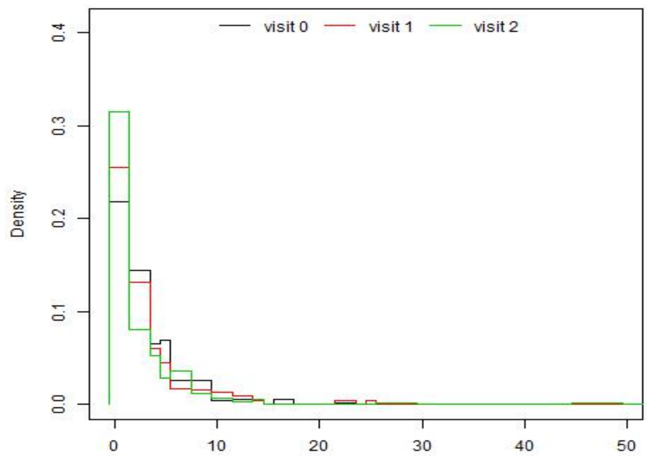

Parent Positive Behavoiur



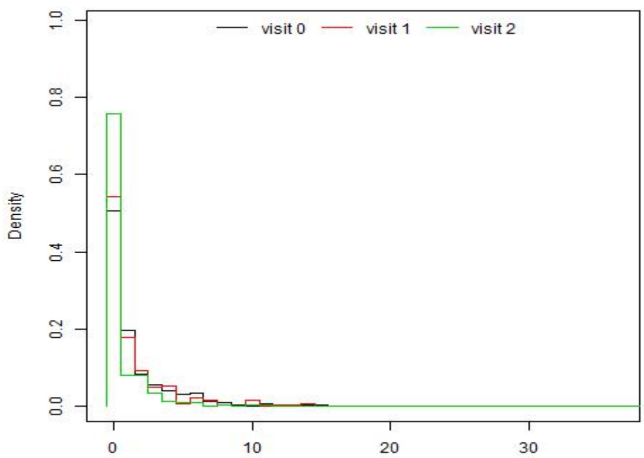

Child Positive Behavoiur



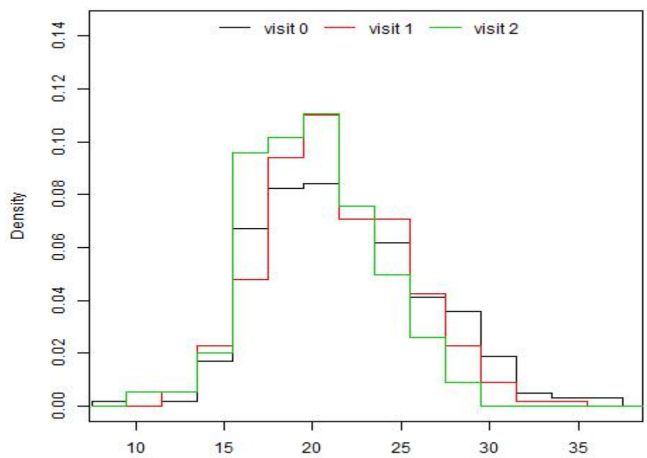

Parent Negative Behavoiur



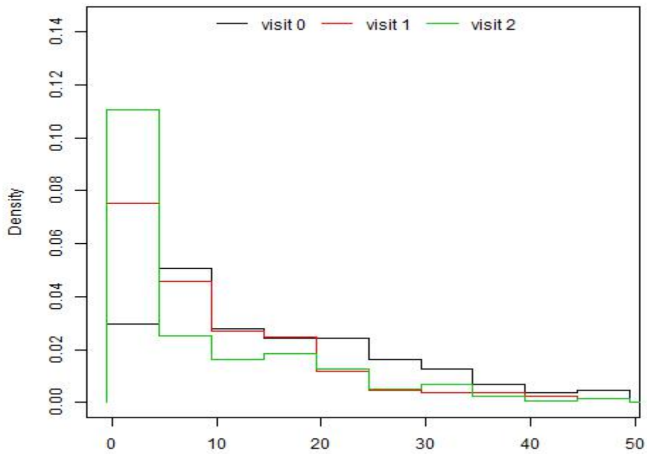

Child Negative Behavoiur



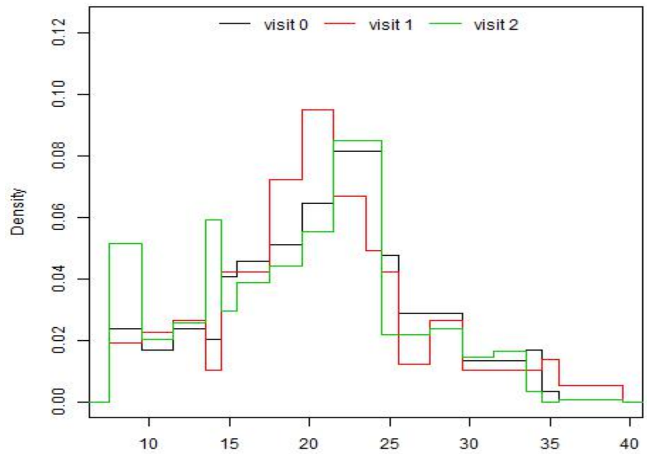

Poor Monitoring & Supervision



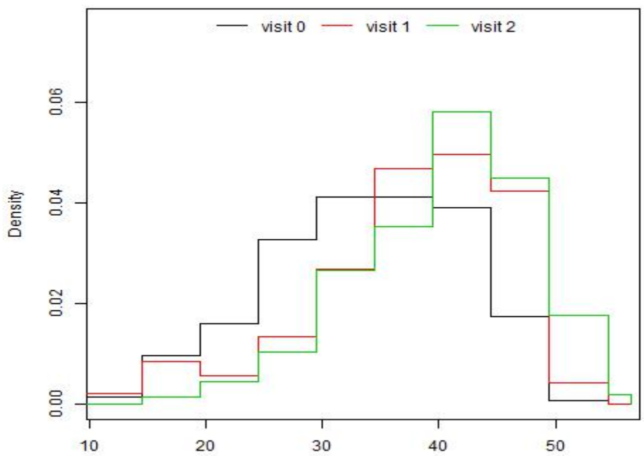
Beck Depression Inventory



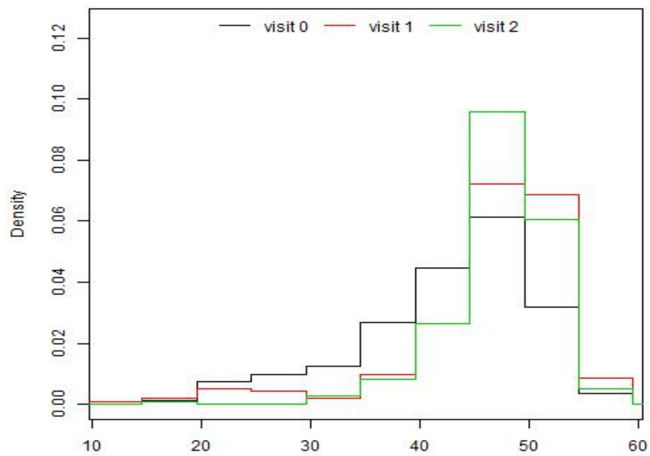

Social Support



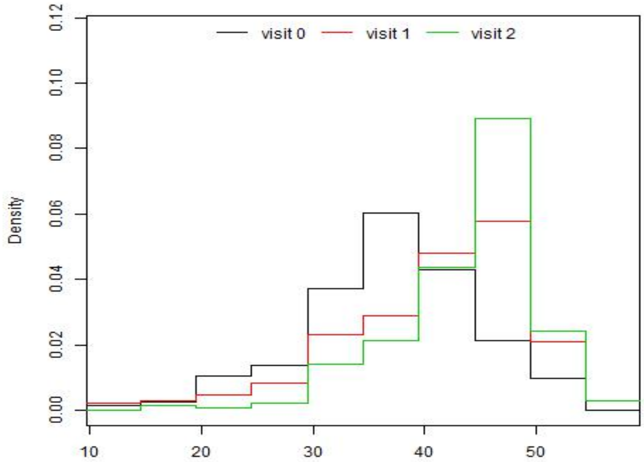

Parental Distress



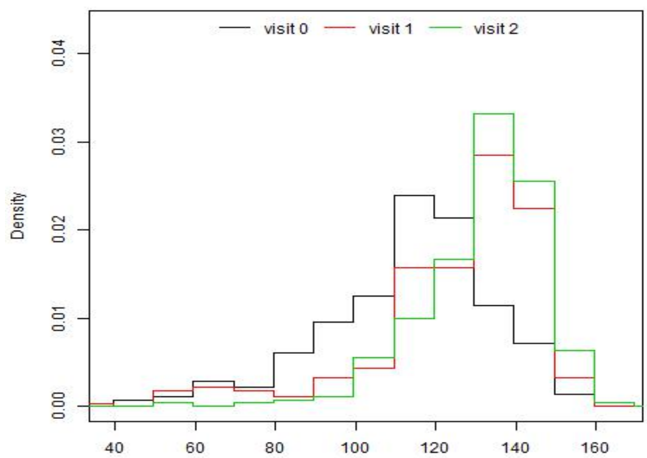

Parent Child Dysfunctional Interraction




Difficult Child




Parenting Stress



## 2.5 The extent of missingness in the data

There are a number of ways in which data could be missing in the study. One of the ways discussed earlier is through loss to follow-up through which 9 complete records are missing at the post-test visit and a further 25 at one year follow-up visit. The reasons for missing data were manifold, including loss to follow-up, non-attendance, loss/ damage of videos and many more. Though participants were encouraged to complete all the questionnaires as far as possible (especially at baseline, as a pre-requisite for selection into the RCT), there were cases where they would either refuse to disclose some information (for example their HIV status, see Table 2.4) or at times they would just not know.

Table 2.8 below summarizes the number of missing data points as a percentage of the cohort size (296) on each outcome of interest over the three time points. As expected, the degree of missing data increased with time, i.e. at baseline, most of the data were captured and as the trial proceeded further from enrolment, the amount of missing data increased. Only 4 (1.35%) sets of observational scores were missing at baseline, this number increased to 29 (9.80%) then 85 (28.72 %) at post-test and at one year follow-up respectively. Data errors like lost or faulty videos were among the reasons for the high missingness in the video-coded (observational) data. Of all the self reported scores, the **BDI** had the highest number of missing values at all time points. It also seems that the missing data problem is more prevalent in the **secondary outcomes** especially at baseline.

Table 2.8: **Degree of missingness by time point**

| | Missing at baseline | Missing at post-test | Missing at follow-up |
|---|---|---|---|
| **Primary Outcomes** | | | |
| | N (%) | N (%) | N (%) |
| ECBI Intensity | 0 (**0.00**) | 9 (**3.04**) | 26 (**8.78**) |
| ECBI Problem | 0 (**0.00**) | 9 (**3.04**) | 26 (**8.78**) |
| Supporting Positive Behaviour Frequency | 0 (**0.00**) | 10 (**3.38**) | 25 (**8.45**) |
| Setting Limits Frequency | 0 (**0.00**) | 9 (**3.04**) | 25 (**8.45**) |
| Positive Parenting Frequency | 0 (**0.00**) | 10 (**3.38**) | 25 (**8.45**) |
| Supporting Positive Behaviour Problem | 1 (**0.34**) | 14 (**4.73**) | 25 (**8.45**) |
| Setting Limits Problem | 2 (**0.68**) | 10 (**3.38**) | 25 (**8.45**) |
| Positive Parenting Problem | 3 (**1.01**) | 14 (**4.73**) | 25 (**8.45**) |
| Non-Violent Discipline | 1 (**0.34**) | 12 (**4.05**) | 26 (**8.78**) |
| Physical Discipline | 1 (**0.34**) | 10 (**3.38**) | 25 (**8.45**) |
| Severe Physical Discipline | 0 (**0.00**) | 9 (**3.04**) | 25 (**8.45**) |
| Psychological Discipline | 0 (**0.00**) | 10 (**3.38**) | 25 (**8.45**) |
| Neglect | 0 (**0.00**) | 9 (**3.04**) | 25 (**8.45**) |
| Parent Positive Behaviour | 4 (**1.35**) | 29 (**9.80**) | 85 (**28.72**) |
| Child Positive Behaviour | 4 (**1.35**) | 29 (**9.80**) | 85 (**28.72**) |
| Parent Negative Behaviour | 4 (**1.35**) | 29 (**9.80**) | 85 (**28.72**) |
| Child Negative Behaviour | 4 (**1.35**) | 29 (**9.80**) | 85 (**28.72**) |
| **Secondary Outcomes** | | | |
| Poor Monitoring And Supervision | 5 (**1.69**) | 14 (**4.73**) | 25 (**8.45**) |
| Beck Depression Inventory | 24 (**8.11**) | 28 (**9.46**) | 27 (**9.12**) |
| Parental Distress | 10 (**3.38**) | 14 (**4.73**) | 25 (**8.45**) |
| Parent Child Dysfunctional Interaction | 6 (**2.03**) | 14 (**4.73**) | 25 (**8.45**) |
| Difficult Child | 7 (**2.36**) | 12 (**4.05**) | 25 (**8.45**) |
| Social Support | 2 (**0.68**) | 12 (**4.05**) | 25 (**8.45**) |

One important conclusion from Table 2.8 is that overall, the amount of missing data is considerably low. The measurement models to be discussed and fitted in Chapter 4 will only include one of the outcomes in this table at a time. A model evaluating the effect of the intervention program on an outcome like **ECBI intensity** would have a total of 35 missing items on a long format response with 888 (296 $\times$ 3) items. Of all the self reported outcomes of interest, the **Beck Depression Inventory** has the highest levels of missingness averaging about 9 % over the three time points. The four observational scores have the highest overall missingness with 115 of the combined 888 (i.e. 13% missing) data items missing on each of the responses.

In most instances where data was missing on a particular score, data would also be missing on all the items that were added up to get the score. This was mainly because of loss to follow-up whereby no data would be collected on the participant. Additionally, there were a few cases where the missingness was on an item level, i.e. the overall score would have a missing value because a subset of the items making up the score had missing data. By way of an example, the **poor monitoring and supervision** score was constructed using the following nine (5-point Likert-type) items:

**Item 1**. "Lets you know where going";

**Item 2**. "Stay out in evening";

**Item 3**. "Friends you don't know";

**Item 4**. "No set time to be home";

**Item 5**. "Out after dark without adult";

**Item 6**. "Forgot what child doing";

**Item 7**. "Check home time";

**Item 8**. "Tell child where going" and

**Item 9**. "Home without adult".

Tables 2.9, 2.10 and 2.11 show the missing data pattens for the aforementioned items at the three study visits. The first column shows the number of participants, the last column shows the number of items on which the participants have missing data and the last row shows the number of participants with missing data on each item. At baseline, the **poor monitoring and supervision** score has 5 missing data points and the number increases to 14 and 25 at post-test and one year follow-up visit respectively. Table 2.9 shows that of the 5 with missing data at baseline, one is missing data only on item 2, two have missing data only on item 3, one has missing data on items 2 & 7 and the last one has missing data on item 2 up to item 6. Table 2.10 confirms that of the fourteen with missing data on this score, nine have were missing completely for all the items (these were part of the lost to follow-up group) and the rest have missing data on arbitrary items. The 25 missing at one year follow-up visit are because of loss to follow-up. For these participants, data were missing completely for all the items (see Table 2.11) and the overall score.

Table 2.9: **Missingness patterns for the poor monitoring and supervision score at the baseline visit**

| Number of Participants | Item 1 | Item 8 | Item 9 | Item 4 | Item 5 | Item 6 | Item 7 | Item 2 | Item 3 | Number of missing items |
|---|---|---|---|---|---|---|---|---|---|---|
| 291 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 0 |
| 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 0 | 1 | 1 |
| 2 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 0 | 1 |
| 1 | 1 | 1 | 1 | 1 | 1 | 1 | 0 | 0 | 1 | 2 |
| 1 | 1 | 1 | 1 | 0 | 0 | 0 | 1 | 0 | 0 | 5 |
| | 0 | 0 | 0 | 1 | 1 | 1 | 1 | 3 | 3 | 10 |

Table 2.10: **Missingness patterns for the poor monitoring and supervision score at the post-test visit**

| Number of Participants | Item 3 | Item 4 | Item 6 | Item 7 | Item 8 | Item 2 | Item 5 | Item 1 | Item 9 | Number of missing items |
|---|---|---|---|---|---|---|---|---|---|---|
| 282 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 0 |
| 2 | 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 | 0 | 1 |
| 1 | 1 | 1 | 1 | 1 | 1 | 1 | 0 | 1 | 0 | 2 |
| 9 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 9 |
| | 9 | 9 | 9 | 9 | 9 | 10 | 10 | 11 | 11 | 87 |

Table 2.11: **Missingness patterns for the poor monitoring and supervision score at the one year follow-up visit**

| Number of Participants | Item 1 | Item 2 | Item 3 | Item 4 | Item 5 | Item 6 | Item 7 | Item 8 | Item 9 | Number of missing items |
|---|---|---|---|---|---|---|---|---|---|---|
| 271 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 0 |
| 25 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 9 |
| | 25 | 25 | 25 | 25 | 25 | 25 | 25 | 25 | 25 | 225 |

## 2.6 Missing data patterns

The **md.pattern()** function in the **R** package **mice** (Van Buuren and Groothuis-Oudshoorn, 2011) was used to produce tables 2.12, 2.13 and 2.14. The top row of these tables show the numbers of participants and the bottom row show the number of outcomes for which the participants counted in the top row have missing data. The ones represent a recorded outcome whilst a zero (highlighted) corresponds with a missing data item. The last column summarizes missingness on each outcome just as in Table 2.8.

At baseline, there are 261 participants with complete information in the outcomes of interest. Nine participants have every record except the **BDI** score and among all participants, at most 6 outcomes had missing values at baseline. It also seems that there is a correlation between missing values for **parenting stress** outcomes and missing values for the **BDI** score.

There is an increased rate of missingness at the post-test visit. A total of 240 participants have complete records and the 9 that were lost to follow-up at this time point are shown in the column with 25 highlighted zeros in Table 2.13. **Observational** scores have the largest rate of missingness at this point followed by the four **positive parenting** outcomes. Only those lost to follow-up had missing values on all self-reported scores and 20 of the participants that did not have observational data still had data recorded on the self reports.

At one year follow-up, the rate of missingness further increased with 207 participants having complete records on the outcomes of interest. As stated before, the increasing missingness might be a result of the fact that at baseline records had to be as complete as possible for eligibility into the RCT. There are 25 patients who missed home visits and thus classified as lost to follow-up though one of them still attended the group session. There are 61 participants that didn't attend group sessions at one-year follow-up but had successful home visits.

Table 2.15 explores the missing data patterns for the observational outcomes at the three time points. There are 200 participants that attended all the group sessions and therefore had their observational data recorded at all the time points. One participant didn't attend any group session and the other three that didn't attend at baseline had attendances at post-test and/ or follow-up. Of all the 29 participants who had missing observational data at post test, only two had missing data at baseline as well . Of the 85 participants with missing observational data at one year follow-up only one had no data at any of the earlier time points.

Table 2.16 shows the missing data patterns for all the outcomes of interest (reported and observed) over the three time points. Overall, 183 (61.82%) participants had complete information on all the data items collected in the study. There are 25 (8.45%) participants (last 3 columns) who have the most missing data with at least 42 missing items across all the study variables.

Table 2.12: **Missing data patterns at baseline**

| | 261 | 1 | 1 | 9 | 1 | 2 | 3 | 1 | 3 | 1 | 1 | 3 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| ECBI intensity | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 0 |
| ECBI problem | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 0 |
| Supporting Positive Behaviour Frequency | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 0 |
| Setting Limits Frequency | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 0 |
| Severe Physical Discipline | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 0 |
| Psychological Discipline | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 0 |
| Neglect | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 0 |
| Positive Parenting Frequency | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 0 |
| Supporting Positive Behaviour Problem | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 0 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 |
| Non-Violent Discipline | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 0 | 1 | 1 |
| Physical Discipline | 1 | 0 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 |
| Setting Limits Problem | 1 | 1 | 1 | 1 | 1 | 0 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 2 |
| Social Support | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 0 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 0 | 2 |
| ipv Chronicity | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 0 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 0 | 1 | 1 | 0 | 1 | 1 | 3 |
| Positive Parenting Problem | 1 | 1 | 1 | 1 | 1 | 0 | 1 | 1 | 1 | 1 | 1 | 1 | 0 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 3 |
| Parent Positive Behaviour | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 0 | 1 | 1 | 1 | 1 | 0 | 1 | 1 | 1 | 1 | 4 |
| Child Positive Behaviour | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 0 | 1 | 1 | 1 | 1 | 0 | 1 | 1 | 1 | 1 | 4 |
| Parent Negative Behaviour | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 0 | 1 | 1 | 1 | 1 | 0 | 1 | 1 | 1 | 1 | 4 |
| Child Negative Behaviour | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 0 | 1 | 1 | 1 | 1 | 0 | 1 | 1 | 1 | 1 | 4 |
| Poor Monitoring And Supervision | 1 | 1 | 0 | 1 | 0 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 0 | 0 | 1 | 1 | 1 | 1 | 1 | 0 | 1 | 5 |
| Parent Child Dysfunctional Interraction | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 0 | 1 | 1 | 1 | 0 | 1 | 1 | 1 | 0 | 0 | 0 | 0 | 6 |
| Difficult Child | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 0 | 1 | 1 | 1 | 0 | 0 | 1 | 0 | 0 | 0 | 0 | 7 |
| Parental Distress | 1 | 1 | 1 | 1 | 1 | 1 | 0 | 1 | 0 | 1 | 1 | 1 | 1 | 1 | 0 | 1 | 1 | 0 | 1 | 0 | 0 | 10 |
| Parenting Stress | 1 | 1 | 1 | 1 | 1 | 1 | 0 | 1 | 0 | 0 | 0 | 1 | 1 | 0 | 0 | 0 | 1 | 0 | 0 | 0 | 0 | 15 |
| Beck Depression Inventory | 1 | 1 | 1 | 0 | 0 | 1 | 1 | 0 | 0 | 0 | 0 | 1 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 1 | 0 | 24 |
| | 0 | 1 | 1 | 1 | 2 | 2 | 2 | 3 | 3 | 3 | 3 | 4 | 4 | 4 | 4 | 4 | 5 | 5 | 5 | 6 | 6 | 96 |

Table 2.13: **Missing data patterns: Post Test**

| | 240 | 2 | 1 | 13 | 2 | 2 | 1 | 16 | 1 | 2 | 1 | 1 | 1 | 2 | 1 | 1 | 9 | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| ECBI intensity | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 0 | 9 |
| ECBI problem | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 0 | 9 |
| SettingLimitsFrequency | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 0 | 9 |
| SeverePhysicalDiscipline | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 0 | 9 |
| Neglect | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 0 | 9 |
| SupportingPositiveBehaviourFrequency | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 0 | 1 | 0 | 10 |
| SettingLimitsProblem | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 0 | 0 | 10 |
| PhysicalDiscipline | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 0 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 0 | 10 |
| PsychologicalDiscipline | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 0 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 0 | 10 |
| PositiveParentingFrequency | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 0 | 1 | 0 | 10 |
| NonViolentDiscipline | 1 | 0 | 1 | 1 | 1 | 1 | 1 | 1 | 0 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 0 | 12 |
| DifficultChild | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 0 | 1 | 0 | 0 | 12 |
| SocialSupport | 1 | 1 | 1 | 1 | 0 | 1 | 1 | 1 | 1 | 1 | 0 | 1 | 1 | 1 | 1 | 1 | 0 | 12 |
| ipvChronicity | 1 | 1 | 1 | 1 | 1 | 0 | 1 | 1 | 1 | 1 | 0 | 1 | 1 | 1 | 1 | 0 | 0 | 13 |
| SupportingPositiveBehaviourProblem | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 0 | 1 | 1 | 0 | 0 | 0 | 0 | 14 |
| PoorMonitoringAndSupervision | 1 | 1 | 0 | 1 | 1 | 1 | 1 | 1 | 0 | 0 | 1 | 0 | 1 | 1 | 1 | 1 | 0 | 14 |
| ParentalDistress | 1 | 1 | 1 | 1 | 1 | 1 | 0 | 1 | 1 | 1 | 1 | 0 | 1 | 0 | 1 | 0 | 0 | 14 |
| ParentChilDysfunctionalInterraction | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 0 | 0 | 0 | 1 | 0 | 0 | 14 |
| PositiveParentingProblem | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 0 | 1 | 1 | 0 | 0 | 0 | 0 | 14 |
| ParentingStress | 1 | 1 | 1 | 1 | 1 | 1 | 0 | 1 | 1 | 1 | 1 | 0 | 0 | 0 | 1 | 0 | 0 | 15 |
| BeckDepressionInventory | 1 | 1 | 1 | 0 | 1 | 1 | 1 | 1 | 0 | 1 | 0 | 0 | 1 | 0 | 1 | 0 | 0 | 28 |
| Parent Positive Behaviour | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 0 | 1 | 0 | 1 | 1 | 0 | 1 | 0 | 1 | 0 | 29 |
| Child Positive Behaviour | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 0 | 1 | 0 | 1 | 1 | 0 | 1 | 0 | 1 | 0 | 29 |
| Parent Negative Behaviour | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 0 | 1 | 0 | 1 | 1 | 0 | 1 | 0 | 1 | 0 | 29 |
| Child Negative Behaviour | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 0 | 1 | 0 | 1 | 1 | 0 | 1 | 0 | 1 | 0 | 29 |
| | 0 | 1 | 1 | 1 | 1 | 1 | 2 | 4 | 5 | 5 | 5 | 5 | 6 | 7 | 8 | 9 | 25 | 373 |

Table 2.14: **Missing data patterns: follow-up**

| | 207 | 1 | 2 | 60 | 1 | 1 | 24 | |
|---|---|---|---|---|---|---|---|---|
| SupportingPositiveBehaviourFrequency | 1 | 1 | 1 | 1 | 1 | 0 | 0 | 25 |
| SettingLimitsFrequency | 1 | 1 | 1 | 1 | 1 | 0 | 0 | 25 |
| SupportingPositiveBehaviourProblem | 1 | 1 | 1 | 1 | 1 | 0 | 0 | 25 |
| SettingLimitsProblem | 1 | 1 | 1 | 1 | 1 | 0 | 0 | 25 |
| PhysicalDiscipline | 1 | 1 | 1 | 1 | 1 | 0 | 0 | 25 |
| SeverePhysicalDiscipline | 1 | 1 | 1 | 1 | 1 | 0 | 0 | 25 |
| PsychologicalDiscipline | 1 | 1 | 1 | 1 | 1 | 0 | 0 | 25 |
| Neglect | 1 | 1 | 1 | 1 | 1 | 0 | 0 | 25 |
| PoorMonitoringAndSupervision | 1 | 1 | 1 | 1 | 1 | 0 | 0 | 25 |
| ParentalDistress | 1 | 1 | 1 | 1 | 1 | 0 | 0 | 25 |
| ParentChilDysfunctionalInterraction | 1 | 1 | 1 | 1 | 1 | 0 | 0 | 25 |
| DifficultChild | 1 | 1 | 1 | 1 | 1 | 0 | 0 | 25 |
| SocialSupport | 1 | 1 | 1 | 1 | 1 | 0 | 0 | 25 |
| ipvChronicity | 1 | 1 | 1 | 1 | 1 | 0 | 0 | 25 |
| PositiveParentingFrequency | 1 | 1 | 1 | 1 | 1 | 0 | 0 | 25 |
| PositiveParentingProblem | 1 | 1 | 1 | 1 | 1 | 0 | 0 | 25 |
| ParentingStress | 1 | 1 | 1 | 1 | 1 | 0 | 0 | 25 |
| ECBI intensity | 1 | 1 | 1 | 1 | 0 | 0 | 0 | 26 |
| ECBI problem | 1 | 1 | 1 | 1 | 0 | 0 | 0 | 26 |
| NonViolentDiscipline | 1 | 0 | 1 | 1 | 1 | 0 | 0 | 26 |
| BeckDepressionInventory | 1 | 1 | 0 | 1 | 1 | 0 | 0 | 27 |
| Parent Positive Behaviour | 1 | 1 | 1 | 0 | 0 | 1 | 0 | 85 |
| Child Positive Behaviour | 1 | 1 | 1 | 0 | 0 | 1 | 0 | 85 |
| Parent Negative Behaviour | 1 | 1 | 1 | 0 | 0 | 1 | 0 | 85 |
| Child Negative Behaviour | 1 | 1 | 1 | 0 | 0 | 1 | 0 | 85 |
| | 0 | 1 | 1 | 4 | 6 | 21 | 25 | 870 |

Table 2.15: **Missing Data Patterns: Observed Scores**

| | 200 | 1 | 9 | 65 | 1 | 1 | 18 | 1 | |
|---|---|---|---|---|---|---|---|---|---|
| Parent Positive Behaviour-baseline | 1 | 0 | 1 | 1 | 0 | 0 | 1 | 0 | 4 |
| Parent Negative Behaviour-baseline | 1 | 0 | 1 | 1 | 0 | 0 | 1 | 0 | 4 |
| Child Positive Behaviour-baseline | 1 | 0 | 1 | 1 | 0 | 0 | 1 | 0 | 4 |
| Child Negative Behaviour-baseline | 1 | 0 | 1 | 1 | 0 | 0 | 1 | 0 | 4 |
| Parent Positive Behaviour-post test | 1 | 1 | 0 | 1 | 0 | 1 | 0 | 0 | 29 |
| Parent Negative Behaviour-post test | 1 | 1 | 0 | 1 | 0 | 1 | 0 | 0 | 29 |
| Child Positive Behaviour-post test | 1 | 1 | 0 | 1 | 0 | 1 | 0 | 0 | 29 |
| Child Negative Behaviour-post test | 1 | 1 | 0 | 1 | 0 | 1 | 0 | 0 | 29 |
| Parent Positive Behaviour-follow-up | 1 | 1 | 1 | 0 | 1 | 0 | 0 | 0 | 85 |
| Parent Negative Behaviour-follow-up | 1 | 1 | 1 | 0 | 1 | 0 | 0 | 0 | 85 |
| Child Positive Behaviour-follow-up | 1 | 1 | 1 | 0 | 1 | 0 | 0 | 0 | 85 |
| Child Negative Behaviour-follow-up | 1 | 1 | 1 | 0 | 1 | 0 | 0 | 0 | 85 |
| | 0 | 4 | 4 | 4 | 8 | 8 | 8 | 12 | 472 |

Table 2.16: **Missing data patterns: all time points**

| | 183 | 1 | 1 | 6 | 1 | 3 | 2 | 3 | 1 | 2 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 55 | 1 | 2 | 1 | 2 | 1 | 23 | 1 | . |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| ECBI intensity | 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 |
| ECBI problem | 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 |
| Supporting Positive Behaviour Frequency | 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 |
| Setting Limits Frequency | 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 |
| Severe Physical Discipline | 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 |
| Psychological Discipline | 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 |
| Neglect | 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 |
| Positive Parenting Frequency | 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 |
| Supporting Positive Behaviour Problem | 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 |
| Non Violent Discipline | 1 | 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 |
| Physical Discipline | 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 |
| Setting Limits Problem | 1 | 1 | 1 | 1 | 0 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 0 | 1 | 1 | 1 | 1 | 1 | 1 | 2 |
| Social Support | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 0 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 0 | 2 |
| ipv Chronicity | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 0 | 1 | 0 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 3 |
| Positive Parenting Problem | 1 | 1 | 1 | 1 | 0 | 1 | 1 | 1 | 1 | 1 | 0 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 0 | 1 | 1 | 1 | 1 | 1 | 1 | 3 |
| Pos Parent | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 0 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 0 | 1 | 1 | 1 | 4 |
| Pos Child | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 0 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 0 | 1 | 1 | 1 | 4 |
| Neg Parent | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 0 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 0 | 1 | 1 | 1 | 4 |
| Neg Child | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 0 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 0 | 1 | 1 | 1 | 4 |
| Poor Monitoring And Supervision | 1 | 1 | 0 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 0 | 0 | 1 | 1 | 1 | 1 | 0 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 5 |
| Parent Chil Dysfunctional Interraction | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 0 | 1 | 1 | 0 | 1 | 0 | 1 | 1 | 0 | 1 | 1 | 1 | 1 | 0 | 6 |
| Difficult Child | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 0 | 1 | 1 | 1 | 0 | 0 | 0 | 1 | 0 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 0 | 7 |
| Parental Distress | 1 | 1 | 1 | 1 | 1 | 0 | 1 | 0 | 1 | 1 | 1 | 1 | 0 | 1 | 0 | 1 | 0 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 0 | 10 |
| Parenting Stress | 1 | 1 | 1 | 1 | 1 | 0 | 1 | 0 | 0 | 1 | 1 | 0 | 0 | 0 | 0 | 1 | 0 | 1 | 1 | 0 | 1 | 1 | 1 | 1 | 0 | 1 |
| Beck Depression Inventory | 1 | 1 | 1 | 0 | 1 | 1 | 1 | 0 | 0 | 1 | 0 | 0 | 0 | 0 | 0 | 0 | 1 | 1 | 1 | 0 | 1 | 0 | 1 | 1 | 0 | 24 |
| Supporting Positive Behaviour Frequency A | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 0 | 0 | 0 | 25 |
| Setting Limits Frequency A | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 0 | 0 | 0 | 25 |
| Supporting Positive Behaviour Problem A | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 0 | 0 | 0 | 25 |
| Setting Limits Problem A | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 0 | 0 | 0 | 25 |
| Physical Discipline A | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 0 | 0 | 0 | 25 |
| Severe Physical Discipline A | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 0 | 0 | 0 | 25 |
| Psychological Discipline A | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 0 | 0 | 0 | 25 |
| Neglect A | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 0 | 0 | 0 | 25 |
| Poor Monitoring And Supervision A | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 0 | 0 | 0 | 25 |
| Parental Distress A | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 0 | 0 | 0 | 25 |
| Parent Chil Dysfunctional Interraction A | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 0 | 0 | 0 | 25 |
| Difficult Child A | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 0 | 0 | 0 | 25 |
| Social Support A | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 0 | 0 | 0 | 25 |
| ipv Chronicity A | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 0 | 0 | 0 | 25 |
| Positive Parenting Frequency A | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 0 | 0 | 0 | 25 |
| Positive Parenting Problem A | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 0 | 0 | 0 | 25 |
| Parenting Stress A | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 0 | 0 | 0 | 25 |
| Supporting Positive Behaviour Frequency B | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 0 | 0 | 0 | 25 |
| Setting Limits Frequency B | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 0 | 0 | 0 | 25 |

| | | | | | | | | | | | | | | | | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| Supporting Positive Behaviour Problem B | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 0 | 0 | 0 | 25 |
| Setting Limits Problem B | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 0 | 0 | 0 | 25 |
| Physical Discipline B | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 0 | 0 | 0 | 25 |
| Severe Physical Discipline B | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 0 | 0 | 0 | 25 |
| Psychological Discipline B | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 0 | 0 | 0 | 25 |
| Neglect B | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 0 | 0 | 0 | 25 |
| Poor Monitoring And Supervision B | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 0 | 0 | 0 | 25 |
| Parental Distress B | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 0 | 0 | 0 | 25 |
| Parent Chil Dysfunctional Interraction B | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 0 | 0 | 0 | 25 |
| Difficult Child B | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 0 | 0 | 0 | 25 |
| Social Support B | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 0 | 0 | 0 | 25 |
| ipv Chronicity B | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 0 | 0 | 0 | 25 |
| Positive Parenting Frequency B | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 0 | 0 | 0 | 25 |
| Positive Parenting Problem B | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 0 | 0 | 0 | 25 |
| Parenting Stress B | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 0 | 0 | 0 | 25 |
| ECBI intensity A | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 0 | 1 | 0 | 0 | 0 | 26 |
| ECBI problem A | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 0 | 1 | 0 | 0 | 0 | 26 |
| Non Violent Discipline A | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 0 | 1 | 1 | 1 | 1 | 1 | 1 | 0 | 0 | 0 | 26 |
| ECBI intensity B | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 0 | 1 | 0 | 0 | 0 | 26 |
| ECBI problem B | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 0 | 1 | 0 | 0 | 0 | 26 |
| Non Violent Discipline B | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 0 | 1 | 1 | 1 | 1 | 1 | 1 | 0 | 0 | 0 | 26 |
| Beck Depression Inventory A | 1 | 1 | 1 | 1 | 1 | 1 | 0 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 0 | 0 | 0 | 27 |
| Beck Depression Inventory B | 1 | 1 | 1 | 1 | 1 | 1 | 0 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 0 | 0 | 0 | 27 |
| Pos Parent A | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 0 | 0 | 0 | 0 | 0 | 1 | 0 | 0 | 85 |
| Pos Child A | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 0 | 0 | 0 | 0 | 0 | 1 | 0 | 0 | 85 |
| Neg Parent A | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 0 | 0 | 0 | 0 | 0 | 1 | 0 | 0 | 85 |
| Neg Child A | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 0 | 0 | 0 | 0 | 0 | 1 | 0 | 0 | 85 |
| Pos Parent B | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 0 | 0 | 0 | 0 | 0 | 1 | 0 | 0 | 85 |
| Pos Child B | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 0 | 0 | 0 | 0 | 0 | 1 | 0 | 0 | 85 |
| Neg Parent B | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 0 | 0 | 0 | 0 | 0 | 1 | 0 | 0 | 85 |
| Neg Child B | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 0 | 0 | 0 | 0 | 0 | 1 | 0 | 0 | 85 |
| | 0 | 1 | 1 | 1 | 2 | 2 | 2 | 3 | 3 | 4 | 4 | 4 | 4 | 4 | 5 | 5 | 6 | 8 | 10 | 11 | 12 | 13 | 42 | 50 | 56 | 1836 |

The following chapter explores the possible imputation methods that could be applied, to such a dataset. It also covers the chosen imputation model, it’s specification, implementation as well as some imputation model diagnostics.

# Chapter 3

# Imputation

The previous chapter covers a discussion on the extent of missing data and also explores the missing data patterns in the SCFP RCT study. One major finding from the above was that there were low levels of missingness across the time points (especially when looking at an outcome-by-outcome basis). This would normally suggest that there is no need for imputation as the complete cases make up the majority of the data (provided there is no bias introduced by using only complete cases, there would be little loss of precision). In this work, an imputation procedure was nevertheless used.

This chapter begins with a quick summary of the ways in which data can go missing in Section 3.1. This is important as it informs on the choice of imputation method. These methods that are discussed in Section 3.2 and with particular focus on multivariate imputation using chained equations (see Section 3.3) which was implemented for this data. Section 3.4 provides more details regarding the imputation model that was implemented and Sections 3.5 and 3.6 go through convergence checks for the imputation models and checking the 'validity' of the imputed values respectively.

## 3.1 Ways in which data can go missing

To begin, it is quite important to understand the ways in which data can go missing. The literature on missing data talks about three types of missingness as defined by their statistical implications. Here data can be classified as either missing completely at random (MCAR), or just missing at random (MAR) or missing not at random (MNAR). The choice of imputation procedure depends on which of the above is assumed as explained below.

If data is assumed to be MCAR, then by definition, the probability of missing data is unrelated to both the observed and unobserved outcomes. A common example of this in clinical research is when certain questionnaires/ procedures are given to a certain random sub-sample of the original pool of participants because of some cost implications. Since this sub-sample is chosen randomly, it implies that the data on those not in the sample would be seen as missing completely at random. The statistical properties of the sub-sample are by implication representative of the whole sample. This is the default assumption that will apply when one does a complete-case analysis.

Data is said to be MAR if given the observed data, the probability of missingness is independent of the unobserved data. An example is if one suspects that older participants are more likely than younger participants to disclose whether they experience intimate partner violence (IPV). The extent of missing IPV data depends on age (which is observed), but does not depend on IPV (which the unobserved in this case). This is the most commonly used assumption in dealing with missing data and as explained later in the chapter, one can implement prediction models like linear regression in order to impute the missing values.

The definition of MNAR is when the probability of missing data depends on the unobserved data conditional on the observed data. In other words, if the way data is missing on a certain variable in the dataset is itself dependent on the (true) value of the same variable even after controlling for other variables then the missingness is considered to be not at random. By way of example, if one suspects that older people are more inclined to disclose their IPV experiences (like in the previous example) but now one also further suspects that those with high IPV experiences are even more likely to not disclose (probably because of fear), then the missing IPV data is considered to be MNAR. One way of imputing data here is to impute assuming MAR and then tweak the imputations by allocating more (or less) favourable values to the imputations and then using those values for model fitting.

## 3.2 Common methods of handling missing data

There are several ways in which missing data can be dealt with, common methods include casewise deletion, single value imputation, likelihood-based imputation and multiple imputation. Some of the aforementioned methods focus on preserving the precision of estimates by maintaining the sample size, others focus on incorporating the additional uncertainty inherent in missing data, some do both the above.

Casewise deletion (also known as complete case analysis) involves omitting observations in which missing values are detected. This deletion can be done on the full dataset i.e. omitting a case if there is any missing value across the full dataset regardless of some variables are used in the analysis. This is not preferred as it reduces statistical power and induce possible bias.

An alternative involves only deleting a case when it has a missing value on a variable included in the current analysis. This leads to different sample sizes for the different analyses but has the advantage of improving the precision from the aforementioned method. As discussed in Section 3.1 above, complete case analysis is only sensible if the missingness is completely at random, else it introduces bias in the results. Again, it underestimates the variability of the entire model by not taking into account any additional uncertainty brought by the fact that there is some missing data.

Single value imputation is another common technique used for dealing with missing data whereby any missing value is replaced by a single value. Here the general (but very restrictive) assumption is that the values used are the true values that would have been observed had the data been complete. One simple way of performing single imputation is called mean imputation whereby the missing cases are replaced with the mean of the non-missing cases. Another single imputation method involves performing a regression on the complete cases and using that to get fitted/ predicted values for the missing cases on a specific variable. These would then become the imputed values for the missing data. Clearly the main problem with single imputation is that it doesn't account for any uncertainty in the imputed values.

Likelihood based imputation is a form of implicit imputation that is also inherent in generalized linear mixed-effects models (GLMMs). Implicit imputation in this case means that data is not actually imputed but the (maximum likelihood) estimates obtained using the incomplete data are enough to summarize the subject-specific trajectories of any particular response being focussed on. If a subject has a missing value at some specific point, their imputed value (if needed) would be predicted from the subject's trajectory and values for the response in subjects with similar covariate profiles. The main advantage is that likelihood-based models implicitly incorporate the (covariance) structure of the estimating model. Restricted maximum likelihood estimation also introduces comparable variance structures.

Multiple imputation methods account for the uncertainty in the imputed values by not limiting the number of replacements for any missing value to one as in single imputation. Multiple complete datasets are created, each having potentially different imputed values then the analyses are done on the multiple datasets and finally the results are pooled using Rubin's rules. There are two general approaches to deal with imputation of multivariate missing data namely the joint modelling (JM) and the chained equations approach. Both of the aforementioned approaches are based in the idea that each imputation is a random draw from the multivariate probability density of the missing data, they mainly differ in how these draws are made. Joint modelling uses Markov Chain Monte Carlo (MCMC) techniques in order to sample imputations from a pre-specified multivariate distribution of the missing data. It is argued to be more appropriate when one can correctly specify the aforementioned distribution. Imputation using the chained equations approach is a good alternative especially when no suitable multivariate distributions can be found for the missing data. Here, imputations are made on a variable-by-variable basis by making use of the conditional densities for each incomplete variable. Unlike JM which uses MCMC techniques, imputation by chained equations makes use of the Gibbs Sampling technique. The latter approach is also known as fully conditional specification (FCS) and in most literature it is referred to as multivariate imputation by chained equations (MICE). As prescribed in the **analysis plan**, imputation was done using the chained equations approach and was implemented using the **R** package **mice** (Van Buuren and Groothuis-Oudshoorn, 2011).

## 3.3 MICE: Details

As discussed before, **mice** uses the Gibbs Sampling method to impute the missing data items. Suppose that $\mathbf{Y}_{N\times p}$ is a partially observed random sample (of size $N$) from the multivariate distribution of the $p$ variables making up the columns of $\mathbf{Y}$, i.e. $P(\mathbf{Y}|\theta)$ where $\theta$ is a set of parameters governing this distribution. The algorithm described below makes use of the conditional distributions to sample imputations and update the dataset iteratively, the idea of the Gibbs Sampling procedure is that eventually one will end up sampling from the above multivariate distribution. The algorithm samples imputations iteratively from the aforementioned distribution in the following manner at the $t^{th}$ iteration:

(i) Obtain a posterior distribution for $\theta_1$ and sample a random perturbation $\theta_1^{*(t)}$ from it. Here, $\theta_1$ the vector of parameters governing the distribution of the variable to be imputed first (conditional on all other variables included in the model), without loss of generality, assume that the $\mathbf{Y}$ is already ordered according to the order of variable imputation. If one uses a regression based imputation method, then $\theta_1$ consists of the regression parameters and the root mean square error for the regression model. In such a case, the posterior distributions for the parameters in $\theta_1$ are the multivariate normal distribution and the chi-square distribution (for the betas and sigma respectively). i.e. draw $\theta_1^{*(t)}$ from $P(\theta_1|\mathbf{Y}_1^{Obs}, \mathbf{Y}_2^{(t-1)}, \ldots, \mathbf{Y}_p^{(t-1)})$

where $\mathbf{Y}_j^{(t)} = (\mathbf{Y}_j^{obs}, \mathbf{Y}_j^{*(t)})$ is the imputed variable at iteration $t$.

(ii) Obtain the imputation for missing items in $\mathbf{Y}_1$ by drawing $\mathbf{Y}_1^{*(t)}$ from
$P(\mathbf{Y}_1|\mathbf{Y}_1^{Obs}, \mathbf{Y}_2^{(t-1)}, \ldots, \mathbf{Y}_p^{(t-1)}, \theta_1^{*(t)})$.

(iii) Perform the above two steps for imputing the rest of the incomplete variables. i.e. to impute $\mathbf{Y}_2$, draw $\theta_2^{*(t)}$ from $P(\theta_2|\mathbf{Y}_1^{(t)}, \mathbf{Y}_2^{Obs}, \mathbf{Y}_3^{(t-1)}, \ldots, \mathbf{Y}_p^{(t-1)})$ and use this to draw $\mathbf{Y}_2^{*(t)}$ from $P(\mathbf{Y}_2|\mathbf{Y}_1^{(t)}, \mathbf{Y}_2^{Obs}, \mathbf{Y}_3^{(t-1)}, \ldots, \mathbf{Y}_p^{(t-1)}, \theta_2^{*(t)})$ and so on.

To impute $\mathbf{Y}_p$, draw $\theta_p^{*(t)}$ from $P(\theta_p|\mathbf{Y}_1^{(t)}, \mathbf{Y}_2^{(t)}, \ldots, \mathbf{Y}_{p-1}^{(t)}, \mathbf{Y}_p^{Obs})$ and use this to draw $\mathbf{Y}_p^{*(t)}$ from $P(\mathbf{Y}_p|\mathbf{Y}_1^{(t)}, \mathbf{Y}_2^{(t)}, \ldots, \mathbf{Y}_{p-1}^{(t)}, \mathbf{Y}_p^{Obs}, \theta_p^{*(t)})$.

Notice that the order of imputation is important as the later variables are imputed based on imputed values of the ones that come earlier in the algorithm.

In short, each iteration in the above algorithm involves imputing the dataset on a variable-by-variable basis whereby for each variable, one needs to (a) generate a random draw from the posterior distribution of $\theta_i$ (for the $i^{th}$ variable) and then (b) use the random draw from above as the parameters of the conditional distribution of $\mathbf{Y}_i$ from which one would draw the random imputation. Shah *et al.* (2014, pp. 764) describe how the default settings in many **mice** implementations achieve points (a) and (b) above in the following three-step algorithm:

(1) Fit some form of regression model on the complete cases to get the estimate and distribution of $\theta_i$. In the case of a linear regression, $\theta$ consists of $\beta$ and $\sigma$ (the regression coefficients and the root mean square error). In such a case, the posterior distribution of $\theta_i$ will be multivariate normal for the $\beta$ and univariate chi square for $\sigma^2$.

(2) Use the above to sample the random perturbations of both $\beta$ and $\sigma$, say $\beta^*$ and $\sigma^*$ from their respective distributions.

(3) Use the random perturbations in (2) as parameters in the conditional distribution of the variable to be imputed. The imputation will be a random generated value from this distribution. For a linear regression model, this would also be coming from a multivariate normal.

Since this is implemented in a multiple imputation framework, this process would have to be implemented a number of times in parallel with each chain being ensured to be independent from the other. These different chains will then form the multiple complete datasets discussed earlier in Section 3.2. The above three step procedure would differ with different data types. Common methods include **logistic regression** for binary data, **predictive mean matching** or linear regression (as described above) for numeric/ continuous data and multinomial logit models for categorical data (more than 2 categories). Van Buuren and Groothuis-Oudshoorn (2011, pp. 16) provides an extensive list of these methods and the cases in which they would apply and the examples are implemented using the **mice** package in **R**. In addition to the aforementioned methods which are all parametric or semi-parametric, one can also choose the **random forest** imputation method. This is a non-parametric method that makes use of a machine learning algorithm called random forest instead of ordinary least squares regression. Shah *et al.* (2014) argue through the use of simulation studies that random forest imputation would produce more efficient estimates when data was artificially made to be missing at random and less biased than the parametric methods if there are non-linearities in the "true" imputation model.

In this analysis, only predictive mean matching, logistic regression and random forest imputation methods were explored. Initially, only logistic regression and predictive mean matching were used but they were both replaced with the random forest method. This was because of two reasons: (1) it was discovered that predictive mean matching was not converging quickly enough (see Section 3.5) and (2) that it was much easier to employ one method on all the variables (even though logistic regression was working well). Subsections 3.3.1, 3.3.2 and 3.3.3 briefly explain how these methods work as discussed in White *et al.* (2011) and Shah *et al.* (2014).

### 3.3.1 Logistic Regression Imputation

Logistic regression is a parametric imputation method appropriate for binary variables. This is very similar to the general method using linear regression discussed above. The three-step algorithm for imputing the $j^{th}$ variable at the $t^{th}$ iteration of the Gibbs sampler is as follows:

(1) If Y is a binary response variable with missing values, use a logistic regression to model the log of the odds (of the missing data item being one of the two classes) as a linear combination of the chosen predictors i.e. $\log\left(\frac{P(Y=1)}{1-P(Y=1)}\right) = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \cdots + \beta_p X_p$. The regression coefficients ($\beta$) are therefore on the logit scale. This should yield $\hat{\beta}$, the regression coefficients and $\mathbf{V}$, the estimated variance-covariance matrix.

(2) Obtain the random perturbation $\beta^*$ by drawing from a multivariate normal ($\hat{\beta}$,$\mathbf{V}$) which approximates the posterior distribution of $\beta$.

(3) For each participant/observation $i$ calculate $p_i^* = [1 + \exp(-x_i\beta^*)]^{-1}$ and impute $y_i^*$ as follows:

$$y_i^* = \begin{cases} 1 & \text{if } \ u_i < p_i^* \\ 0 & \text{otherwise} \end{cases}$$

where $u_i$ is a randomly generated number from the $U(0,1)$ distribution, $x_i$ is the vector of (observed and imputed) predictors and $y_i^*$ is the imputed value for the $i^{th}$ participant with incomplete data on the specific variable.

The above is illustrated on a variable-by-variable and case-by-case basis, of course one can combine all the cases for each variable and impute whole vectors of missing values on the variable.

### 3.3.2 Predictive Mean Matching Imputation

Predictive mean matching is a semi-parametric imputation method that ultimately samples the imputed values from the values on the complete cases. The first two steps of the algorithm are just like in the general case described on page 35, the only differences come in the remaining steps. The algorithm is as follows:

(1) Fit a linear regression model to the complete cases with the incomplete variable in question as the dependent and the predictors being as outlined in Section 3.4. This should yield estimates and distributions for $\hat{\beta}$ and $\hat{\sigma}$.

(2) Obtain random perturbations of $\hat{\beta}$ and $\hat{\sigma}$ (say $\beta^*$ and $\sigma^*$) by drawing from their posterior distributions i.e. MVN for $\beta$ and Chi-square for $\sigma$ with linear regression.

(3) Use the above perturbations ( $\beta^*$ and $\sigma^*$) to get predicted values of the response both on the observed and unobserved data.

(4) The method then chooses the k (usually 5 or more) nearest complete data points based on the predictions, i.e. choose the k complete data points whose predicted values are nearest to the predicted value of the incomplete data.

(5) The imputed value is the observed (not predicted) value of a randomly selected candidate from the k above.

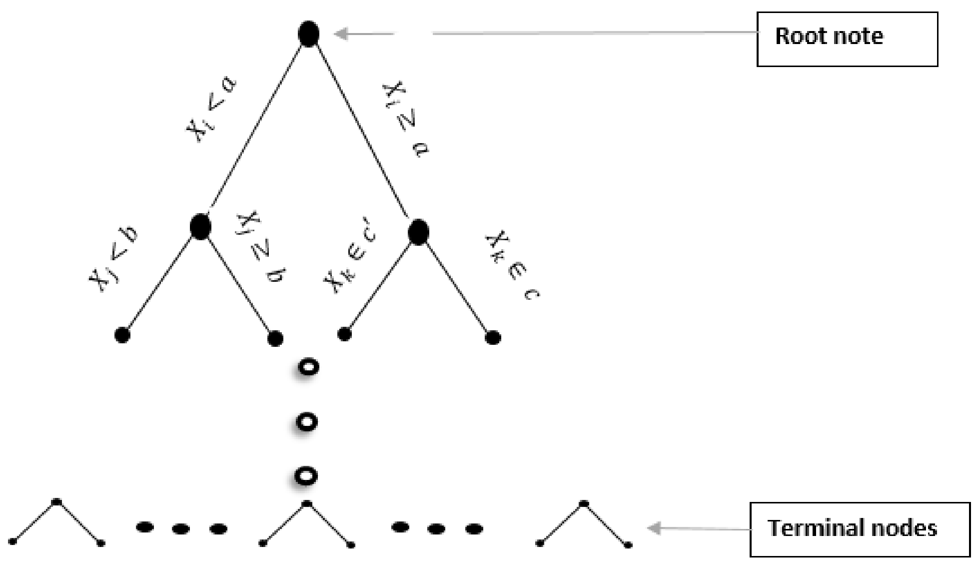
The resulting imputed values share the same properties as the observed data. The method is appropriate if the completely observed data has enough variability to cover the range of outcomes for the variable.

### 3.3.3 Random Forest Imputation

Random forest imputation (Liaw and Wiener, 2001) is a non-parametric method adapted from tree based machine learning techniques. It involves fitting multiple trees on different bootstrap samples of the complete cases in the data set. The fitted values are then obtained by taking the average (in case of a numerical variable) or mode (for categorical data) of all the individual trees' fitted values over all the trees in the random forest. The aforementioned trees are called regression trees if the variable to be imputed is numeric or classification trees if the variable is categorical.

The process of fitting a tree involves the use of a 'greedy algorithm' called recursive binary splitting. From the root node (i.e. top of the tree), data is split into two groups based on a variable and cut-off value which minimizes some sort of predetermined error function. Some commonly used error functions include the residual sum of squares (numeric data) and the Gini-index (categorical data). The second step of fitting the tree involves going to each of the two separate nodes that came out of the root node and again splitting these subsets into two smaller sets (per node) based on the variable and cut-off point (of the variable) that gives the **greatest reduction** in the error function. It is possible that one can have different variables and cut-off values for the different tree nodes at any level of the tree. This process is carried on until some stopping criteria is reached or until each observation is on it's own node (i.e. the biggest possible tree has N terminal nodes for a sample of size N).

Figure 3.1: Example of a tree



The predicted/ fitted value at each node is either the mean or mode depending on the data type as discussed above. Figure 3.1 illustrates an example of a tree constructed using the recursive binary splitting algorithm with the first two stages of the process also indicated. The continuous variable $X_i$ was used to split the root node and the cut-off value used was $a$ in such a way that those with values less than $a$ are split from those with values greater than $a$ on the variable $X_i$ (this achieves tree depth 1). These two nodes are then further split (one by one) using the recursive binary splitting algorithm, here one node is split using the continuous variables $X_j$ and the other is split using the categorical variable $X_k$. At this stage (depth 2), the sample now has been split into the following four classes:

1. $\{X_i < a \cap X_j < b\}$
2. $\{X_i < a \cap X_j \geq b\}$
3. $\{X_i \geq a \cap X_k \in c\}$
4. $\{X_i \geq a \cap X_k \notin c\}$

If one only had the root node then the predicted value will be the mean or mode (depending on the data type) of the entire sample. After any split, the predicted values will be based on the sub-samples attained on that stage. Trees are non-linear classifiers in general which gives them the advantage over regression based methods in that they can accommodate any patterns in the data but this also makes them prone to the problem of over-fitting if stopping criteria is left open. Over-fitting is also associated with the model under-performing when predicting out of sample. Another problem with fitting trees is that one has no way to estimate the variation of the predictions.

Random forests address the shortcomings of trees by fitting multiple small trees (usually of depth 2), each one on a different bootstrap samples and also making sure that each tree is fitted using a ransom sub-sample of the set of predictors available. Fitting the small trees is the random forests' mechanism of avoiding over-fitting. The bootstrapping will help trees to handle sampling variation as the multiple bootstraps are meant to be representative of other potential samples from the population. It also helps with the computation of the out-of-bag error which can be used as a measure of the variance for the predictions. Using random sub-samples of the predictor set to generate the different trees also helps achieve some variation in the tree structure itself. In some datasets, there may be some dominant variables that are always chosen first through the greedy splitting algorithm. If variables are not randomly selected this way, there is a chance that all the trees in the forest would be similar and therefore having the similar predictions which would underestimate the variance.

In the case of a numeric variable, the imputed value for the missing data item using the random forest method would be the mean prediction over all the trees. For categorical variables, the imputation would be the mode of the predicted classes over all the fitted trees for the specific study participant with incomplete data. The out-of-bag error can be used as a proxy for the variability of the random forest predictions. This is calculated in the following way:

- Since each tree is grown on a different bootstrap sample, there is a sub-sample of the complete cases that is always left out. The observations in the aforementioned sub-samples are called out-of-bag samples. Of course since the process of bootstrapping is random by construction, these out-of-bag items are random and cannot be expected to be the same for all the trees.

- After each tree has been fitted, predictions are then made for out-of-bag samples.

- These out-of-bag predictions are then aggregated over all the trees (i.e. the **mean** is taken for numeric data and **mode** for categorical data). This yields the aggregated predicted values for each time the observation was out of bag. Ideally for a large number of bootstraps, one would expect each complete-case observation to be left out of the bag at least once.

- The out of bag error is then computed by using all the complete cases that have an out-of-bag prediction. This can be the residual sum of squares for numeric data or some classification error for categorical data.

The algorithm for imputing the $j^{th}$ incomplete variable is as follows:

(1) Run a random forest on bootstrap samples of the complete cases to get predictions of the missing values (average of every tree's prediction of the unobserved item).

(2) Use the **out of bag error** as a proxy of the variability of the prediction above.

(3) If response is numeric, use the above as parameters of the new multivariate (normal) distribution from which you'll randomly draw the imputation for the missing data item.

(4) If response is binary, just randomly choose one of the trees in the forest and take its prediction as the imputation. This is similar from randomly picking a 1 with probability being the average prediction over all trees, 0 otherwise. This can be extended to multi-categorical case whereby one would pick the imputed value with probability being the proportion of trees that predicted the class in hand.

## 3.4 Derivation of the imputation model

The imputation model implemented in this analysis closely follows that which is outlined in the analysis plan (see Appendix B pp. 6-9). This plan assumes that data is missing at random (MAR) and as such, the models discussed in this section are centred around finding the best methods and predictors in order to model the missingness. Basically, this plan specifies the rules to selecting predictors for the imputation of each incomplete variable in the data set. It also states which imputation method would be used for each variable, i.e. predictive mean matching for Likert type data and logistic regression for binary data (these two data types comprise a majority of the variables). As discussed earlier in the chapter, the two aforementioned methods prescribed in the analysis plan were later replaced with random forest imputation (which works with any data type).

Imputation was done at the item level. The imputation model for a given item included the following predictors:

1. Some basic variables from the assumed measurement model:
   - Child sex (male or female),
   - Child age group (2-5 years old or 6-9 years old),
   - Study arm (control or intervention),
   - Programme group (0 for the control group then 1-11 in the intervention arm),
   - Programme wave (wave 1: Khayelitsha or wave 2: Nyanga);
2. Other items making up the score at the current and all previous study visits; and
3. Composite scores in the same group of primary or secondary outcomes at the current and all previous visits.

By way of example, consider the item "*Involve your child in household chores behaviour*" that forms part of eight items summed to make up **Supporting Positive Behaviour Problem (subscore)**. This subscore is also summed up with **Setting Limits Problem (subscore)** will make up the **Positive Parenting Problem score**. The three aforementioned scores and together with **Supporting Positive Behaviour Frequency (subscore)**, **Setting Limits Frequency (subscore)** and **Positive Parenting Frequency score** form a group of primary outcomes called **Positive Parenting**. The predictors in the imputation model would include:

1. All the 'basic' variables listed above (child age & sex, study arm, programme group and wave);
2. The other seven items that have to be summed up with the current item to come up with **Supporting Positive Behaviour Problem (subscore)** at current and past time points. The item "*Involve your child in household chores behaviour*" at all past time points is also included; and
3. The sum scores **Supporting Positive Behaviour Frequency (subscore)**, **Setting Limits Frequency (subscore)**, and **Setting Limits Problem (subscore)** at the current and previous time points. The inclusion of **Supporting Positive Behaviour Problem (subscore)** at any previous time points will be redundant and will also introduce collinearity since it can be recreated using the items in (2) above. Including

**Positive Parenting Frequency score** or **Positive Parenting Problem score** will have the same effects since these are derived by adding up the subscores mentioned above.

Passive imputation was used for all the summed scores. Due to computer memory problems, it was decided that the imputation would be implemented in batches corresponding to each time point. The impact of this approach was that only observed measurements (and not imputed values) from previous visits were used in the imputation at subsequent visits.

## 3.5 Checking the convergence of the imputation algorithm

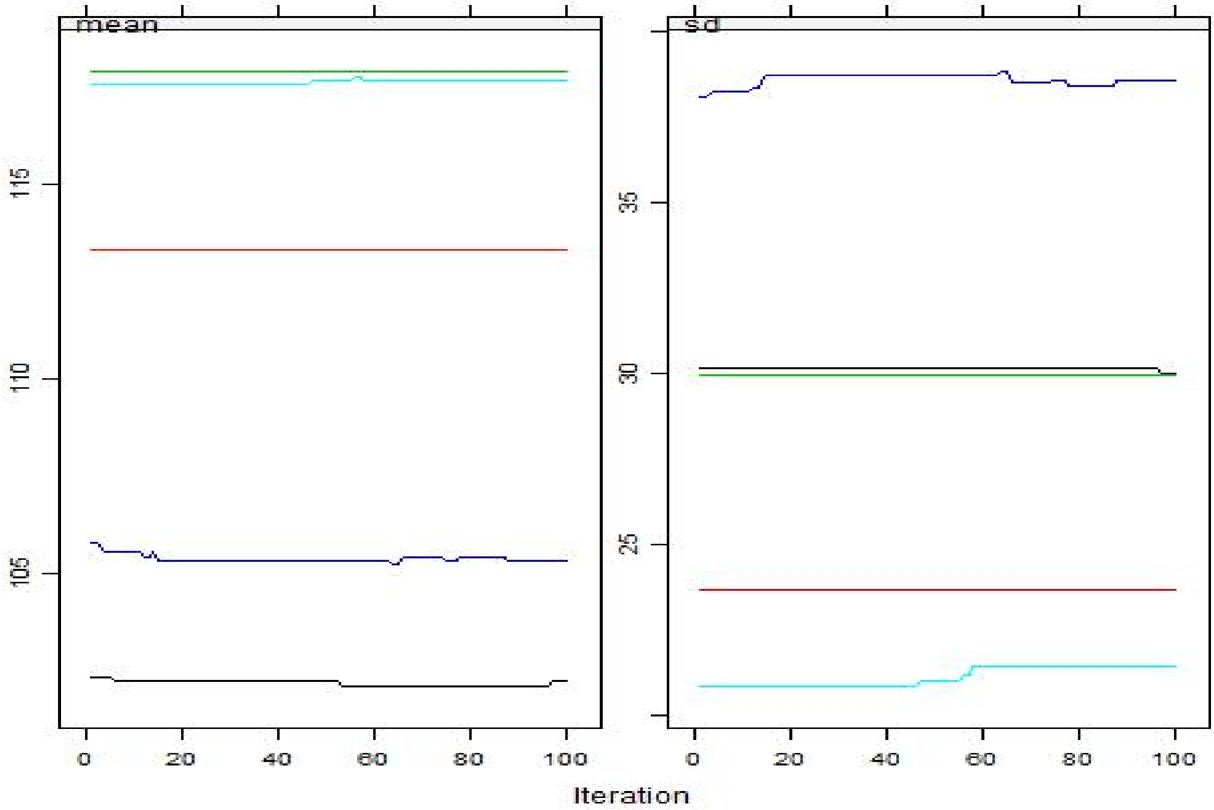
To recap, the idea behind **mice** is that imputations can be obtained sampling from the multivariate distribution of the missing data using the Gibbs Sampler. The Gibbs Sampler basically iteratively samples from conditional densities (instead of the full joint distribution which might be difficult to specify fully) as described in Section 3.3 above. The idea here is that in the long run (or at convergence), one would be sampling imputations from the true multivariate distribution of the missing data. As such, it is an important step to always check that the imputation algorithm has converged and hence the imputed values are now coming from the true multivariate distribution as required.

According to Van Buuren and Groothuis-Oudshoorn (2011), one good way to check for convergence is to assess if the chains are mixing well in such a way that one can't tell the difference between observations from the different chains i.e. at convergence, "should not notice any trends and also the between-chain variation should not exceed the within-chain variation". The plots below show the chain mean and standard deviations at each of the 100 iterations that the algorithm was run. These plots serve to illustrate two main points: (1) predictive mean matching would not converge and in some cases, the chains didn't move away from their starting values, (2) random forest imputation and logistic regression were very quick to converge and ultimately the former was chosen since it also could be applied on any data type.

Figures 3.2 and 3.3 show two different scenarios that summarizes the convergence of the chosen imputation methods. Notice that these are plots of the composite scores which would have been passively imputed, the methods displayed here are what was used to impute the individual variables/ items that would then be summed to get the composite score. In Figure 3.2 (a), clearly, the predictive mean matching method hasn't led to convergence within the first 100 iterations. In fact, it seems that the different chains are stuck at their respective initial values and they're clearly not mixing. In contrast, Figure 3.2 (b) shows healthy convergence for the random forest imputation method. Both methods seem to be converging well in figures 3.3 (a) and (b). The rest of the iteration plots are in Appendix C and they all show how the random forest method and logistic regression always converged within the first 100 iterations whilst most of the times predictive mean matching did not. The labels for the y-axis appear in the top left hand of each plot. It is true that the logistic regression imputation converge at a different value from the random forest imputation. Of prime importance is the degree of overlap of the different line graphs. These plots clearly illustrate that the random forest imputation was the more acceptable method. The imputed value in all these cases would be the a random pick from the chain once it has converged, in this case one could set it to be the $100^{th}$ imputed value on each chain since convergence would have been reached using the random forest method.

Figure 3.2: **Checking for convergence: ECBI Intensity at post-test**

(a) **Using predictive mean matching**




(b) **Using random forest imputation**

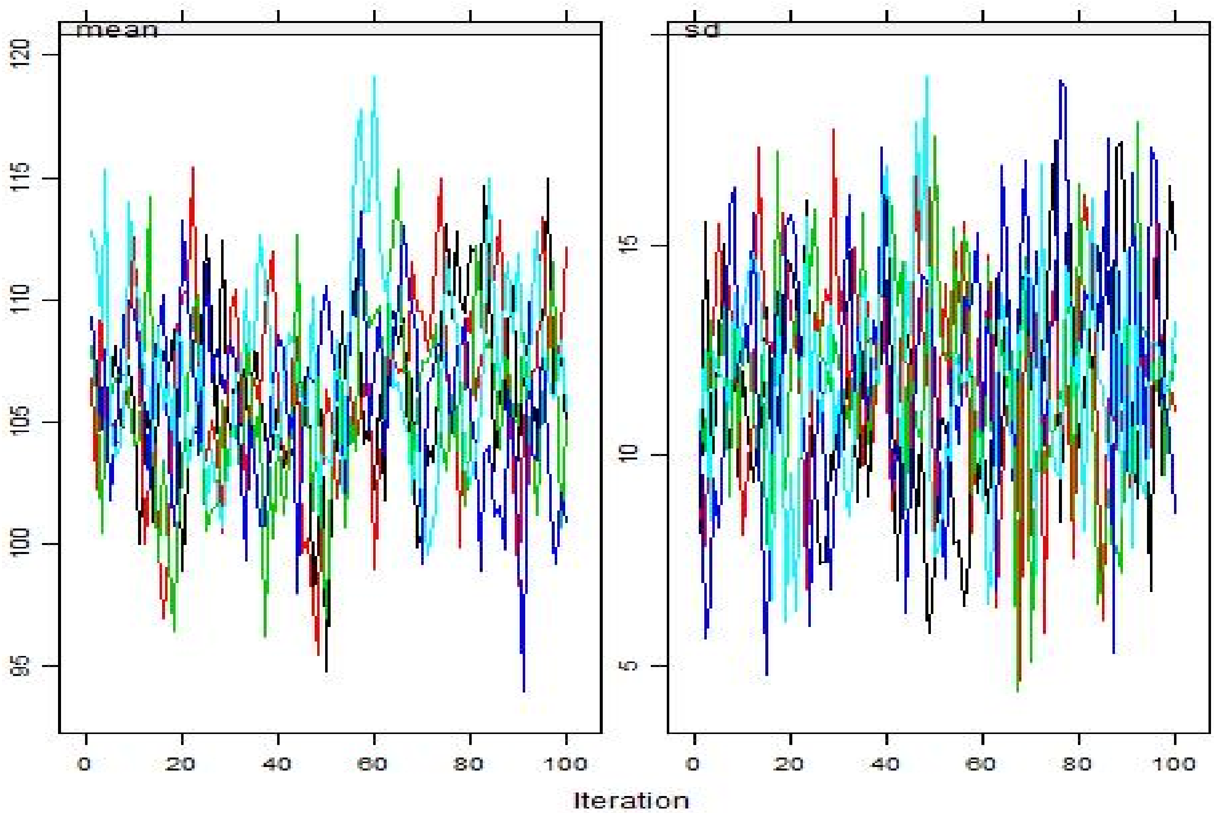
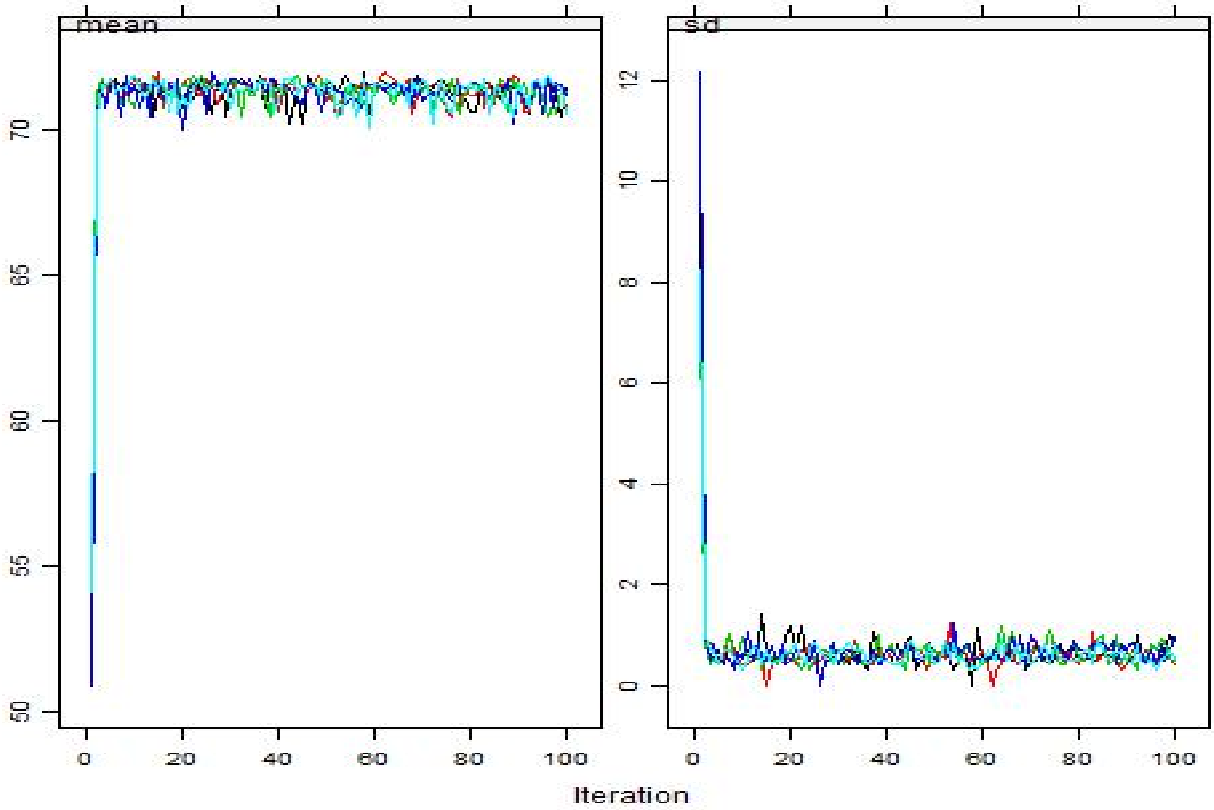
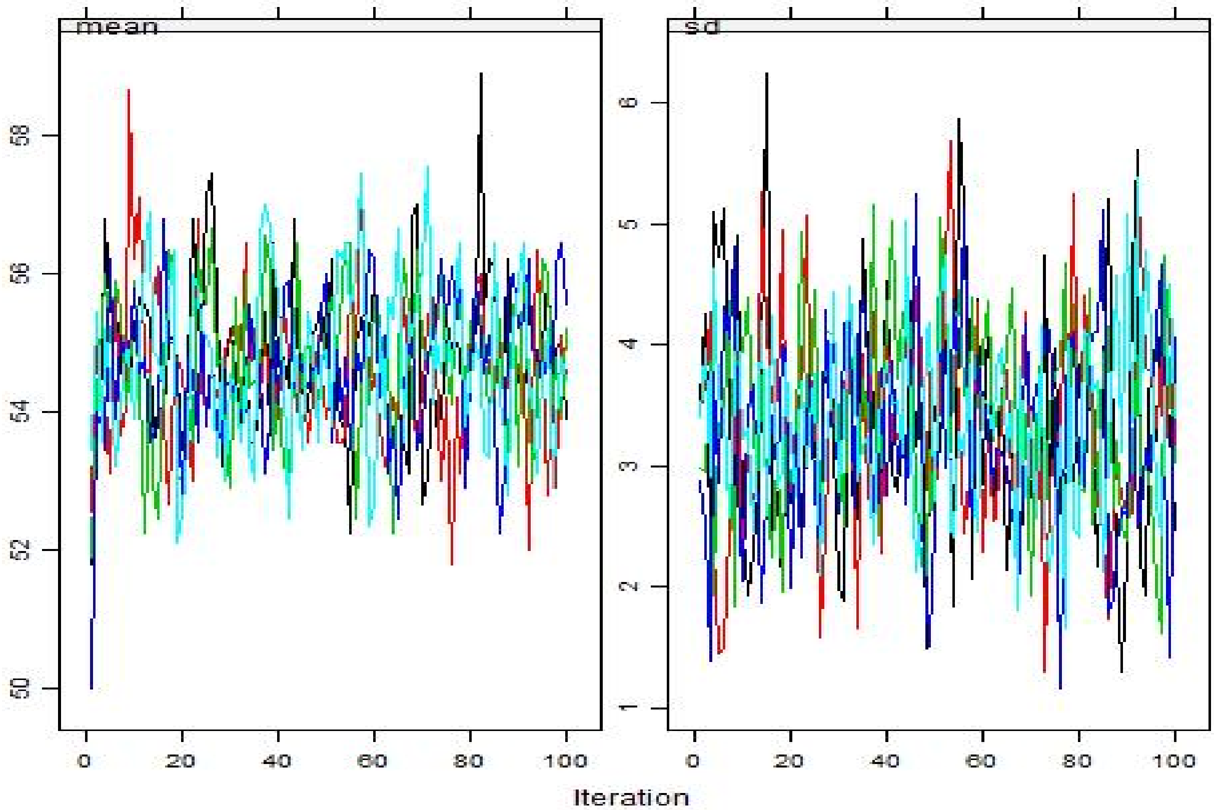


Figure 3.3: **Checking for convergence: ECBI Problem at post-test**

(a) **Using logistic regression**




(b) **Using random forest imputation**



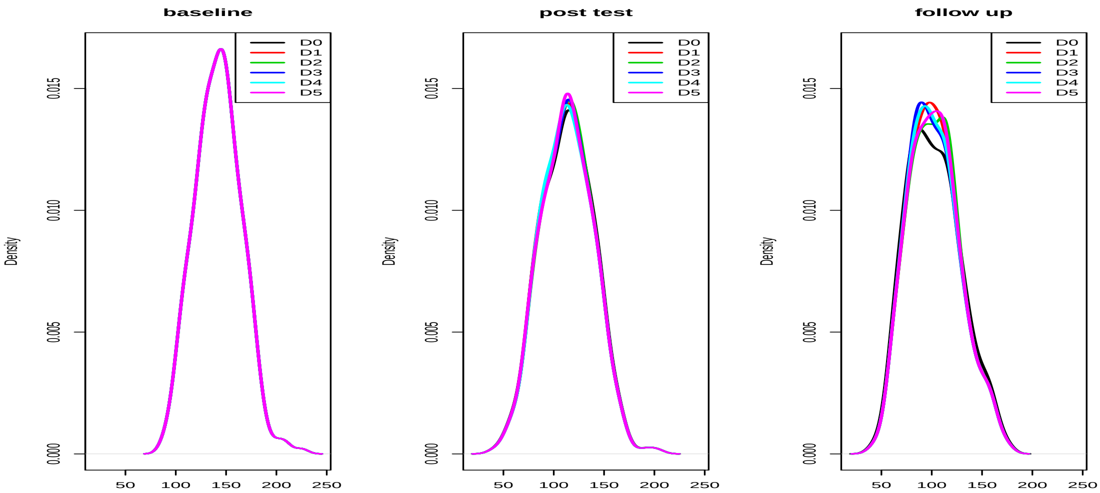
## 3.6 Checking the imputations

Van Buuren and Groothuis-Oudshoorn (2011, pp. 42) describe a good imputation as "a value that could have been observed had it not been missing". In other words, it is ideal that the imputed value looks like it still comes from the same distribution as the observed values. The plots in Figures 3.4 (a) - 3.4 (j) show the kernel densities for the unimputed dataset as well as the five 'complete' datasets based on random forest imputation. Notice that the plots include a few variables that were complete at baseline (e.g. ECBI Intensiy, see the rest in Table 2.8), for all these outcomes, the kernel densities would obviously be the same over the six datasets. As the missingness increases, one would expect more between chain variability and thus more clear differences between the different kernel densities. This is more visible at the one year follow-up visit where a larger proportion of data is missing. The observational outcomes also have the most missing values at this time point and would therefore be expected to have more visible differences. The plots in Figures 3.4 (a) - 3.4 (j) below do not indicate any imputed distributions chains that were significantly different from the others (including the unimputed set). As such, we can conclude that it is reasonable to assume that the imputations come from a similar distribution as the observed/ reported data.

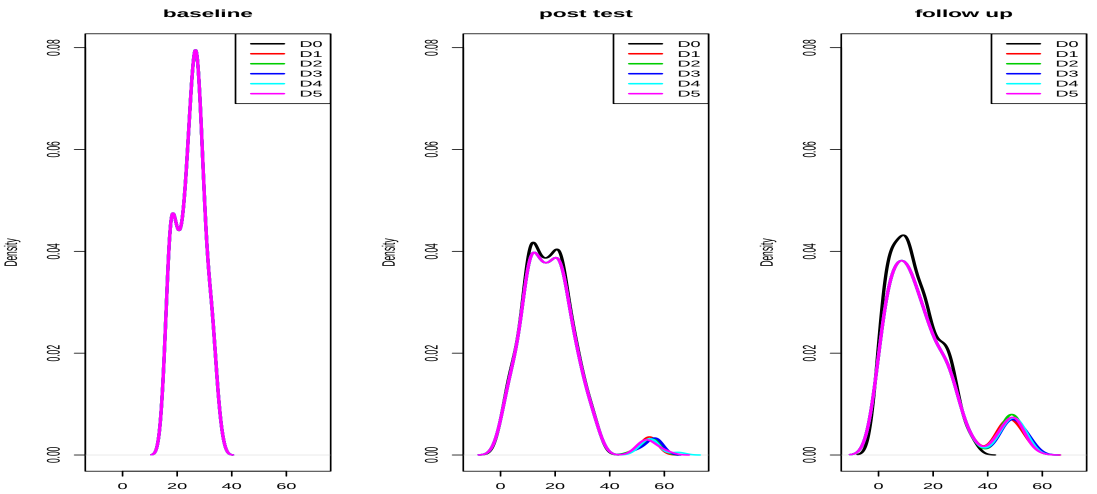
—

Figure 3.4 (a): **Kernel Density Plots for Imputed and Unimputed Datasets**

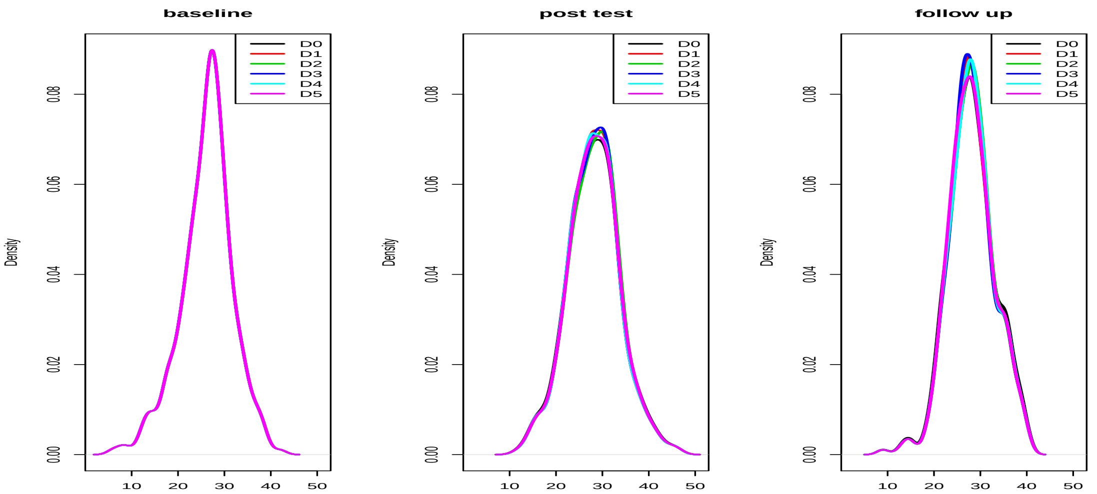
ECBI Intensity




ECBI Problem

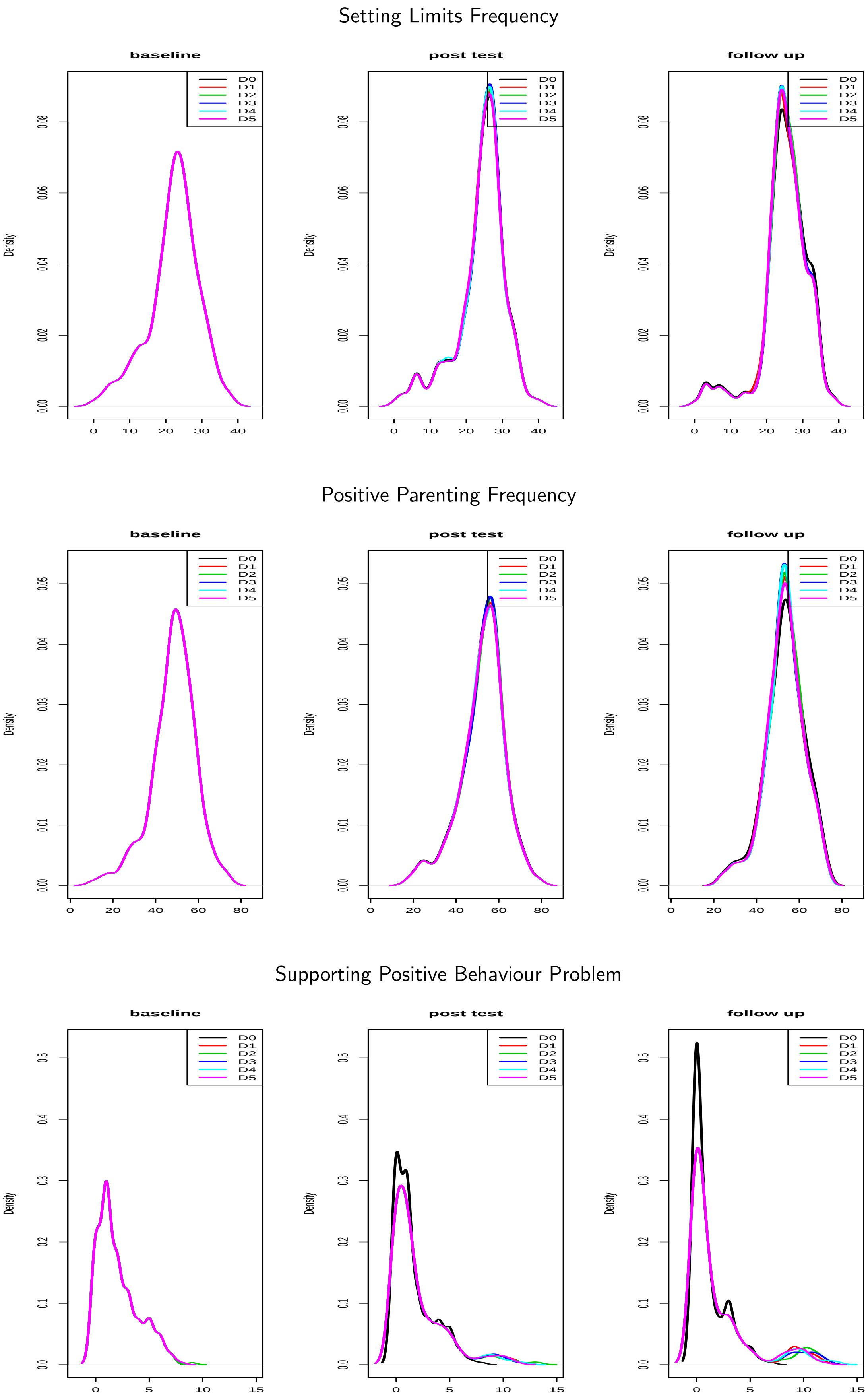



Supporting Positive Behaviour Frequency



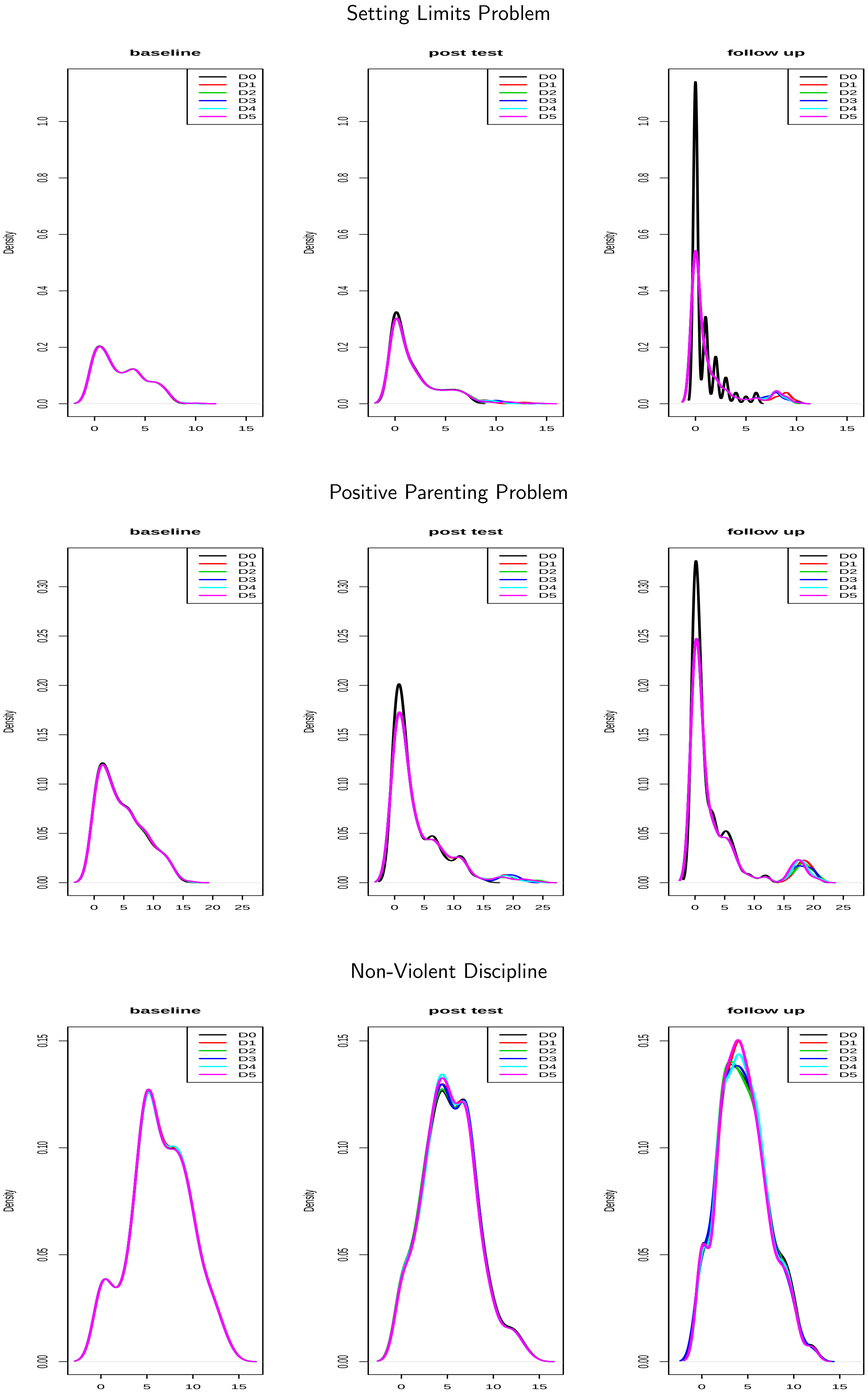

*D0 is unimputed and D1 - D5 are the 5 imputed datasets

Figure 3.4 (b): **Kernel Density Plots for Imputed and Unimputed Datasets**




*D0 is unimputed and D1 - D5 are the 5 imputed datasets

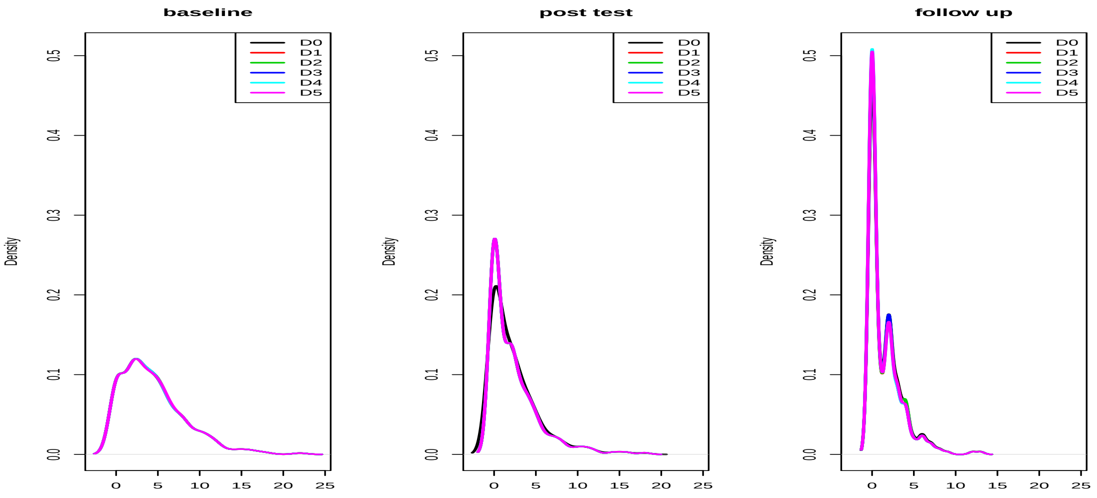
Figure 3.4 (c): **Kernel Density Plots for Imputed and Unimputed Datasets**



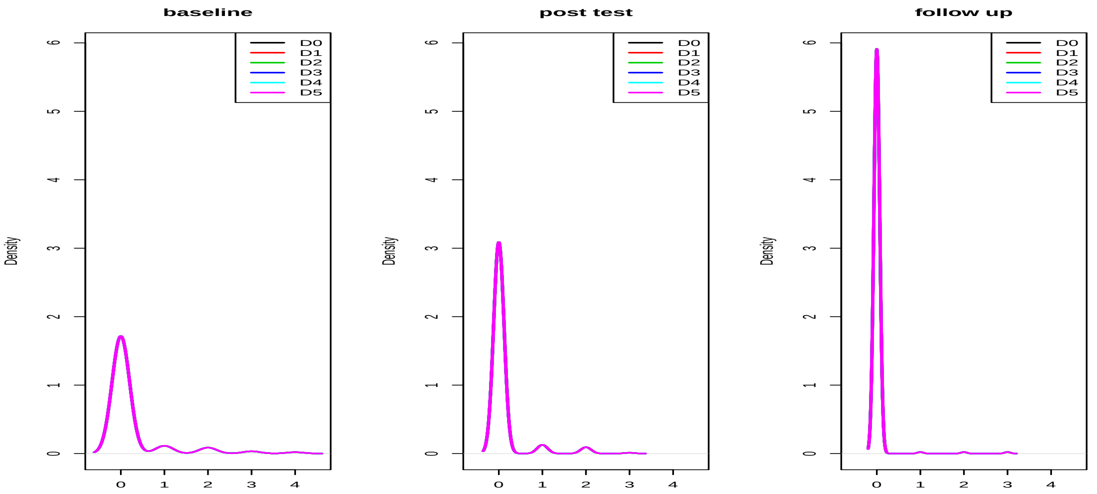

*D0 is unimputed and D1 - D5 are the 5 imputed datasets

Figure 3.4 (d): **Kernel Density Plots for Imputed and Unimputed Datasets**

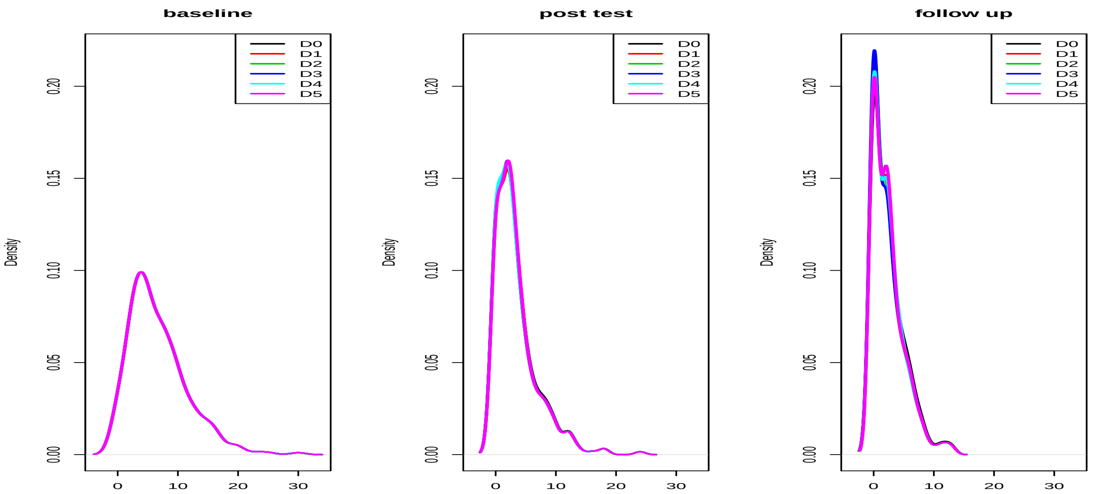
Physical Discipline




Severe Physical Discipline

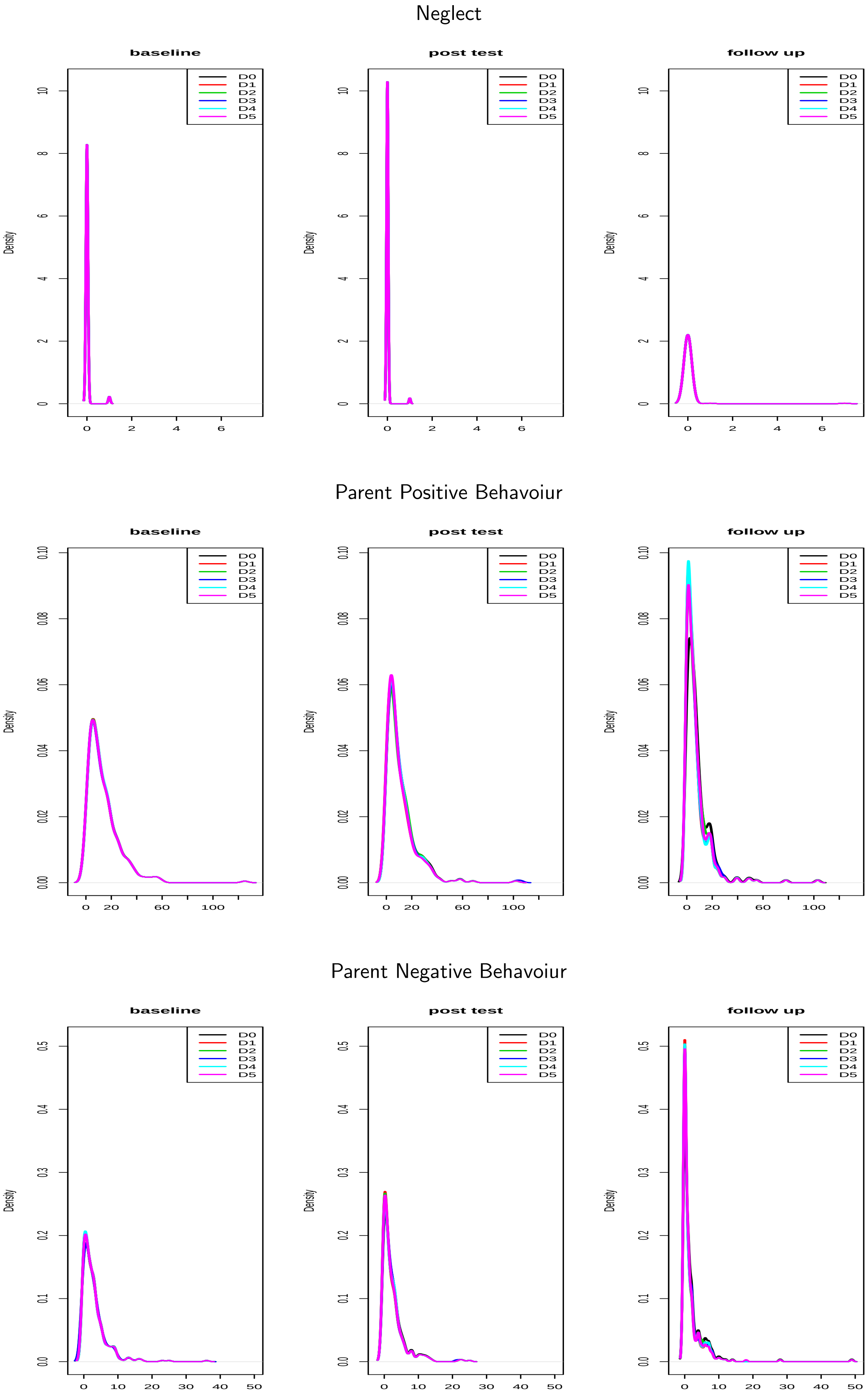



Psychological Discipline




*D0 is unimputed and D1 - D5 are the 5 imputed datasets

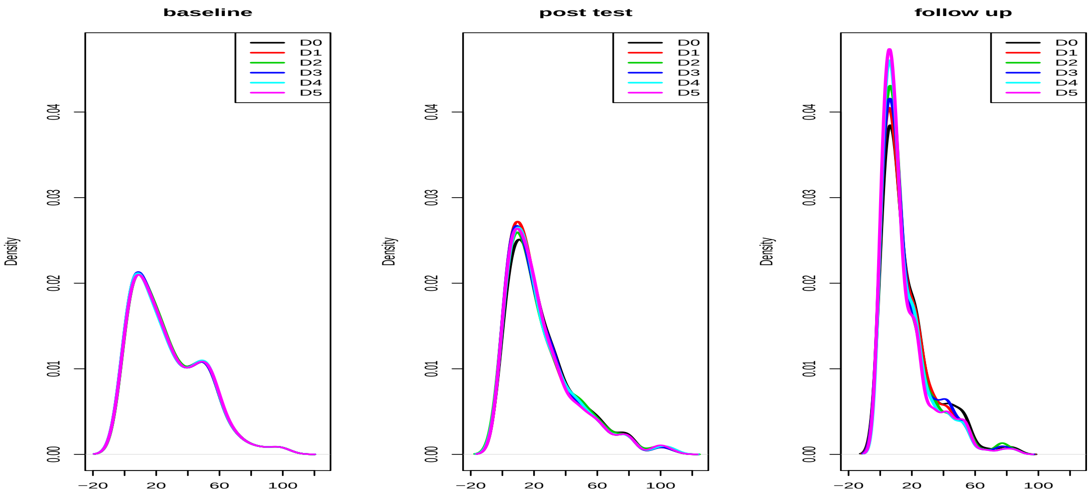
Figure 3.4 (e): **Kernel Density Plots for Imputed and Unimputed Datasets**



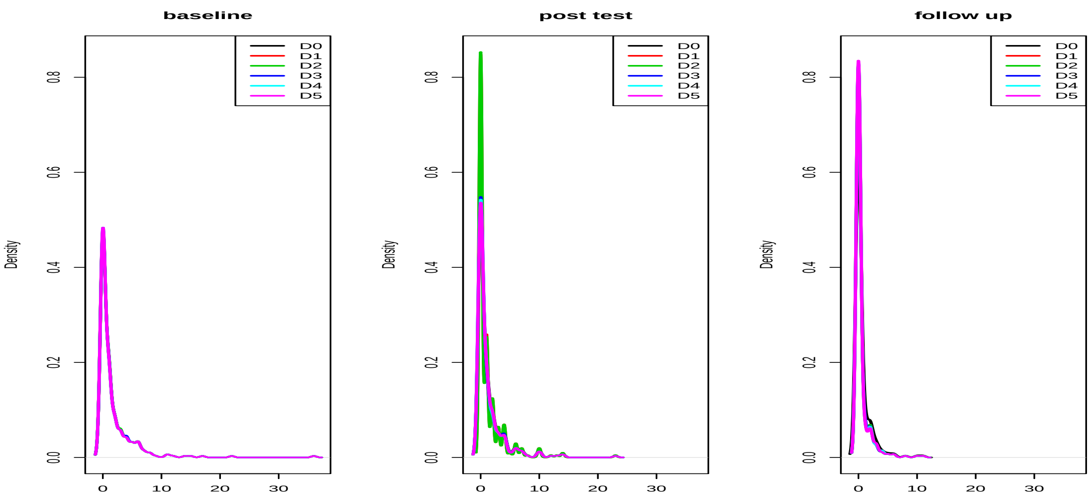

*D0 is unimputed and D1 - D5 are the 5 imputed datasets

Figure 3.4 (h): **Kernel Density Plots for Imputed and Unimputed Datasets**

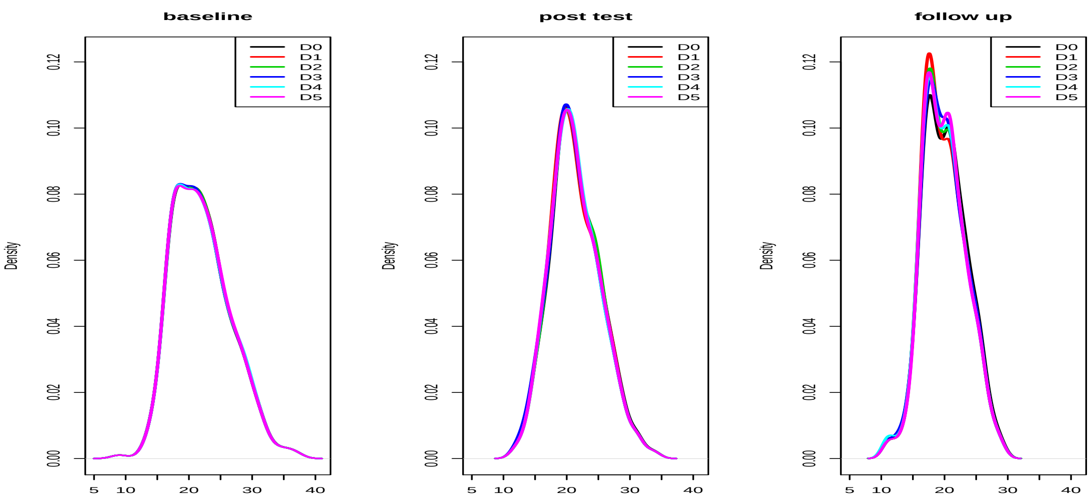
Child Positive Behavoiur




Child Negative Behavoiur

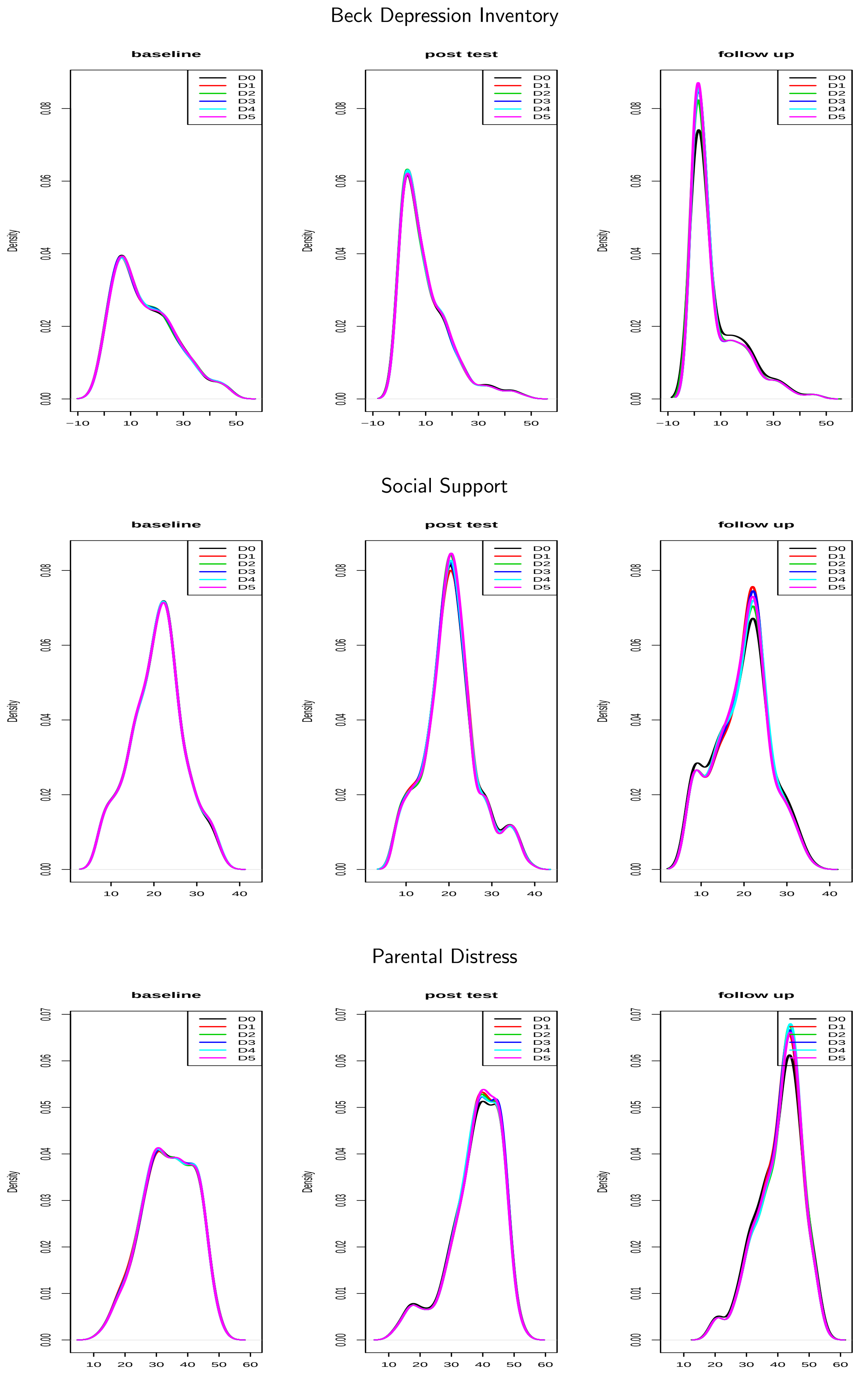



Poor Monitoring and Supervision



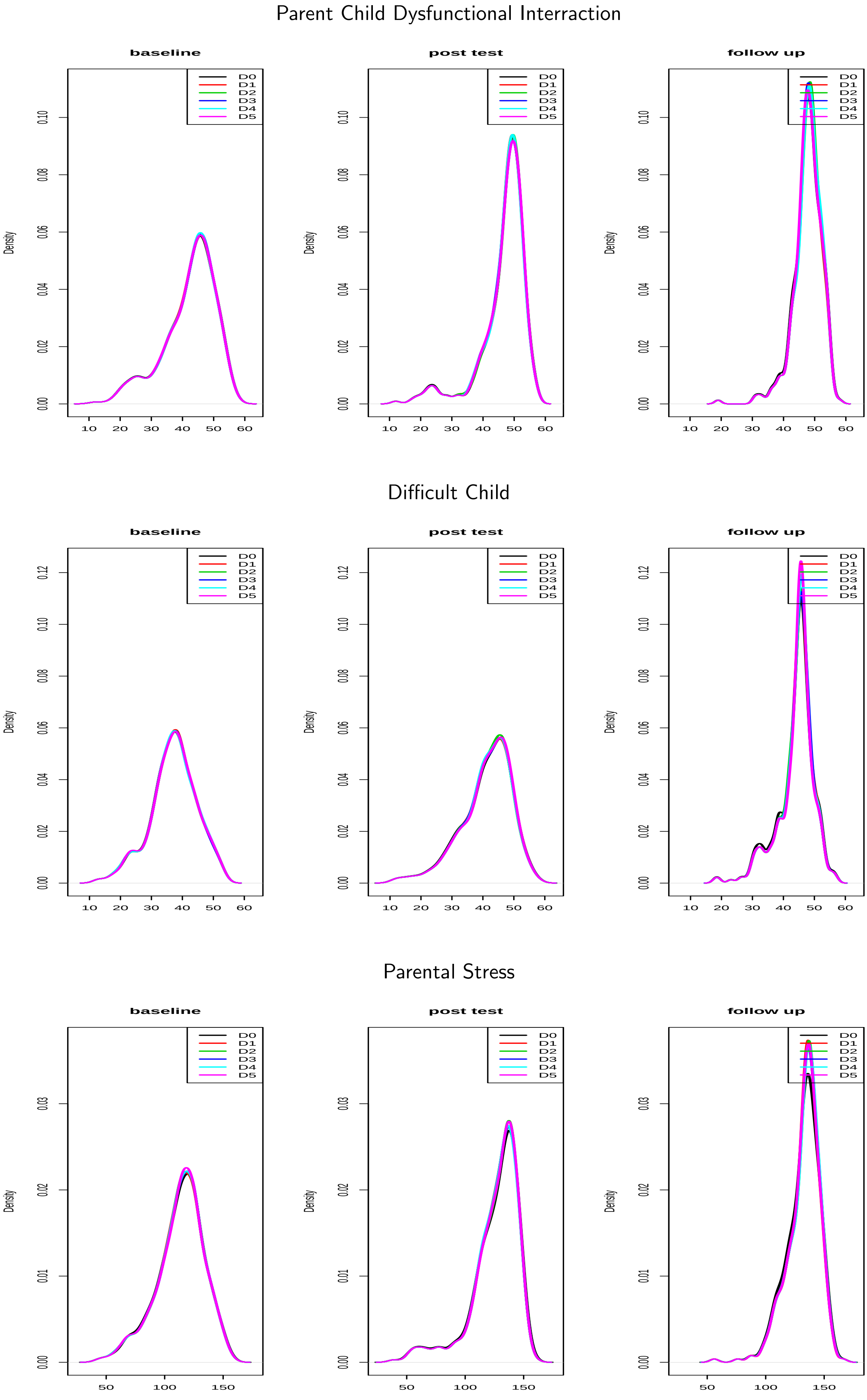

*D0 is unimputed and D1 - D5 are the 5 imputed datasets

Figure 3.4 (i): **Kernel Density Plots for Imputed and Unimputed Datasets**




*D0 is unimputed and D1 - D5 are the 5 imputed datasets

Figure 3.4 (j): **Kernel Density Plots for Imputed and Unimputed Datasets**




*D0 is unimputed and D1 - D5 are the 5 imputed datasets

## 3.7 Imputation: Chapter Summary and Conclusion

To summarize, this chapter dealt with imputing the missing data from the RCT implemented to evaluate the effectiveness of the Sinovuyo Caring Families Programme. One major assumption made throughout this chapter was that the data was missing at random (MAR) which implied that the patterns in the observed data were enough to inform on the patterns of the missing data. Imputation methods like predictive mean matching, logistic regression and random forest imputation were applied as a result. It was found that all the models where predictive mean matching was applied struggled to converge and some of the chains didn't even move away from their starting points in 100 iterations of the **mice** program. Random forests were ultimately chosen for imputing all the variables regardless of data type, a decision that simplified the coding of the imputation program without compromising the convergence of the chains, random forests proved to always converge quickly. One issue of concern which may be considered for future research here is to analyse the missing not at random (MNAR) assumption as suggested in the analysis plan. For this analysis, it was deemed sufficient to assume MAR. The results aren't expected to differ much here since there was a low incidence of missingness in the data. We could also have chosen to impute separately within the two randomized arms as opposed to including arm in the model and thus allowing for models to differ within these arms.

# Chapter 4

# Modelling study outcomes using generalized linear mixed models (GLMM)

Up to now, focus has been drawn on data features that were dealt with in an attempt to understand and prepare the dataset for analysis. Chapter 2 dealt with exploring the data by focussing on issues like score construction and reliability,describing the sample at baseline and exploring the extent of missingness in the data. Methods for handling missing data were discussed in Chapter 3, in particular how multiple imputation using chained equations was implemented for this particular dataset. After the multiple imputation process, there were five "complete" datasets ready for analysis and pooling of results.

This chapter will explore the methods used in analysing the aforementioned data. It also outlines how the analysis and pooling of results was done and also includes a discussion of the results. From Chapter 2 it was seen that most of the outcomes of interest had skewed densities and some were very highly peaked at zero (zero-inflated). This introduces the need to model data using generalized linear models (GLMs). Sections 4.1 and 4.2 briefly outlines the theory behind linear models and GLMs as an extension of linear models. Since the dataset at hand is longitudinal (i.e. was collected on the same set of participants on three different timepoints), generalized linear mixed-effects models (GLMMs) are also discussed in Section 4.3 as an extension of GLMs aimed at to addressing the dependency in data. Further dependencies that required the use of GLMMs arose from the design of the RCT that implemented a group-based intervention.

The model specifications and model fitting procedures are discussed in Sections 4.5 and 4.6 respectively. The former explores the two types of GLMMs that were fitted as differentiated by their treatment of the intervention session attendance and also their treatment of some interaction terms. This section also discusses the pooling of results using Rubin's rules. The latter deals with how the aforementioned models were actually implemented on the multiple datasets in order to get the results. The rest of the chapter discusses the model results as well as model diagnostics.

## 4.1 Linear Models

The simplest models assume a linear relationship between the response ($\mathbf{Y}$) and exposures ($\mathbf{X}_1, \mathbf{X}_2, \ldots, \mathbf{X}_p$) i.e.

$$\mathbf{Y} = \beta_0 + \beta_1\mathbf{X}_1 + \beta_2\mathbf{X}_2 + \cdots + \beta_p\mathbf{X}_p + \epsilon \tag{4.1}$$

where:

- $\boldsymbol{\epsilon}$ is a vector of error terms, $\boldsymbol{\epsilon} \sim$ Normal$(\mathbf{0}, \sigma^2 I_n)$ i.e. the errors are independent (and so are the individual responses), normally distributed with a zero mean and an equal variance.
- $\beta_0, \beta_1, \ldots, \beta_p$ are the regression parameters governing the relationship between the response and the exposures. Combined with the above, one can estimate these using maximum likelihood (which in this case would be equivalent to least squares estimation) in such a way that $\boldsymbol{\hat{\beta}_{MLE} = (X^T X)^{-1} X^T} Y$ where $\boldsymbol{\hat{\beta}} = (\hat{\beta}_0, \hat{\beta}_1, \ldots, \hat{\beta}_p)^T$ and $\mathbf{X} = (1, \mathbf{X}_1, \mathbf{X}_2, \ldots, \mathbf{X}_p)$, the latter is also known as the design matrix.

The above would be suitable when modelling data that satisfies the following assumptions:

1. The response is suspected to be normally distributed. This can be visually checked by plotting histograms of the sampled responses to get an idea of the distribution.
2. There is no suspicion of dependence between observations. This eliminates all repeated measures data where dependence between observations is within the subject.
3. One suspects that there is an identity relationship between the expected response and the linear combination of the predictors, i.e.

$$E(\mathbf{Y}) = \beta_0 + \beta_1\mathbf{X}_1 + \beta_2\mathbf{X}_2 + \cdots + \beta_p\mathbf{X}_p \tag{4.2}$$

4. The error variance $\sigma^2$ is constant. This can be checked by performing residual analysis.

The models in the following sections relax some of the above-mentioned assumptions so as to allow for linear models to be fitted for data with more complex structures.

## 4.2 Generalized Linear Models (GLMs)

GLMs (Dobson and Barnett, 2008) are an extension of the least squares linear regression models discussed in Section 4.1 above. They relax the assumption (made in linear regression) that the dependent variable must be normally distributed and also relax the identity link (that the mean is linearly related to the predictors) assumed with linear models. As such, GLMs are used to model/ fit the mean or some monotonic transformation of the mean as a linear combination of the predictors for differently distributed responses. That is, given a dependent variable $\mathbf{Y}$, one can model:

$$g(E(\mathbf{Y})) = \beta_0 + \beta_1\mathbf{X}_1 + \cdots + \beta_p\mathbf{X}_p \tag{4.3}$$

for some link function $g()$ with predictors $\mathbf{X}_1, \ldots, \mathbf{X}_p$ where $\beta_0, \beta_1, \ldots, \beta_p$ are the regression parameters (also known as fixed fixed effects, see Section 4.3) summarizing the linear relationship between the predictors and (transformed) mean of $\mathbf{Y}$. The equation for subject/ participant $i$, using this formulation becomes:

$$g(E(\mathbf{Y}_i)) = \beta_0 + \beta_1\mathbf{X}_{i,1} + \cdots + \beta_p\mathbf{X}_{i,p} \tag{4.4}$$

where $\mathbf{Y}_i$ is the $i^{th}$ entry in the vector $\mathbf{Y}$ and $\mathbf{X}_{i,j}$ is the $i^{th}$ subject's value for the $j^{th}$ covariate.

The choice for this link function usually depends on the assumed distribution of the response/ dependent variable $\mathbf{Y}$ and the shape of the relationship between mean of $\mathbf{Y}$ and linear combination of predictors. A key feature of GLMs is that the distribution of the response must be from the exponential family. A probability density function $f(y;\theta)$ is said to be part of the exponential family if it can be expressed in the following 'exponential' form:

$$f(y;\theta) = \exp\{a(y).b(\theta) + c(\theta) + d(y)\} \tag{4.5}$$

where $b()$ is known as the natural parameter of the distribution in the case where $a(y) = y$ is the identity, see Dobson and Barnett (2008, pp. 51). Examples include the Poisson, Binomial, Normal, Exponential and the Geometric distribution. The following three examples illustrate how the binomial, Poisson and Normal distributions are part of the exponential family:

$$\begin{aligned} Y &\sim \text{Binomial}(n,\theta) \\ P(Y=y;\theta) &= \binom{n}{y}\theta^y(1-\theta)^{n-y} \\ &= \exp\left\{\ln\left(\binom{n}{y}\right) + y\ln(\theta) + (n-y)\ln(1-\theta)\right\} \\ &= \exp\left\{\ln\left(\binom{n}{y}\right) + y\ln\left(\frac{\theta}{1-\theta}\right) + n\ln(1-\theta)\right\} \end{aligned} \tag{4.6}$$

here, $a(y) = y, b(\theta) = \ln\left(\frac{\theta}{1-\theta}\right), c(\theta) = n\ln(1-\theta)$ and $d(y) = \ln\left(\binom{n}{y}\right)$.

$$\begin{aligned} Y &\sim \text{Poisson}(\theta) \\ P(Y=y;\theta) &= \frac{\theta^y e^{-\theta}}{y!} \\ &= \exp\{y\ln(\theta) - \theta - \ln(y!)\} \end{aligned} \tag{4.7}$$

here, $a(y) = y, b(\theta) = \ln(\theta), c(\theta) = -\theta$ and $d(y) = -\ln(y!)$.

$$\begin{aligned} Y &\sim \text{Normal}(\mu,\sigma^2) \\ f(y;\mu,\sigma^2) &= \frac{1}{\sqrt{2\pi\sigma^2}}\exp\left\{-\frac{1}{2\sigma^2}(y-\mu)^2\right\} \\ &= \exp\left\{-\frac{y\mu}{\sigma^2} + \frac{\mu^2}{2\sigma^2} - \frac{\ln(2\pi\sigma^2)}{2} + \frac{y^2}{2\sigma^2}\right\} \end{aligned} \tag{4.8}$$

here $\sigma^2$ is being treated as a nuisance parameter, $a(y) = y, b(\mu) = -\frac{\mu}{\sigma^2}, c(\mu) = \frac{\mu^2}{2\sigma^2} - \frac{\ln(2\pi\sigma^2)}{2}$ and $d(y) = \frac{y^2}{2\sigma^2}$. A common choice for the aforementioned link function for distributions in the exponential family is $b(\theta)$. Following from the above illustration, if the response is deemed to be binomial, then the log of the odds (also known as the logit transform) is then modelled as the linear combination of the predictors i.e.

$$\ln\left(\frac{\theta}{1-\theta}\right) = \beta_0 + \beta_1 X_1 + \cdots + \beta_p X_p. \tag{4.9}$$

For most of the common distributions, the link function is usually the log, logit or inverse. All these functions typically have the additional necessary characteristic that their ranges span the entire real number set, $\mathbb{R}$. This ensures that the values for both sides of equation (4.9) have the same range. Referring back to the binomial example, it would be inappropriate to just model $\theta$ as the linear combination of the predictors $X_1, \ldots, X_p$ as the former is only defined in [0, 1] whilst the latter can take on any values in $\mathbb{R}$. The aforementioned logit transformation will have the desired range and can still be mapped back to $p$ since it is invertible (i.e. 1:1 mapping). This is similar to when the parameter $\theta$ takes on non-negative values excluding zero, a log transformation would have the desired range and is invertible.

The regression parameters are estimated using maximum likelihood via a process called iteratively re-weighted least squares outlined in Dobson and Barnett (2008, pp. 68- 73). The main assumption needed here is that the observations are statistically **independent**. Since the dataset being considered is longitudinal, it will be inappropriate to fit GLMs to this dataset as the independence assumption is violated. More clearly, the data was collected in such a way that there are repeated measurements (in time) for every subject in the study and it would be inappropriate to just assume that there are no within-subject correlations in the data especially across time points. Furthermore, there is suspicion that individuals' responses to the intervention programme are also affected by the group that the participant was placed in. This also further shows why assuming independence between subjects might be inappropriate for analysing this specific dataset. For these reasons, Generalized linear Mixed Models (GLMMs) were used.

## 4.3 Generalized Linear Mixed-Effects Models (GLMM)

In short, GLMMs (McCulloch, 2003) are an extension of GLMs whose main distinguishing feature is the relaxation of the independence (between observations) assumption that both GLMs and linear regression models are based on. This is done through the addition of random effects to the already discussed fixed effects in the regression equation. These random effects enable the models to capture the dependence structure imposed by some grouping variables in the dataset. In doing so, GLMMs allow for regression parameters to vary with any grouping variables that are suspected to be causing differences in responses. For the study at hand, GLMMs were used to allow for variation to be dependent on **each subject/ child-caregiver dyad** (since there was a repeated measurements component to the study) and **programme group** since one could easily suspect that the impact of the intervention also relied on the other participants that one was grouped with and the facilitator pair assigned to a group.

A general case for the mixed effects model is when one adds a random effect to each of the terms of the model. This is usually the case when one expects the different individuals to react have different intercepts/ starting points and different reactions to all the covariates in the model. Such a model can be represented as follows:

$$g(E(Y_i)) = \beta_0 + b_{i,0} + (\beta_1 + b_{i,1})X_{i,1} + \cdots + (\beta_p + b_{i,p})X_{i,p} \tag{4.10}$$

where $\beta_k,\ \ k = 1, 2, \ldots, p$ are called the fixed effects and the $b_{i,j},\ \ i = 1, 2, \ldots, N; j = 0, 1, 2, \ldots, p$ are the subject specific random effects assumed to be random draws from a normal distribution with a zero mean and a variance-covariance structure to be estimated along with the betas and any other additional parameters needed to fully define the distribution. Together the fixed and random effects form the mixed effects model. As mentioned before, the random effects can also be defined on a group level rather than a subject/ participant level and in some models one can have both subject-specific and group specific random effects. The SCFP is a good example of this scenario, see the model formulation in sections 4.5.1 and 4.5.2 below. Estimation of the regression parameters is usually done through maximizing the likelihood function, which is preferred especially when considering implicit imputation of data using the measurement and correlation models. Restricted maximum likelihood (REML) is also common mainly because it produces less biased estimates of variance parameters by limiting the effect of some nuisance parameters. One can also estimate the parameters using the penalized likelihood function where a shrinkage factor is introduced with the aim of improving the stability of estimates in exchange for introducing some bias. Another approach involves maximizing the penalized quasi-likelihood which in addition to improving the stability of estimates, also accounts for over-dispersion that might occur especially in grouped data.

## 4.4 Distributions Used for Modelling the Outcomes

The outcomes of interest were introduced in Chapter 2 as sums of different reported or observed variables that were either measured using Likert scales or were binary indicators of presence/absence of a trait or behaviour. These scores can be viewed as count or continuous variables either including zeros or strictly positive (refer to Table 2.5).

Possible distributions that were considered include the Normal distribution (for variables that had symmetric empirical distributions), Negative Binomial (for variables with empirical distributions showing an excess of zeros and over-dispersed data), the Poisson (for count data), Generalized Gamma and Log-normal (for skew data) and the Zero-Altered Gamma. These for this project, only the Normal, Negative Binomial, Poisson and Generalized Gamma were used in the final models. Their density functions, means and variances are shown below:

1. **Normal Distribution**: If $Y \sim \mathrm{N}(\mu, \sigma^2)$ then the density function is

$$f(y;\mu,\sigma^2) = \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left\{-\frac{(y-\mu)^2}{2\sigma^2}\right\} = \exp\left\{-\frac{1}{2}\ln(2\pi\sigma^2) - \frac{y^2}{2\sigma^2} + \frac{\mu y}{\sigma^2} - \frac{\mu^2}{2\sigma^2}\right\} \tag{4.11}$$

Where $-\infty < y < \infty$; $-\infty < \mu < \infty$ and $0 < \sigma^2 < \infty$, Here the mean is $\mu$ and the variance is $\sigma^2$. With ordinary linear regression one is only interested in modelling the mean as a function of the predictors, i.e. considering the variance to be a nuisance parameter.

2. **Negative Binomial** type I: If $Y \sim \mathrm{NBI}(\mu, \sigma)$ then the probability mass function (pmf) as in Stasinopoulos *et al.* (2016, pp. 63) would be:

$$P(Y = y;\mu,\sigma) = \frac{\Gamma\left(y+\frac{1}{\sigma}\right)}{\Gamma\left(\frac{1}{\sigma}\right)\Gamma(y+1)} \left(\frac{\sigma\mu}{1+\sigma\mu}\right)^y \left(\frac{1}{1+\sigma\mu}\right)^{\frac{1}{\sigma}} \tag{4.12}$$

where y = 0,1,2,. . . , the mean is $\mu > 0$ and variance is $(1 + \sigma\mu)\mu$ where $\sigma > 0$.

3. The **Generalized Gamma distribution (GG)** has 3-parameters and if $Y \sim GG(\mu, \sigma, \nu)$ then the probability density function as in Stasinopoulos *et al.* (2016, pp. 238) would be:

$$f(y;\mu,\sigma,\nu) = \frac{|\nu|\theta^\theta z^\theta \exp\{-\theta z\}}{\Gamma(\theta) y} \tag{4.13}$$

with $y > 0, \mu > 0, \sigma > 0$ and $-\infty < \nu < \infty$; $z = \left(\frac{y}{\mu}\right)^\nu$ and $\theta = \left(\frac{1}{\sigma\nu}\right)^2$. The mean of the distribution is: $\frac{\mu\Gamma\left(\theta+\frac{1}{\nu}\right)}{\theta^{\frac{1}{\nu}}\Gamma(\theta)}$ and the variance is: $\mu^2\left\{\frac{\Gamma(\theta)\Gamma\left(\theta+\frac{1}{\nu}\right)-\left[\Gamma\left(\theta+\frac{2}{\nu}\right)\right]^2}{\theta^{\frac{2}{\nu}}[\Gamma(\theta)]^2}\right\}$

4. The **Poisson Distribution** is such that if $Y \sim \text{Poisson}(\mu)$ them the pmf is:

$$P(Y = y; \mu) = \frac{e^{-\mu}\mu^y}{y!} \tag{4.14}$$

where $y = 0, 1, 2, \ldots,\ \mu > 0$ and also the mean and variance are all equal to $\mu$.

## 4.5 Model Specification

There are two types of models fitted in analysing the SCFP data: (1) the binary-intervention models and (2) the dose-response models. The former focused on comparing the control arm to the intervention arm with the hope that the participants in the intervention arm would have significantly more desirable outcomes at post-test and even at the one year follow-up visits. The latter focused on comparing whether participants' level of attendance to the group sessions (in the intervention arm) made significant differences. Ideally the dose-response models were hoped to show that increased attendance was associated significant improvements in the outcomes over time.

### 4.5.1 Binary-Intervention models

As discussed above, the binary-intervention models consider whether the participants in the intervention arm significantly differ from those in the control arm with respect to the outcomes at the post-test visit and at one year follow-up. Here, the purpose was to evaluate whether the intervention program actually worked as intended and this was done by including a binary indicator in the model showing whether the participant was in the control or in the intervention arm. The hope was that participants in the intervention arm had significantly better outcomes over time i.e. if the a higher score was the more desirable then the clinicians would expect the intervention arm to have significantly higher scores and vice versa. This model was implemented in an intention-to-treat (ITT) analysis as well as a per-protocol analysis (PP). The ITT analyses are used provide overall unbiased comparisons of the control and intervention arms as 'intended' in the protocol. However, one of the common occurrences in RCTs is that not all participants get the intended 'treatment' because of non-attendance and as such the ITT analyses might not be reflecting the true effect of the treatment. Per-Protocol analyses are meant to compare those who actually received the treatment in sufficient dosage to the controls so as to see the true effect of the treatment. The PP analyses are only unbiased there is no evidence of biased non-attendance in the RCT. For the ITT analysis, all the 296 participants from the two arms were included. A total of 71 participant dyads in the intervention arm who attended six or fewer group sessions (out of the 12) were excluded from the PP analysis effectively making the number in the intervention arm 77 for this part. The control arm remained the same in these two types of analyses.

For **child-caregiver dyad** $j$ in **group** $i$, the binary intervention model specifies:

$$\begin{aligned} log(E(Y_{i,j})) = &(\beta_0 + b_{0,i,j}) + \beta_{wave}.\delta_{wave,i,j} + \beta_{sex}.\delta_{sex,i,j} + \beta_{age}.\delta_{age,i,j} + \beta_{time1}.\delta_{time1,i,j} \\ &+\beta_{time2}.\delta_{time2,i,j} + (\beta_{armtime1} + b_{armtime1,i})\,.\delta_{arm,i,j}.\delta_{time1,i,j} \\ &+ (\beta_{armtime2} + b_{armtime2,i})\,.\delta_{arm,i,j}.\delta_{time2,i,j} \end{aligned} \tag{4.15}$$

where

- $i = 0, 1, 2, \ldots, 11$; all participants in the Control Arm are in group 0 and participants in the intervention arm fall in one of groups 1 to 11;
- $j = 0, 1, 2, \ldots, n_i$ and $n_i$ is the number of participant pairs in group $i$;
- $\delta_{time1,i,j} = 0, 1$ an indicator attaining 1 if time-point is the post-test visit and 0 otherwise;
- $\delta_{time2,i,j} = 0, 1$ an indicator attaining 1 if time-point is the one year follow-up visit and 0 otherwise;
- $\delta_{arm,i,j} = 0, 1$ for the control arm and intervention arm respectively;
- $\delta_{wave,i,j} = 0, 1$ indicating the different waves, i.e. Nyanga (0) and Khayelitsha (1);
- $\delta_{sex,i,j} = 0, 1$ for the child gender: female (0) and male (1);
- $\delta_{age,i,j} = 0, 1$ for the child's age: 2 - 5 years (0) and 6 - 9 years (1);
- the terms $\beta_0, \beta_{time1}, \beta_{time2}, \beta_{wave}, \beta_{sex}, \beta_{age}$, $\beta_{armtime1}$ and $\beta_{armtime2}$ are treated as fixed effects whilst $b_{0,i,j}$ is a subject-specific random effect and $b_{at,i}$ is a group-specific random effect, assumed to be normally distributed with a zero mean and a covariance structure to be estimated together with all the other parameters in the model.

Of particular interest are $\beta_{armtime1}$ and $\beta_{armtime2}$ here as they inform on how the intervention arm differs from the control arm with respect to the changes in outcomes from baseline to post-test and from baseline to follow-up, respectively. As described in Chapter 3, all these measurement models are being adjusted for child's age category and gender, wave of the study and the study arm. The group-specific random effect on the arm-time interaction adjusts the slope of the response for the group-effect in the two aforementioned time intervals . The random effect placed on the intercept ensures that the model treats each dyad as if they have a different starting point allowing some subjects to always score score higher/ lower than others.

### 4.5.2 Dose-Response models

The second model considers the fact that not all the participants in the intervention group received the same level of treatment. Table 4.1 below shows the total number of group sessions that the different participants attended. About 48% of the participants in the intervention arm only attended six or fewer group sessions and 27.7% didn't attend any. The dose-response models take session attendance (as a continuous random variable) into account instead of just looking broadly at the arm of the study. The models specifically assume some linear relationship between session attendance and the score improvements, this could be one of the weakness of the models especially considering that the sessions focussed on addressing different aspects of behaviour. These models also take into account that **baseline intimate partner violence (ipv)** as well as **wave** may affect the participants' responses to the intervention program.

Table 4.1: Group session attendance among the participants in the intervention arm

| **Number of Sessions Attended** | 0 | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 | 12 |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| **Number of participants** | 41 | 6 | 5 | 3 | 9 | 6 | 1 | 11 | 16 | 12 | 11 | 14 | 13 |
| **Cumulative percentage** | 27.7 | 31.8 | 35.1 | 37.2 | 43.2 | 47.3 | 48.0 | 55.4 | 66.2 | 74.3 | 81.8 | 91.2 | 100 |

For **child-caregiver dyad** $j$ in **group** $i$, the dose-response model specifies:

$$\begin{aligned} log(E(Y_{i,j})) = & (\beta_0 + b_{0,i,j}) + \beta_{time1}.\delta_{time1,i,j} + \beta_{time2}.\delta_{time2,i,j} \\ & + (\beta_{att:time1} + b_{att:time1,i}) .X_{att,i,j}.\delta_{time1,i,j} + (\beta_{att:time2} + b_{att:time2,i}) .X_{att,i,j}.\delta_{time2,i,j} \\ & + \beta_{wave}.\delta_{wave,i,j} + \beta_{ipv}.\delta_{ipv,i,j} + \beta_{sex}.\delta_{sex,i,j} + \beta_{age}.\delta_{age,i,j} \\ & + \beta_{wave:time1}.\delta_{wave,i,j}.\delta_{time1,i,j} + \beta_{wave:time2}.\delta_{wave,i,j}.\delta_{time2,i,j} \\ & + \beta_{ipv:time1}.\delta_{ipv,i,j}.\delta_{time1,i,j} + \beta_{ipv:time2}.\delta_{ipv,i,j}.\delta_{time2,i,j} \\ & + \beta_{ipv:att:time1}.\delta_{ipv,i,j}.X_{att,i,j}.\delta_{time1,i,j,k} \\ & + \beta_{ipv:att:time2}.\delta_{ipv,i,j}.X_{att,i,j}.\delta_{time2,i,j,k} \\ & + \beta_{wave:att:time1}.\delta_{wave,i,j}.X_{att,i,j}.\delta_{time1,i,j,k} \\ & + \beta_{wave:att:time2}.\delta_{wave,i,j}.X_{att,i,j}.\delta_{time2,i,j,k} \end{aligned} \tag{4.16}$$

where

- $i = 0, 1, 2, \ldots, 11$; all participants in the Control Arm are in group 0 and participants in the intervention arm fall in one of groups 1 to 11;
- $j = 0, 1, 2, \ldots, n_i$ and $n_i$ is the number of participant pairs in group $i$;
- $\delta_{time1,j,k} = 0, 1$ an indicator for time 1;
- $\delta_{time1,j,k} = 0, 1$ an indicator for time 2;
- $X_{att,i,j} = 0, 1, 2, \ldots, 12$ for the number of group sessions attended with controls having a value of 0 by default;
- $\delta_{wave,i,j} = 0, 1$ indicating the different waves, i.e. Nyanga (0) and Khayelitsha (1);
- $\delta_{sex,i,j} = 0, 1$ for the child gender: female (0) and male (1);
- $\delta_{age,i,j} = 0, 1$ for the child's age: 2 - 5 years (0) and 6 - 9 years (1);
- $\delta_{ipv,i,j} = 0, 1$ for the caregiver's baseline experience of ipv: None (0) and Some (1);
- the terms $\beta_0, \beta_{time1}, \beta_{time2}, \beta_{wave}, \beta_{sex}, \beta_{age}, \beta_{ipv}, \beta_{wave:time1}, \beta_{wave:time2}, \beta_{ipv:time1}, \beta_{ipv:time2}, \beta_{att:time1}, \beta_{att:time2}, \beta_{ipv:att:time1}, \beta_{ipv:att:time2}, \beta_{wave:att:time1}$, and $\beta_{wave:att:time2}$, are treated as fixed effects whilst $b_{0,i,j}$ is a subject-specific random effect and $b_{att:time1,i}$ and $b_{att:time2,i}$ are group-specific random effects. These random effects would all be assumed to be normally distributed with a zero mean and a covariance structure to be estimated together with all the other parameters in the model.

Here, the coefficients for the third order interaction terms (i.e. $\beta_{ipv:att:time1}$, $\beta_{ipv:att:time2}$, $\beta_{wave:att:time1}$, and $\beta_{wave:att:time2}$,) would show whether **ipv** or **wave** impact the impact the slope of the response (with respect to attendance) within the two aforementioned time-periods. The group-specific random effect placed on the attendance-time effect enforces that each group have a different 'slope' of the response with respect to attendance (referred to as the response to dosage henceforth), all else being equal.

## 4.6 Model Fitting Details

The models described in the above section were all fitted using two **R** packages: **glmmPQL** (Venables and Ripley, 2002) and **gamlss** (Rigby and Stasinopoulos, 2005). The former (glmmPQL) obtains parameter estimates by maximizing the penalized quasi-likelihood and the latter applies the Fisher scoring algorithm to maximize the penalized likelihood function. Table 4.2 lists all 22 the outcomes of interest that were assessed using the binary intervention models and their specifications. Residual analyses were done in order to check the model fit. The model fitting procedure also ensured that count distributions (the Poisson and Negative Binomial) were used for count variables (mainly the problem scores with a narrower support/range and the observed variables). The Negative binomial proved to sufficiently account for the peaks at zero. For the more complex dose-response models, the researchers identified ten primary outcomes for analysis. As discussed in Section 4.4 only four distributions (Gaussian, Negative Binomial: type 1, Generalized Gamma and Poisson) were used for the final models. All were fitted with the log link so that the exponentiated betas could be interpreted as relative changes not as absolute effects. Only the Gaussian models were modelled using **glmmPQL**.

Table 4.2: Summary of model specifications in **R**

| Model Number | Response | family | link | R function used |
|---|---|---|---|---|
| 1 | ECBI intensity* | gaussian | log | glmmPQL |
| 2 | ECBI problem* | gaussian | log | glmmPQL |
| 3 | Supporting Positive Behaviour Frequency | gaussian | log | glmmPQL |
| 4 | Setting Limits Frequency reversed | generalized gamma | log | gamlss |
| 5 | Positive Parenting Frequency* | gaussian | log | glmmPQL |
| 6 | Non-Violent Discipline | NBI | log | gamlss |
| 7 | Poor Monitoring And Supervision | NBI | log | gamlss |
| 8 | Parental Distress reversed | NBI | log | gamlss |
| 9 | Difficult Child reversed | NBI | log | gamlss |
| 10 | Parenting Stress reversed | NBI | log | gamlss |
| 11 | Social Support | gaussian | log | glmmPQL |
| 12 | Physical Discipline* | NBI | log | gamlss |
| 13 | Psychological Discipline* | NBI | log | gamlss |
| 14 | Supporting Positive Behaviour Problem | Poisson | log | gamlss |
| 15 | Setting Limits Problem | NBI | log | gamlss |
| 16 | Positive Parenting Problem* | NBI | log | gamlss |
| 17 | Beck Depression Inventory | NBI | log | gamlss |
| 18 | PCDI reversed | NBI | log | gamlss |
| 19 | Parent Positive Behaviour* | Poisson | log | gamlss |
| 20 | Parent Negative Behaviour* | Poisson | log | gamlss |
| 21 | Child Negative Behaviour* | Poisson | log | gamlss |
| 22 | Child Positive Behaviour* | Poisson | log | gamlss |

* Response also included in the Dose-Response analysis

One problem encountered in fitting the models especially with **gamlss** was that the estimation algorithms would not converge in some circumstances. This was very common especially when the model was complicated especially the dose-response models where many parameters had to be estimated. In many cases, this was mainly because of bad starting points that would lead to the algorithms being stuck in regions where the deviance (the error measure used in the maximization algorithms) would not decrease in order to meet the the convergence criteria. This was easily solved by fitting simpler models (usually with a more simplified random effects structure) and using it's estimates as starting points to the more complicated models. In some cases, convergence issues were solved by simplifying the correlation structure of the random effects such that they would become independent.

The models were fitted on the unimputed data as well as the five imputed data sets. Final estimates from the imputed data sets were then pooled using Rubin's rules. All the pooling was done on the coefficient scale (i.e. the log scale) and the pooled results were then exponentiated to get the effects. Suppose that $\tilde{\beta}_j$ and $\tilde{V}_j; j = 1, \ldots, 5$ are the respective estimates and estimates variance (of the regression coefficients) obtained from the five imputed data sets, the pooled estimator $\tilde{\beta}_{MI}$ and $\tilde{V}_{MI}$ are obtained using Rubin's rules as follows:

1. The average over all data sets was used for betas, i.e. $\tilde{\beta}_{MI} = \frac{1}{5}\sum_{j=1}^{5}\tilde{\beta}_j$.

2. The sum of between- and within-imputation variances was used for the pooled variance, i.e. $\tilde{V}_{MI} = \bar{V} + \left(1 + \frac{1}{5}\right) B$

   where $\bar{V} = \frac{1}{5}\sum_{j=1}^{5} \tilde{V}_j$ and $B = \frac{1}{5-1}\sum_{j=1}^{5}\left(\tilde{\beta}_j - \tilde{\beta}_{MI}\right)^2$.

3. Normal approximations to the t-distribution were used (since the degrees of freedom for each estimate were above 200) and 95% confidence intervals were computed using the critical value of 1.959964 (the 97.5% quantile of the standard normal) was used in the calculation.

4. These estimates and confidence intervals were then exponentiated in order to get the relative effects since a log link was used.

The summary statistics for observed outcomes are presented in Section 4.7 and the model results are illustrated in Sections 4.8 and 4.9 below.

## 4.7 Summary Statistics for Observed Outcomes

Table 4.4 summarizes the observed outcomes of interest between the control arm, the intervention arm as in the ITT analysis and the intervention arm as in the PP analysis over the three time points. Although the actual models reported in this section are based on within-subject/ group trajectories, this table shows an indication of the general direction of the scores over time and also shows the differences between the PP and ITT samples. Table 4.4 shows a substantial drop in child behaviour problems between the baseline and post-test visit and the lower score is maintained through to the one-year follow-up which is to be expected since this was used as the screening tool. The profiles for the ITT and PP samples are smaller and both the ITT and PP samples seem to show a steeper decline than the control group. Positive parenting scores do not exhibit any noticeable differences between the groups and also don't show any substantial differences over time. There are differences for setting limits problem and for positive parenting problem Harsh parenting scores seem to have substantial improvements from the baseline scores and there is a little difference between the control, ITT and PP samples.

Observed parenting and child behaviour scores also do not show clear differences over time but noticeably the control arm seems to have substantially lower scores compared to both versions of the intervention arm on the observed parent positive behaviour score. Both the depression, social support and monitoring and supervision scores also don't seem to differ between groups and over time. Noticeably, parenting stress scores seem to be undesirably higher at the later visits as compared to the baseline but the differences between the control, ITT and PP groups doesn't seem big. Overall, the majority of the scores show some improvement over time except for parenting stress and the ITT and PP samples don't seem to be substantially different which hints that the PP results are not biased.

Table 4.4: Summaries (means and standard deviations) of observed outcomes for samples used in PP vs ITT analyses

| | | Baseline | Post-Test | Follow-up |
|---|---|---|---|---|
| **Child Behaviour Problems** | | | | |
| ECBI Intensity | Control | 143.00 (23.14) | 115.64 (27.5) | 104.16 (26.81) |
| | ITT | 141.21 (22.85) | 111.58 (24.82) | 100.59 (26.64) |
| | PP | 141.75 (22.56) | 108.11 (26.60) | 98.21 (28.07) |
| ECBI Problem | Control | 25.16 (4.94) | 18.03 (8.46) | 13.01 (8.73) |
| | ITT | 24.61 (5.07) | 16.05 (7.93) | 12.78 (8.48) |
| | PP | 24.55 (4.74) | 14.27 (8.07) | 11.89 (8.33) |
| **Positive Parenting** | | | | |
| Supporting Positive Behaviour Frequency | Control | 26.57 (6.10) | 27.07 (5.50) | 28.16 (5.08) |
| | ITT | 25.76 (5.06) | 29.29 (5.65) | 28.16 (5.18) |
| | PP | 26.04 (4.96) | 30.91 (5.39) | 28.89 (5.66) |
| Setting Limits Frequency | Control | 22.66 (7.46) | 22.78 (7.58) | 25.46 (5.97) |
| | ITT | 21.79 (6.34) | 25.57 (5.89) | 24.89 (7.18) |
| | PP | 22.36 (5.55) | 27.01 (4.61) | 24.33 (7.70) |
| Positive Parenting Frequency | Control | 49.23 (11.42) | 49.88 (11.33) | 53.62 (8.99) |
| | ITT | 47.55 (9.31) | 54.86 (9.37) | 53.05 (9.99) |
| | PP | 48.40 (8.26) | 57.92 (8.02) | 53.21 (10.98) |
| Supporting Positive Behaviour Problem | Control | 2.16 (1.94) | 1.74 (1.81) | 1.02 (1.44) |
| | ITT | 1.91 (1.74) | 1.30 (1.65) | 1.01 (1.55) |
| | PP | 1.94 (1.63) | 0.83 (1.26) | 0.94 (1.44) |
| Setting Limits Problem | Control | 2.61 (2.34) | 1.86 (2.27) | 0.74 (1.26) |
| | ITT | 2.50 (2.10) | 1.45 (1.90) | 0.87 (1.48) |
| | PP | 2.40 (2.01) | 0.97 (1.60) | 0.97 (1.54) |
| Positive Parenting Problem | Control | 4.78 (3.96) | 3.66 (3.78) | 1.76 (2.41) |
| | ITT | 4.38 (3.36) | 2.76 (3.31) | 1.89 (2.75) |
| | PP | 4.29 (3.11) | 1.8 (2.62) | 1.91 (2.65) |
| **Harsh Parenting** | | | | |
| Non-Violent Discipline | Control | 6.34 (3.29) | 5.18 (2.94) | 4.21 (2.46) |
| | ITT | 6.42 (3.24) | 5.43 (2.93) | 4.92 (2.93) |
| | PP | 6.66 (3.00) | 5.47 (2.63) | 4.84 (2.86) |
| Physical Discipline | Control | 4.67 (4.19) | 2.77 (3.25) | 1.50 (1.94) |
| | ITT | 4.22 (3.47) | 1.99 (2.83) | 1.49 (2.27) |
| | PP | 4.26 (3.69) | 1.64 (2.38) | 1.33 (2.38) |
| Psychological Discipline | Control | 6.73 (5.28) | 3.72 (3.94) | 2.44 (2.94) |
| | ITT | 6.66 (4.37) | 3.14 (3.25) | 2.49 (2.50) |
| | PP | 6.44 (4.28) | 2.76 (2.85) | 2.29 (2.57) |

Table 4.4 (continued): Summaries (means and standard deviations) of outcomes for samples used in PP vs ITT analyses

| | | **Baseline** | **Post-Test** | **Follow-up** |
|---|---|---|---|---|
| **Observed Parenting and Child Behaviour** | | | | |
| Parent Positive Behaviour | Control | 12.63 (10.32) | 8.79 (8.45) | 6.79 (6.62) |
| | ITT | 15.27 (15.54) | 14.23 (14.95) | 10.83 (15.57) |
| | PP | 16.61 (18.23) | 14.97 (13.39) | 14.00 (19.84) |
| Child Positive Behaviour | Control | 27.94 (23.94) | 24.11 (21.17) | 16.39 (14.9) |
| | ITT | 26.76 (20.07) | 26.10 (22.57) | 18.57 (18.55) |
| | PP | 28.48 (20.83) | 30.18 (22.38) | 18.71 (18.69) |
| Parent Negative Behaviour | Control | 2.86 (3.37) | 2.49 (3.73) | 2.49 (5.74) |
| | ITT | 3.14 (4.59) | 2.67 (3.52) | 1.88 (2.84) |
| | PP | 3.13 (3.82) | 2.53 (2.92) | 1.95 (2.95) |
| Child Negative Behaviour | Control | 1.75 (2.92) | 1.42 (3.17) | 0.60 (1.64) |
| | ITT | 1.71 (3.98) | 1.51 (2.33) | 0.68 (1.50) |
| | PP | 2.30 (5.18) | 1.51 (2.03) | 0.80 (1.84) |
| **Monitoring and Supervision** | | | | |
| Poor Monitoring And Supervision | Control | 22.03 (4.6) | 21.06 (4.08) | 20.01 (3.65) |
| | ITT | 21.99 (4.53) | 21.85 (3.82) | 20.19 (3.27) |
| | PP | 22.89 (5.03) | 21.53 (3.27) | 19.76 (3.13) |
| **Depression** | | | | |
| Beck Depression Inventory | Control | 15.39 (12.12) | 10.90 (10.48) | 8.41 (10.62) |
| | ITT | 15.74 (10.90) | 8.75 (8.44) | 7.75 (9.26) |
| | PP | 16.29 (10.92) | 7.64 (7.22) | 6.72 (8.13) |
| **Parenting Stress** | | | | |
| Parental Distress | Control | 33.70 (8.50) | 36.65 (8.07) | 39.76 (7.62) |
| | ITT | 34.14 (7.92) | 38.96 (8.45) | 40.92 (6.97) |
| | PP | 34.55 (7.57) | 39.38 (9.10) | 41.63 (6.58) |
| Parent Child Dysfunctional Interaction | Control | 41.61 (8.78) | 46.08 (7.80) | 46.96 (5.29) |
| | ITT | 42.91 (8.05) | 46.89 (7.41) | 48.29 (4.01) |
| | PP | 42.92 (7.98) | 46.43 (8.40) | 48.59 (3.54) |
| Difficult Child | Control | 36.82 (8.02) | 40.41 (8.52) | 43.53 (6.54) |
| | ITT | 37.37 (7.16) | 41.63 (8.15) | 44.21 (5.70) |
| | PP | 37.32 (7.54) | 42.04 (8.22) | 44.91 (5.86) |
| Parenting Stress | Control | 111.92 (21.53) | 123.12 (21.22) | 130.25 (15.76) |
| | ITT | 114.59 (19.33) | 127.56 (21.16) | 133.42 (13.20) |
| | PP | 115.26 (19.32) | 128.01 (23.02) | 135.13 (12.64) |
| **Social Support** | | | | |
| Social Support | Control | 20.58 (6.29) | 20.60 (6.57) | 19.41 (6.68) |
| | ITT | 21.03 (5.98) | 20.51 (6.02) | 19.60 (6.53) |
| | PP | 21.05 (6.03) | 19.79 (6.19) | 18.96 (6.84) |

## 4.8 Results for Binary-Intervention Models

The following subsections present the GLMM results for each of the outcomes of interest that was modelled. For each outcome, the results are shown for both ITT and PP models implemented on the imputed and unimputed data sets. The inclusion of both the imputed and unimputed data is to show the impact of the imputation procedure discussed in Chapter 3. As discussed before, there was little missing data and as such it was not expected that these models would differ between the two datasets used. The tables show the following estimates of the fixed effects as included in equation 4.15:

- Intercept = $\beta_0$,
- wave (Khayelitsha) = $\beta_{wave}$ - compares wave 1 (Khayelitsha) to wave 2 (Nyanga) the reference level,
- sex (Male) = $\beta_{sex}$ - compares the dyads with boys to those with girls, reference category being girls,
- age (6-9 yrs) = $\beta_{age}$ - compares the dyads with older children (6-9 years old) to those with younger children (2 - 5 years old),
- time1 (control) = $\beta_{time1}$ - indicates the relative change in score the between baseline and post-test visits within the **control arm** which was the reference category in the model equation as laid out in 4.15,
- time2 (control) = $\beta_{time2}$ - indicates the relative change in score the between baseline and one year follow-up visits within the **control arm** which was the reference category in the model equation as laid out in 4.15,
- armtime1 = $\beta_{arm:time1}$ - this can be looked at in two ways: (1) the difference between the control and intervention arms with regards to the relative change in the outcome between the baseline and post-test visits **or equivalently** since this is an RCT and scores are similar across the arms at baseline (2) the relative difference with respect to the outcome between the control and the intervention arm at time 1 / post-test visit,
- armtime2 = $\beta_{arm:time2}$ this can be looked at in two ways: (1) the difference between the control and intervention arms with regards to the relative change in the outcome between the baseline and one year follow-up visits **or equivalently** since this is an RCT and scores are similar across the arms at baseline (2) the relative difference with respect to the outcome between the control and the intervention arm at time 2 / one year follow-up visit,
- time1 (intervention) = $\beta^*_{time1}$ - indicates the relative change in score the between baseline and post-test visits within the **intervention arm**. This can be computed in two ways: (1) recode the model data such that the reference category becomes the intervention arm or (2) multiply the $\beta_{time1}$ and $\beta_{arm:time1}$ from the above.
- time2 (intervention) = $\beta^*_{time2}$ - indicates the relative change in score the between baseline and one year follow-up visits within the **intervention arm**.This is achieved by setting the intervention arm as the reference category in the model. This can be computed in two ways: (1) recode the model data such that the reference category becomes the intervention arm or (2) multiply the $\beta_{time2}$ and $\beta_{arm:time2}$ from the above.

Since the main hypothesis of the study was to evaluate whether the programme was effective, plots of the relative time effects and their 95% confidence intervals are also shown to illustrate the differences between the two arms over time. In each of these plots, the relative time effects and their confidence intervals comparing the proportional change in the score between: (1) visit 1 vs visit 0 in the control arm, (2) visit 2 vs visit 0 in the control arm, (3) visit 1 vs visit 0 in the intervention arm and (4) visit 2 vs visit 0 in the intervention arm. Visually, the main study hypothesis aims to compare (1) to (3) and (2) to (4) effectively recovering the arm:time interactions over the two time periods. Results are presented fully on one arbitrarily chosen primary outcome (ECBI Intensity, see Subsection 4.8.1) and another arbitrarily chosen secondary outcome (BDI, see Subsection 4.8.2), and the rest are summarized in Subsection 4.8.3.

### 4.8.1 ECBI Intensity Score

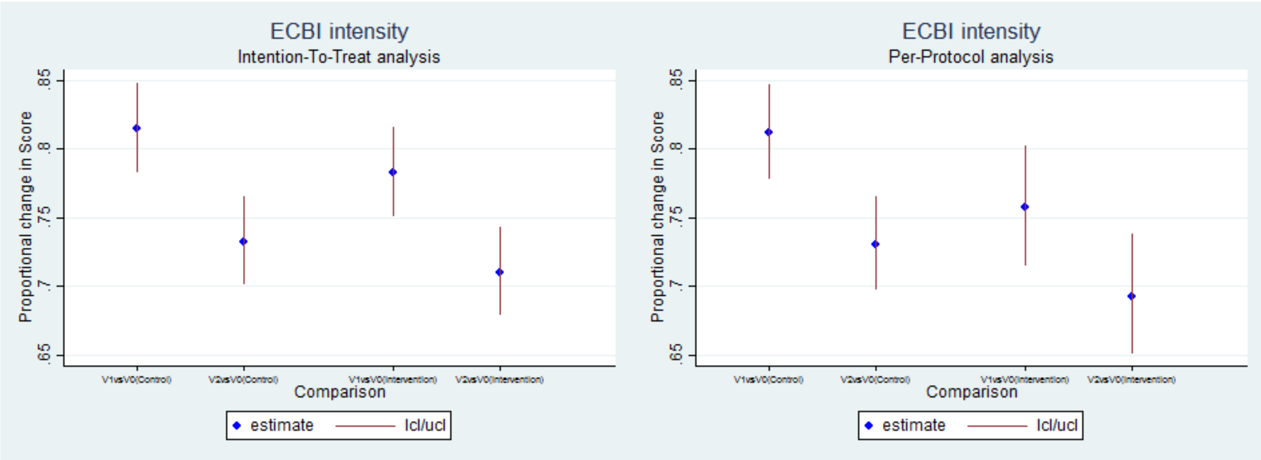
Table 4.3: Imputed Data

| | Intention to Treat | | | Per-Protocol | | |
|---|---|---|---|---|---|---|
| | Exp(Beta) | 95% CI | P-value | Exp(Beta) | 95% CI | P-value |
| (Intercept) | 140.0992 | ( 136.0069 - 144.3146 ) | < 0.0001 | 142.7116 | ( 137.9187 - 147.6711 ) | < 0.0001 |
| wave (Khayelitsha) | 1.0274 | ( 0.9993 - 1.0562 ) | 0.0561 | 1.0100 | ( 0.9778 - 1.0433 ) | 0.5469 |
| sex (Male) | 1.0171 | ( 0.9896 - 1.0454 ) | 0.2255 | 1.0120 | ( 0.9799 - 1.0452 ) | 0.4688 |
| age (6 - 9 yrs) | 0.9828 | ( 0.9559 - 1.0105 ) | 0.2215 | 0.9744 | ( 0.9433 - 1.0065 ) | 0.1164 |
| time1 (control) | 0.8150 | ( 0.7831 - 0.8482 ) | < 0.0001 | 0.8115 | ( 0.7777 - 0.8467 ) | < 0.0001 |
| time2 (control) | 0.7329 | ( 0.7014 - 0.7657 ) | < 0.0001 | 0.7303 | ( 0.6972 - 0.7649 ) | < 0.0001 |
| armtime1 | 0.9607 | ( 0.9137 - 1.0102 ) | 0.1178 | 0.9335 | ( 0.8758 - 0.9951 ) | 0.0347 |
| armtime2 | 0.9688 | ( 0.9164 - 1.0243 ) | 0.2647 | 0.9488 | ( 0.8846 - 1.0177 ) | 0.1418 |
| time1 (intervention) | 0.7830 | ( 0.7512 - 0.8161 ) | < 0.0001 | 0.7575 | ( 0.7149 - 0.8027 ) | < 0.0001 |
| time2 (intervention) | 0.7100 | ( 0.6787 - 0.7428 ) | < 0.0001 | 0.6929 | ( 0.6508 - 0.7378 ) | < 0.0001 |

Table 4.4: Unimputed Data

| | Intention to Treat | | | Per-Protocol | | |
|---|---|---|---|---|---|---|
| | Exp(Beta) | 95% CI | P-value | Exp(Beta) | 95% CI | P-value |
| (Intercept) | 140.2077 | ( 136.0155 - 144.5292 ) | < 0.0001 | 142.7965 | ( 137.9208 - 147.8445 ) | < 0.0001 |
| wave (Khayelitsha) | 1.0279 | ( 0.9988 - 1.0577 ) | 0.0600 | 1.0113 | ( 0.9781 - 1.0457 ) | 0.5065 |
| sex (Male) | 1.0165 | ( 0.9881 - 1.0457 ) | 0.2572 | 1.0113 | ( 0.9784 - 1.0454 ) | 0.5033 |
| age (6 - 9 yrs) | 0.9814 | ( 0.9537 - 1.0099 ) | 0.1970 | 0.9725 | ( 0.9407 - 1.0054 ) | 0.0996 |
| time1 (control) | 0.8150 | ( 0.7825 - 0.8488 ) | < 0.0001 | 0.8115 | ( 0.7772 - 0.8473 ) | < 0.0001 |
| time2 (control) | 0.7340 | ( 0.7013 - 0.7681 ) | < 0.0001 | 0.7314 | ( 0.6972 - 0.7672 ) | < 0.0001 |
| armtime1 | 0.9636 | ( 0.9152 - 1.0147 ) | 0.1592 | 0.9350 | ( 0.8761 - 0.9979 ) | 0.0431 |
| armtime2 | 0.9645 | ( 0.9092 - 1.0232 ) | 0.2296 | 0.9412 | ( 0.8736 - 1.0140 ) | 0.1106 |
| time1 (intervention) | 0.7854 | ( 0.7525 - 0.8197 ) | < 0.0001 | 0.7588 | ( 0.7152 - 0.8050 ) | < 0.0001 |
| time2 (intervention) | 0.7079 | ( 0.6752 - 0.7422 ) | < 0.0001 | 0.6883 | ( 0.6443 - 0.7354 ) | < 0.0001 |

Figure 4.1: Model estimates of the intervention effect on ECBI Intensity




As expected, tables 4.3 and 4.4 show that the differences between the imputed and unimputed datasets are not substantial. The effects based on the aforementioned datasets are not very different and their p-values are also very similar to each other. The models show no significant associations between the response: **ECBI Intensity Score** and covariates: **wave**, **sex** and **age**, their respective p-values for both ITT and PP models large. The **time** effects are all significantly different from 1 (the point of parity) suggesting that both visit 1 (post-test) and visit 2 (one year follow-up) scores were significantly different from the baseline (visit 0) scores. The **time** effects were all less one indicating that relative to baseline scores, both visit 1 and visit 2 **ECBI Intensity** scores were lower. The downward trend is also shown in the plots for the **time** effects and indicates the desirable movement of the scores as previously discussed. The fact that all time effects are significant indicates that the scores improved (in this case) in both arms during the study and the improvement was further maintained to the one year follow-up visit. The per-protocol sample shows a significantly greater decrease of the ECBI scores from baseline to post-test for the intervention arm as compared to the control arm [$\exp(\beta_{arm:time1}) = 0.9335; 95\%CI = (0.8758 - 0.9951)$ on the imputed data]. This effect is not picked up from baseline to one year follow-up. In other words, the model shows that there is greater intervention impact on child behaviour in those who attended seven or more intervention group sessions as compared to the controls but this effect faded with time.

### 4.8.2 Beck Depression Inventory Score

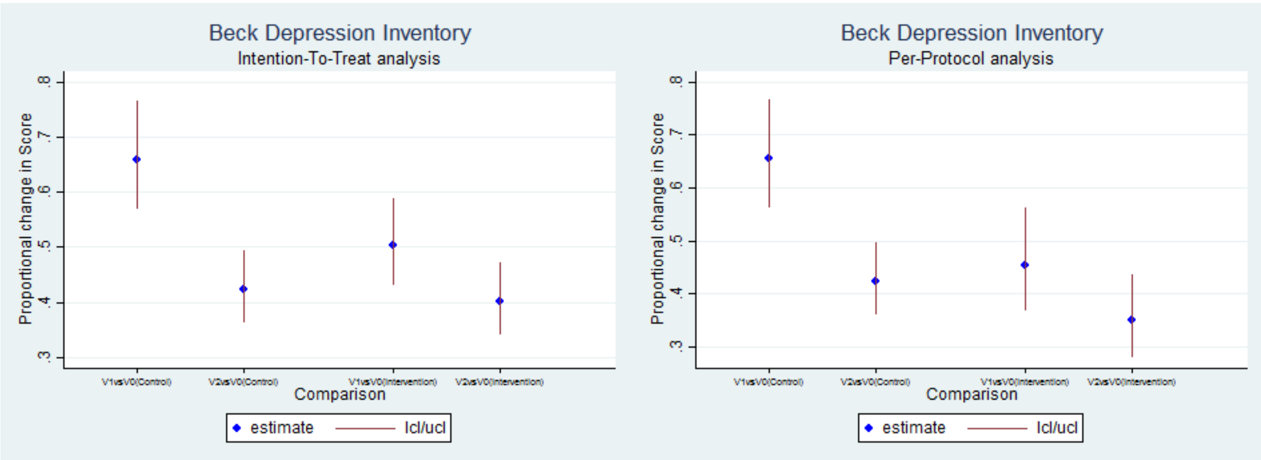
Table 4.5: (Imputed Data)

| | Intention to Treat | | | Per-Protocol | | |
|---|---|---|---|---|---|---|
| | Exp(Beta) | 95% CI | P-value | Exp(Beta) | 95% CI | P-value |
| (Intercept) | 12.3794 | ( 11.0160 - 13.9115 ) | < 0.0001 | 12.1423 | ( 10.6209 - 13.8816 ) | < 0.0001 |
| wave (Khayelitsha) | 1.1871 | ( 1.0694 - 1.3177 ) | 0.0013 | 1.1857 | ( 1.0532 - 1.3349 ) | 0.0048 |
| sex (Male) | 0.9938 | ( 0.8972 - 1.1007 ) | 0.9045 | 1.0350 | ( 0.9197 - 1.1648 ) | 0.5677 |
| age (6 - 9 yrs) | 1.0879 | ( 0.9804 - 1.2073 ) | 0.1125 | 1.1210 | ( 0.9952 - 1.2628 ) | 0.0600 |
| time1 (control) | 0.6597 | ( 0.5687 - 0.7653 ) | < 0.0001 | 0.6568 | ( 0.5625 - 0.7670 ) | < 0.0001 |
| time2 (control) | 0.4226 | ( 0.3619 - 0.4934 ) | < 0.0001 | 0.4234 | ( 0.3605 - 0.4973 ) | < 0.0001 |
| armtime1 | 0.7542 | ( 0.6313 - 0.9010 ) | 0.0019 | 0.6549 | ( 0.5275 - 0.8130 ) | 0.0001 |
| armtime2 | 0.9165 | ( 0.7597 - 1.1056 ) | 0.3621 | 0.7068 | ( 0.5500 - 0.9084 ) | 0.0067 |
| time1 (intervention) | 0.5038 | ( 0.4308 - 0.5892 ) | < 0.0001 | 0.4552 | ( 0.3688 - 0.5619 ) | < 0.0001 |
| time2 (intervention) | 0.4014 | ( 0.3408 - 0.4727 ) | < 0.0001 | 0.3496 | ( 0.2801 - 0.4364 ) | < 0.0001 |

Table 4.6: (Unimputed Data)

| | Intention to Treat | | | Per-Protocol | | |
|---|---|---|---|---|---|---|
| | Exp(Beta) | 95% CI | P-value | Exp(Beta) | 95% CI | P-value |
| (Intercept) | 12.0525 | ( 10.7206 - 13.5500 ) | < 0.0001 | 11.9923 | ( 10.5104 - 13.6832 ) | < 0.0001 |
| wave (Khayelitsha) | 1.1937 | ( 1.0750 - 1.3254 ) | 0.0010 | 1.1893 | ( 1.0555 - 1.3400 ) | 0.0046 |
| sex (Male) | 1.0060 | ( 0.9070 - 1.1158 ) | 0.9104 | 1.0552 | ( 0.9376 - 1.1877 ) | 0.3733 |
| age (6 - 9 yrs) | 1.1140 | ( 1.0035 - 1.2366 ) | 0.0432 | 1.1217 | ( 0.9965 - 1.2626 ) | 0.0579 |
| time1 (control) | 0.6410 | ( 0.5513 - 0.7452 ) | < 0.0001 | 0.6370 | ( 0.5447 - 0.7450 ) | < 0.0001 |
| time2 (control) | 0.4467 | ( 0.3824 - 0.5218 ) | < 0.0001 | 0.4455 | ( 0.3791 - 0.5234 ) | < 0.0001 |
| armtime1 | 0.7990 | ( 0.6675 - 0.9565 ) | 0.0148 | 0.6679 | ( 0.5369 - 0.8308 ) | 0.0003 |
| armtime2 | 0.9343 | ( 0.7752 - 1.1260 ) | 0.4756 | 0.6674 | ( 0.5292 - 0.8416 ) | 0.0007 |
| time1 (intervention) | 0.5182 | ( 0.4416 - 0.6080 ) | < 0.0001 | 0.4528 | ( 0.3650 - 0.5617 ) | < 0.0001 |
| time2 (intervention) | 0.4330 | ( 0.3683 - 0.5091 ) | < 0.0001 | 0.3700 | ( 0.2966 - 0.4617 ) | < 0.0001 |

Figure 4.2: Model estimates of the intervention effect on the BDI




The model results for the **BDI** also confirm that the imputed and unimputed data give similar results due to the low incidence of missingness discussed in Chapter 2. Tables 4.5 and 4.6 show no substantial differences in the effects and their p-values. In this model (especially in the PP analysis), **wave** shows a significant association with the **BDI** score, **age** is also potentially significant and **sex** is not significant. By construction, lower **BDI** scores are desirable and the above graphs and tables also show that the whole cohort made significant improvements in this score over the two time points compared to the baseline scores. Both the ITT and PP models show that from visit 0 to visit 1, the proportional changes in the **BDI** score were significantly larger in the intervention arm as compared to the control arm [$\exp(\beta_{arm:time1}) = 0.7542; 95\%CI = (0.6313 - 0.9010)$ ITT model on the imputed data]. This effect is not maintained on the ITT models but is maintained on the PP models. In other words, the intervention had a significant impact on the depression scores which improved (lowered) significantly between visit 0 and visit 1 but the effect was maintained on those that attended more intervention group sessions.

### 4.8.3 Summary of the rest of the models

Table 4.7 below shows the **arm*time** interaction effects that are meant to address the main hypothesis on whether the intervention program works. The full tables and plots for these models are included in Appendix A. As stated before, the programme would be deemed to work if the score improvements (from visit 0 to 1 and/ or visit 0 to 2) for participants in the intervention arm were significantly bigger than those in the control arm. From the table, the participants in the two arms weren't different with regards to the proportional score changes from visit 0 to visit 1 on a total of six scores namely (1) **Non-Violent Discipline**, (2) **Poor Monitoring and Supervision**, (3) **Parent-Child Dysfunctional Interaction**, (4) **Social Support**, (5) **Parent Negative Behaviour** and (6) **Child Negative Behaviour**. For these scores, the models show that the two arms were not significantly different from each other.

The intervention arm had significantly larger score improvements compared to the control arm from visit 0 to visit 1 on the following outcomes: (a) ECBI Problem score, (b) all frequency scores that form part of **Positive Parenting** scores, (c) **Physical and Psychological Discipline** scores, (d) **Parenting Distress** and (e) the video coded **positive parent and child behaviour** scores. For each of these, the intervention arm showed more significantly more desirable movements, i.e. more positive changes for scores where positive changes were desirable and vice versa. Of these, only **Parenting Distress** score (reversed) and the video coded **positive parent and child behaviour** scores show a sustained effect up to visit 2. The rest of the scores (the problem scores on **Positive Parenting** and the remaining **Parenting Stress** scores) show an unexpected larger score worsening from visit 0 to visit 1 in the intervention arm as compared to the control arm.

Table 4.7: Model Summaries (Imputed Data)

| | Intention to Treat | | | Per-Protocol | | |
|---|---|---|---|---|---|---|
| | Exp(Beta) | 95% CI | P-value | Exp(Beta) | 95% CI | P-value |
| ECBI Problem Score | | | | | | |
| armtime1 | 1.0189 | ( 0.8983 - 1.1557 ) | 0.7708 | 0.8502 | ( 0.7225 - 1.0005 ) | 0.0507 |
| armtime2 | 1.0407 | ( 0.9013 - 1.2015 ) | 0.5868 | 0.9651 | ( 0.8129 - 1.1460 ) | 0.6855 |
| Supporting Positive Behaviour Frequency Score | | | | | | |
| armtime1 | 1.0825 | ( 1.0361 - 1.1309 ) | 0.0004 | 1.1376 | ( 1.0812 - 1.1969 ) | $< 0.0001$ |
| armtime2 | 1.0034 | ( 0.9604 - 1.0484 ) | 0.8790 | 1.0221 | ( 0.9699 - 1.0772 ) | 0.4135 |
| Setting Limits Frequency Reverse Score | | | | | | |
| armtime1 | 0.8872 | ( 0.8435 - 0.9332 ) | $< 0.0001$ | 0.8376 | ( 0.7891 - 0.8890 ) | $< 0.0001$ |
| armtime2 | 1.0072 | ( 0.9570 - 1.0600 ) | 0.7832 | 1.0320 | ( 0.9713 - 1.0965 ) | 0.3086 |
| Positive Parenting Frequency Score | | | | | | |
| armtime1 | 1.1016 | ( 1.0547 - 1.1506 ) | $< 0.0001$ | 1.1591 | ( 1.1022 - 1.2190 ) | $< 0.0001$ |
| armtime2 | 0.9927 | ( 0.9511 - 1.0360 ) | 0.7354 | 0.9923 | ( 0.9418 - 1.0454 ) | 0.7713 |
| Supporting Positive Behaviour Problem Score | | | | | | |
| armtime1 | 0.9283 | ( 0.7852 - 1.0974 ) | 0.3836 | 0.5478 | ( 0.4278 - 0.7013 ) | $< 0.0001$ |
| armtime2 | 1.1387 | ( 0.9543 - 1.3587 ) | 0.1495 | 1.1364 | ( 0.9115 - 1.4168 ) | 0.2557 |
| Setting Limits Problem Score | | | | | | |
| armtime1 | 0.9967 | ( 0.7847 - 1.2659 ) | 0.9782 | 0.6348 | ( 0.4579 - 0.8801 ) | 0.0064 |
| armtime2 | 1.1682 | ( 0.8960 - 1.5231 ) | 0.2507 | 1.3092 | ( 0.9497 - 1.8048 ) | 0.1000 |
| Positive Parenting Problem Score | | | | | | |
| armtime1 | 0.9121 | ( 0.7360 - 1.1302 ) | 0.4003 | 0.5660 | ( 0.4247 - 0.7544 ) | 0.0001 |
| armtime2 | 1.1237 | ( 0.8951 - 1.4106 ) | 0.3150 | 1.2276 | ( 0.9283 - 1.6233 ) | 0.1503 |
| Non-Violent Discipline Score | | | | | | |
| armtime1 | 1.0360 | ( 0.9377 - 1.1446 ) | 0.4872 | 1.0338 | ( 0.9164 - 1.1663 ) | 0.5890 |
| armtime2 | 1.1533 | ( 1.0336 - 1.2867 ) | 0.0107 | 1.1366 | ( 0.9981 - 1.2942 ) | 0.0534 |
| Physical Discipline Score | | | | | | |
| armtime1 | 0.6772 | ( 0.5430 - 0.8446 ) | 0.0005 | 0.5856 | ( 0.4446 - 0.7714 ) | 0.0001 |
| armtime2 | 0.9530 | ( 0.7379 - 1.2308 ) | 0.7121 | 0.8293 | ( 0.6018 - 1.1427 ) | 0.2525 |
| Psychological Discipline Score | | | | | | |
| armtime1 | 0.8440 | ( 0.7161 - 0.9948 ) | 0.0432 | 0.7787 | ( 0.6340 - 0.9564 ) | 0.0171 |
| armtime2 | 1.0263 | ( 0.8492 - 1.2402 ) | 0.7885 | 0.9375 | ( 0.7389 - 1.1895 ) | 0.5952 |
| Poor Monitoring And Supervision Score | | | | | | |
| armtime1 | 1.0359 | ( 0.9858 - 1.0885 ) | 0.1635 | 1.0178 | ( 0.9586 - 1.0806 ) | 0.5644 |
| armtime2 | 1.0071 | ( 0.9559 - 1.0611 ) | 0.7895 | 0.9788 | ( 0.9183 - 1.0434 ) | 0.5115 |
| Parenting Distress Reverse Score | | | | | | |
| armtime1 | 0.9001 | ( 0.8504 - 0.9526 ) | 0.0003 | 0.8822 | ( 0.8238 - 0.9448 ) | 0.0003 |
| armtime2 | 0.9504 | ( 0.8951 - 1.0090 ) | 0.0958 | 0.9204 | ( 0.8559 - 0.9898 ) | 0.0253 |
| Parent-Child Dysfunctional Interaction Reverse Score | | | | | | |
| armtime1 | 0.9444 | ( 0.8729 - 1.0218 ) | 0.1546 | 0.9761 | ( 0.8892 - 1.0715 ) | 0.6109 |
| armtime2 | 0.9292 | ( 0.8554 - 1.0093 ) | 0.0818 | 0.9135 | ( 0.8270 - 1.0090 ) | 0.0745 |

| | Intention to Treat | | | Per-Protocol | | |
|---|---|---|---|---|---|---|
| | Exp(Beta) | 95% CI | P-value | Exp(Beta) | 95% CI | P-value |
| Difficult Child Reverse Score | | | | | | |
| armtime1 | 0.9389 | ( 0.8804 - 1.0012 ) | 0.0543 | 0.9173 | ( 0.8486 - 0.9915 ) | 0.0296 |
| armtime2 | 0.9691 | ( 0.9055 - 1.0373 ) | 0.3659 | 0.9302 | ( 0.8564 - 1.0103 ) | 0.0861 |
| Parenting Stress Reverse Score | | | | | | |
| armtime1 | 0.9226 | ( 0.8778 - 0.9697 ) | 0.0015 | 0.9123 | ( 0.8595 - 0.9684 ) | 0.0026 |
| armtime2 | 0.9538 | ( 0.9062 - 1.0039 ) | 0.0702 | 0.9257 | ( 0.8704 - 0.9845 ) | 0.0140 |
| Social Support Score | | | | | | |
| armtime1 | 0.9974 | ( 0.9305 - 1.0692 ) | 0.9418 | 0.9639 | ( 0.8831 - 1.0520 ) | 0.4099 |
| armtime2 | 1.0101 | ( 0.9383 - 1.0874 ) | 0.7895 | 0.9782 | ( 0.8916 - 1.0732 ) | 0.6413 |
| Parent Positive Behaviour Score | | | | | | |
| armtime1 | 1.3185 | ( 1.2033 - 1.4448 ) | $< 0.0001$ | 1.4368 | ( 1.2949 - 1.5943 ) | $< 0.0001$ |
| armtime2 | 1.2042 | ( 1.0903 - 1.3299 ) | 0.0002 | 1.3448 | ( 1.1354 - 1.5929 ) | 0.0006 |
| Parent Negative Behaviour Score | | | | | | |
| armtime1 | 0.9691 | ( 0.8331 - 1.1273 ) | 0.6839 | 1.0084 | ( 0.8405 - 1.2099 ) | 0.9281 |
| armtime2 | 0.7519 | ( 0.6035 - 0.9369 ) | 0.0111 | 0.8273 | ( 0.6280 - 1.0899 ) | 0.1778 |
| Child Positive Behaviour Score | | | | | | |
| armtime1 | 1.0644 | ( 0.9717 - 1.1660 ) | 0.1796 | 1.2623 | ( 1.1496 - 1.3859 ) | $< 0.0001$ |
| armtime2 | 1.1254 | ( 1.0410 - 1.2166 ) | 0.0030 | 1.1079 | ( 1.0147 - 1.2096 ) | 0.0222 |
| Child Negative Behaviour Score | | | | | | |
| armtime1 | 0.9910 | ( 0.8084 - 1.2149 ) | 0.9310 | 0.9069 | ( 0.7167 - 1.1476 ) | 0.4159 |
| armtime2 | 1.0719 | ( 0.7611 - 1.5095 ) | 0.6911 | 1.0718 | ( 0.7265 - 1.5811 ) | 0.7267 |

## 4.9 Results for the Dose-Response Models

This section presents the model results for the **ten** responses that were selected for the dose-response models. Just as in the previous section, the tables shown here summarise the relative effects of the different covariates selected to model the outcomes of interest. An important difference from the previous models is that now attendance is modelled as a continuous (numeric) variable (indicated as **att** in the tables) rather than a binary one. As stated in Section 4.5.2, this assumes linearity in the score change for every additional group-session attended. In addition, the dose-response models also include three-way interactions between (1) **wave**, **attendance** and **time** and (2) **ipv** , **attendance** and **time**. The **wave - attendance - time** interaction terms are included to test whether the relative changes in the scores with respect to attendance differ between the two waves over the two time periods (i.e. baseline to visit1 and baseline to visit 2). Similarly, the **ipv - attendance - time** interaction terms are included to test whether the relative changes in the scores with respect to attendance differ between the participants who reported baseline ipv and those who didn't over the two time periods (i.e. baseline to visit1 and baseline to visit 2).

The plots given for each model in this section present the four main strata based on the wave-ipv level combinations. The different slopes of the lines in these plots illustrate the three way interactions discussed above. The baseline stratum is a group of participants in Nyanga and had no reported baseline ipv **(wave0, ipv0)**. The other strata include ipv-exposed group in Nyanga **(wave0, ipv1)**, those who reported no baseline ipv group in Khayelitsha **(wave1, ipv0)** and the ipv-exposed group in Khayelitsha **(wave1, ipv1)**. These plots show the point estimates and point confidence intervals (sometimes perturbed slightly in order to be clear) of the proportional changes expected in the scores from baseline to visit 1 and again from baseline to visit 2. The aforementioned point estimates and confidence intervals are plotted over the (integer) values of attendance, linearity is assumed in-between these integer values. All these values are estimated within the multiple imputation framework as explained in Section 4.6 above. As a general rule, if the plots of the profiles are not parallel then there is some indication of some of the aforementioned three-way interaction effects.

### 4.9.1 ECBI Intensity Score

Table 4.8: Dose-Response Model Results

| | Imputed Data | | | Unimputed Data | | |
|---|---|---|---|---|---|---|
| | Exp(Beta) | 95% CI | P-value | Exp(Beta) | 95% CI | P-value |
| (Intercept) | 141.8323 | ( 136.9798 - 146.8567 ) | < 0.0001 | 141.8846 | ( 136.9316 - 147.0166 ) | < 0.0001 |
| wave1 | 0.9867 | ( 0.9479 - 1.0270 ) | 0.5114 | 0.9868 | ( 0.9473 - 1.0279 ) | 0.5220 |
| time1 (control) | 0.7485 | ( 0.7072 - 0.7923 ) | < 0.0001 | 0.7492 | ( 0.7069 - 0.7939 ) | < 0.0001 |
| time2 (control) | 0.7109 | ( 0.6699 - 0.7544 ) | < 0.0001 | 0.7115 | ( 0.6683 - 0.7574 ) | < 0.0001 |
| ipv1 | 1.0229 | ( 0.9801 - 1.0675 ) | 0.2987 | 1.0229 | ( 0.9793 - 1.0685 ) | 0.3065 |
| sex (Male) | 1.0167 | ( 0.9895 - 1.0448 ) | 0.2316 | 1.0164 | ( 0.9881 - 1.0454 ) | 0.2576 |
| age (6 - 9 yrs) | 0.9825 | ( 0.9558 - 1.0100 ) | 0.2096 | 0.9820 | ( 0.9545 - 1.0104 ) | 0.2104 |
| att.time1 | 1.0045 | ( 0.9949 - 1.0142 ) | 0.3612 | 1.0048 | ( 0.9949 - 1.0148 ) | 0.3401 |
| att.time2 | 0.9997 | ( 0.9893 - 1.0102 ) | 0.9513 | 0.9991 | ( 0.9883 - 1.0101 ) | 0.8771 |
| wave1:time1 | 1.1524 | ( 1.0723 - 1.2386 ) | 0.0004 | 1.1534 | ( 1.0715 - 1.2414 ) | 0.0002 |
| wave1:time2 | 1.0385 | ( 0.9604 - 1.1229 ) | 0.344 | 1.0398 | ( 0.9582 - 1.1283 ) | 0.3488 |
| time1:ipv1 | 1.0504 | ( 0.9717 - 1.1354 ) | 0.2162 | 1.0518 | ( 0.9712 - 1.1391 ) | 0.2137 |
| time2:ipv1 | 1.0316 | ( 0.9470 - 1.1238 ) | 0.4754 | 1.0325 | ( 0.9445 - 1.1286 ) | 0.4810 |
| wave1:att.time1 | 0.9872 | ( 0.9755 - 0.9990 ) | 0.0341 | 0.9871 | ( 0.9751 - 0.9992 ) | 0.0375 |
| ipv1:att.time1 | 0.9860 | ( 0.9737 - 0.9984 ) | 0.0275 | 0.9854 | ( 0.9728 - 0.9982 ) | 0.0257 |
| wave1:att.time2 | 0.9985 | ( 0.9855 - 1.0117 ) | 0.8252 | 0.9989 | ( 0.9853 - 1.0127 ) | 0.8760 |
| ipv1:att.time2 | 0.9927 | ( 0.9793 - 1.0064 ) | 0.2967 | 0.9916 | ( 0.9772 - 1.0062 ) | 0.2556 |

Figure 4.3: ECBI Intensity change from baseline to visit 1

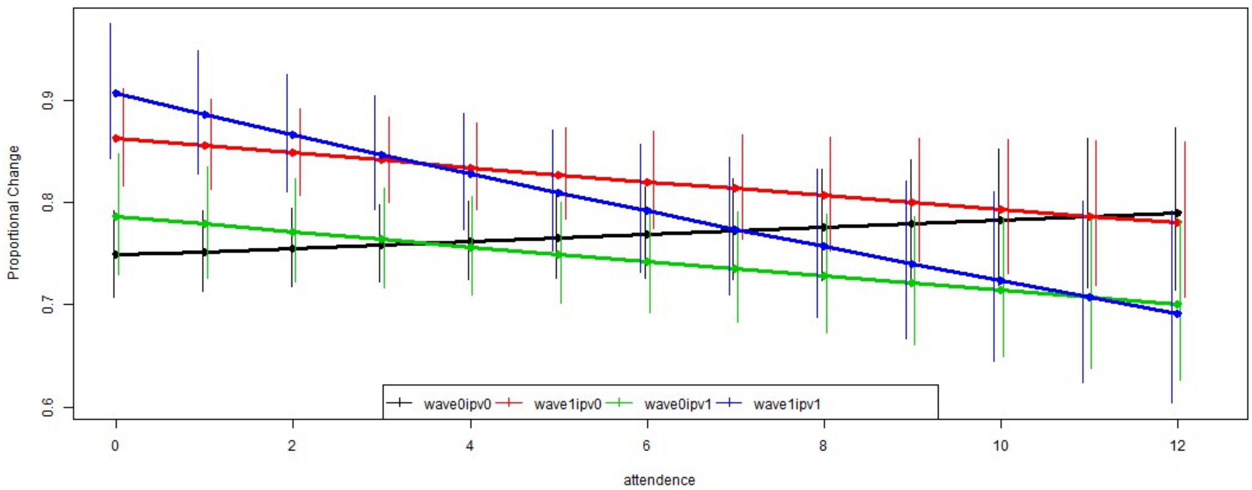
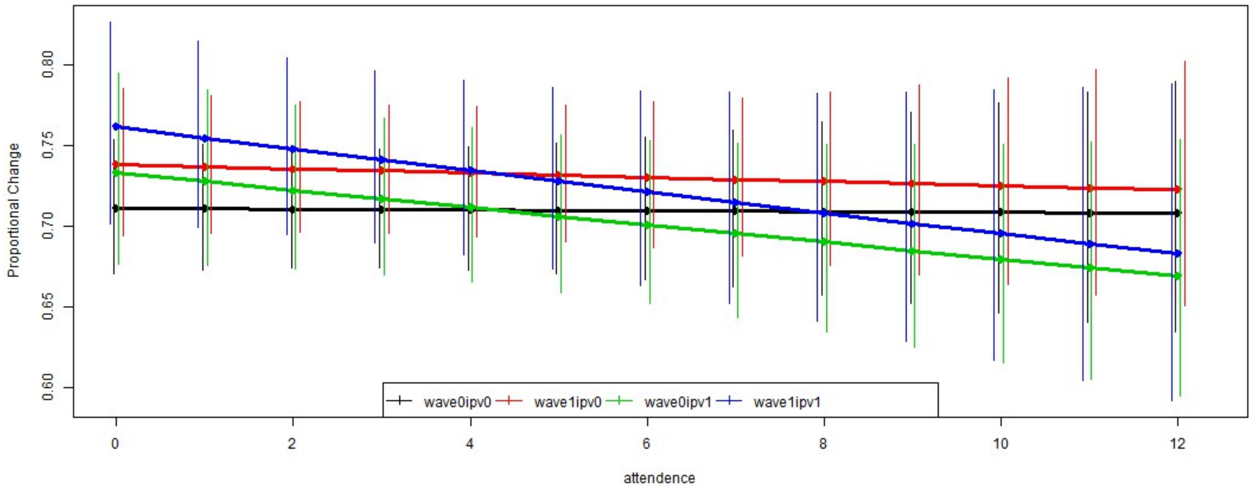


Figure 4.4: ECBI Intensity change from baseline to visit 2




As previously discussed, lower **ECBI intensity** scores are more desirable and as such, one would expect relatively lower scores for participants who attended more visits. At visit 1 (post-test), the model summary table above shows that those in Khayelitsha had significantly better (i.e. lower) **ECBI intensity** scores per session, compared to those in Nyanga $[\exp(\beta_{wave1:att:time1}) = 0.9872; 95\%CI = (0.9755 - 0.9990)]$. Also considering the same time-point, the model also shows significantly lower scores per session for the ipv-exposed group compared to the non-exposed group $[\exp(\beta_{ipv1:att:time1}) = 0.9860; 95\%CI = (0.9737 - 0.9984)]$. In other words, this model shows that the intervention was able to have a greater relative impact per session amongst those that were in Khayelitsha and those that had baseline ipv-exposure. However, these effects are not realized at visit 2.

### 4.9.2 ECBI Problem Score

Table 4.9: Dose-Response Model Results

| | Imputed Data | | | Unimputed Data | | |
|---|---|---|---|---|---|---|
| | Exp(Beta) | 95% CI | P-value | Exp(Beta) | 95% CI | P-value |
| (Intercept) | 24.4074 | ( 22.4582 - 26.5259 ) | < 0.0001 | 24.4027 | ( 22.9484 - 25.9491 ) | < 0.0001 |
| wave1 | 1.0026 | ( 0.9129 - 1.1011 ) | 0.9568 | 1.0009 | ( 0.9349 - 1.0716 ) | 0.9799 |
| time1 (control) | 0.7142 | ( 0.6226 - 0.8192 ) | < 0.0001 | 0.6761 | ( 0.6089 - 0.7507 ) | < 0.0001 |
| time2 (control) | 0.5990 | ( 0.5110 - 0.7021 ) | < 0.0001 | 0.4733 | ( 0.4097 - 0.5468 ) | < 0.0001 |
| ipv1 | 1.0078 | ( 0.9113 - 1.1146 ) | 0.8796 | 1.0068 | ( 0.9357 - 1.0833 ) | 0.8550 |
| sex (Male) | 1.0625 | ( 0.9934 - 1.1364 ) | 0.0773 | 1.0527 | ( 0.9991 - 1.1091 ) | 0.0538 |
| age (6 - 9 yrs) | 0.9593 | ( 0.8962 - 1.0268 ) | 0.2309 | 0.9741 | ( 0.9242 - 1.0267 ) | 0.3268 |
| att.time1 | 1.0078 | ( 0.9842 - 1.0319 ) | 0.5210 | 0.9947 | ( 0.9749 - 1.0149 ) | 0.6062 |
| att.time2 | 0.9988 | ( 0.9705 - 1.0279 ) | 0.9338 | 0.9964 | ( 0.9695 - 1.0242 ) | 0.7988 |
| wave1:time1 | 1.1022 | ( 0.9250 - 1.3133 ) | 0.2766 | 1.1075 | ( 0.9699 - 1.2647 ) | 0.1310 |
| wave1:time2 | 1.2250 | ( 1.0046 - 1.4937 ) | 0.0449 | 1.2257 | ( 1.0264 - 1.4637 ) | 0.0247 |
| time1:ipv1 | 1.0974 | ( 0.9080 - 1.3263 ) | 0.3362 | 1.1172 | ( 0.9686 - 1.2886 ) | 0.1276 |
| time2:ipv1 | 0.9655 | ( 0.7746 - 1.2034 ) | 0.7547 | 1.0512 | ( 0.8678 - 1.2733 ) | 0.6093 |
| wave1:att.time1 | 0.9748 | ( 0.9442 - 1.0063 ) | 0.1161 | 0.9859 | ( 0.9609 - 1.0116 ) | 0.2787 |
| ipv1:att.time1 | 0.9636 | ( 0.9308 - 0.9976 ) | 0.0359 | 0.9664 | ( 0.9397 - 0.9940 ) | 0.0173 |
| wave1:att.time2 | 0.9826 | ( 0.9495 - 1.0169 ) | 0.3159 | 0.9980 | ( 0.9661 - 1.0309 ) | 0.9019 |
| ipv1:att.time2 | 1.0163 | ( 0.9812 - 1.0527 ) | 0.3671 | 0.9914 | ( 0.9577 - 1.0262 ) | 0.6225 |

Figure 4.5: ECBI Problem change from baseline to visit 1

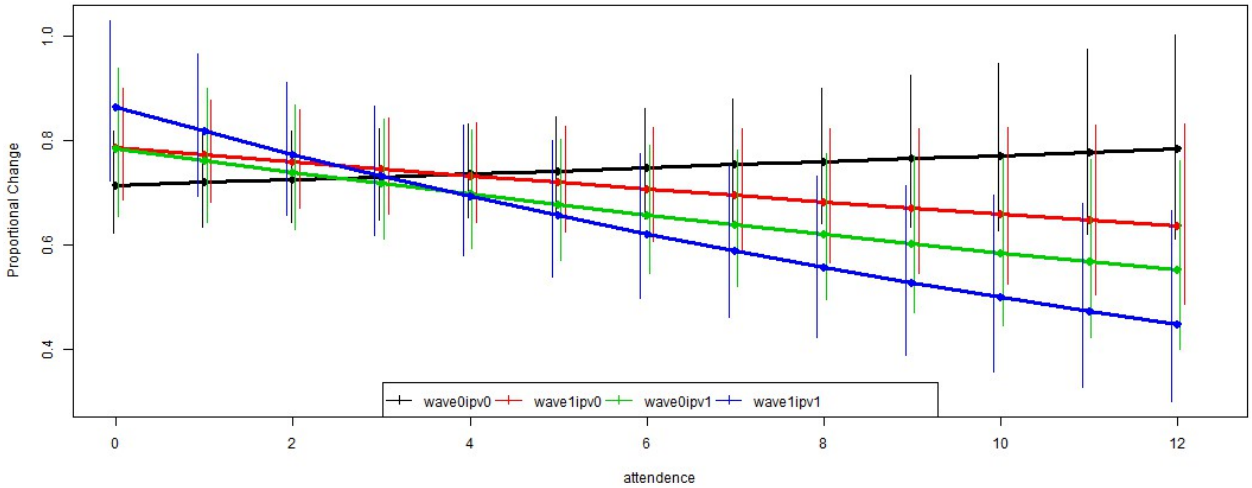
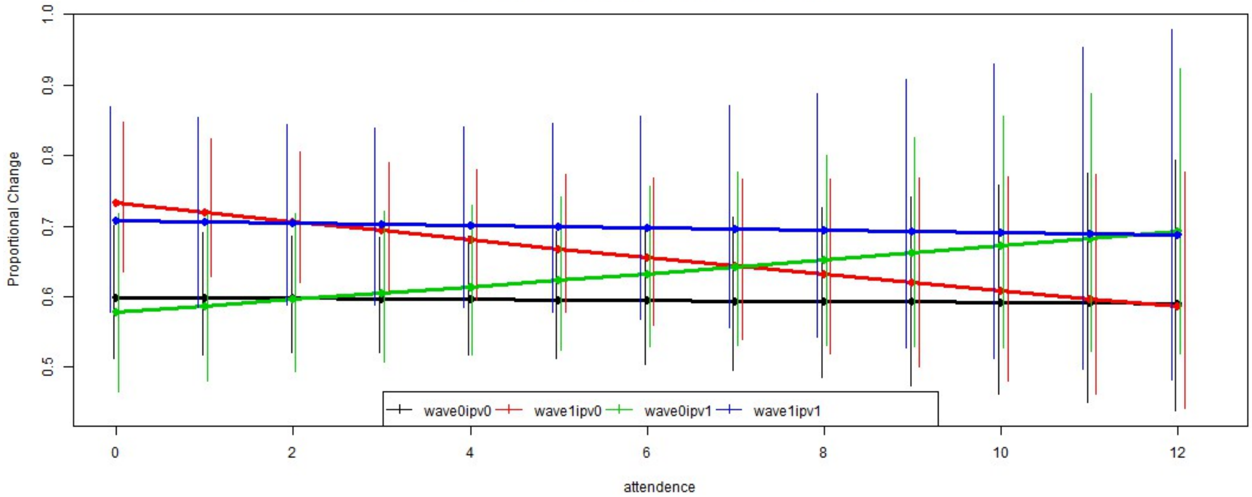



Figure 4.6: ECBI Problem change from baseline to visit 2




The **ECBI problem** score was also expected to have a negative relationship with attendance by construction. From the above summary table, it seems that at visit 1, the ipv-exposed group has significantly lower **ECBI problem** scores per session compared to the non-exposed group [$\exp(\beta_{ipv1:att:time1}) = 0.9636; 95\%CI = (0.9308 - 0.9976)$], this effect loses significance when considering visit 2 (one year follow-up). The table also shows that at visit 2 (one year follow-up), **ECBI problem** scores in wave 1 (Khayelitsha) were significantly higher than those in wave 2 (Nyanga) [$\exp(\beta_{wave1:time2}) = 1.2250; 95\%CI = (1.0046 - 1.4937)$].

### 4.9.3 Positive Parenting Frequency Score

Table 4.10: Dose-Response Model Results

| | Imputed Data | | | Unimputed Data | | |
|---|---|---|---|---|---|---|
| | Exp(Beta) | 95% CI | P-value | Exp(Beta) | 95% CI | P-value |
| (Intercept) | 46.6185 | ( 44.7992 - 48.5116 ) | < 0.0001 | 46.6478 | ( 44.7923 - 48.5802 ) | < 0.0001 |
| wave1 | 1.0714 | ( 1.0227 - 1.1223 ) | 0.0036 | 1.0713 | ( 1.0217 - 1.1233 ) | 0.0045 |
| time1 (control) | 1.0879 | ( 1.0301 - 1.1490 ) | 0.0025 | 1.0857 | ( 1.0267 - 1.1482 ) | 0.0040 |
| time2 (control) | 1.1448 | ( 1.0853 - 1.2075 ) | < 0.0001 | 1.1449 | ( 1.0832 - 1.2100 ) | < 0.0001 |
| ipv1 | 0.9910 | ( 0.9422 - 1.0424 ) | 0.7267 | 0.9910 | ( 0.9412 - 1.0433 ) | 0.7288 |
| sex (Male) | 0.9786 | ( 0.9540 - 1.0038 ) | 0.0960 | 0.9765 | ( 0.9511 - 1.0026 ) | 0.0772 |
| age (6 - 9 yrs) | 1.0429 | ( 1.0164 - 1.0700 ) | 0.0014 | 1.0441 | ( 1.0165 - 1.0723 ) | 0.0016 |
| att.time1 | 1.0076 | ( 0.9999 - 1.0154 ) | 0.0543 | 1.0082 | ( 1.0002 - 1.0162 ) | 0.0451 |
| att.time2 | 0.9975 | ( 0.9896 - 1.0055 ) | 0.5358 | 0.9974 | ( 0.9892 - 1.0057 ) | 0.5386 |
| wave1:time1 | 0.9040 | ( 0.8414 - 0.9712 ) | 0.0058 | 0.9041 | ( 0.8400 - 0.9732 ) | 0.0074 |
| wave1:time2 | 0.9374 | ( 0.8743 - 1.0050 ) | 0.0690 | 0.9421 | ( 0.8766 - 1.0126 ) | 0.1053 |
| time1:ipv1 | 0.9858 | ( 0.9100 - 1.0680 ) | 0.7269 | 0.9862 | ( 0.9085 - 1.0705 ) | 0.7396 |
| time2:ipv1 | 0.9688 | ( 0.8959 - 1.0477 ) | 0.4277 | 0.9658 | ( 0.8907 - 1.0474 ) | 0.3999 |
| wave1:att.time1 | 1.0082 | ( 0.9987 - 1.0178 ) | 0.0912 | 1.0080 | ( 0.9982 - 1.0178 ) | 0.1108 |
| ipv1:att.time1 | 1.0065 | ( 0.9965 - 1.0167 ) | 0.2032 | 1.0063 | ( 0.9959 - 1.0167 ) | 0.2364 |
| wave1:att.time2 | 1.0061 | ( 0.9962 - 1.0162 ) | 0.2277 | 1.0064 | ( 0.9960 - 1.0169 ) | 0.2252 |
| ipv1:att.time2 | 0.9970 | ( 0.9864 - 1.0076 ) | 0.5760 | 0.9960 | ( 0.9848 - 1.0074 ) | 0.4936 |

Figure 4.7: Positive Parenting Frequency change from baseline to visit 1

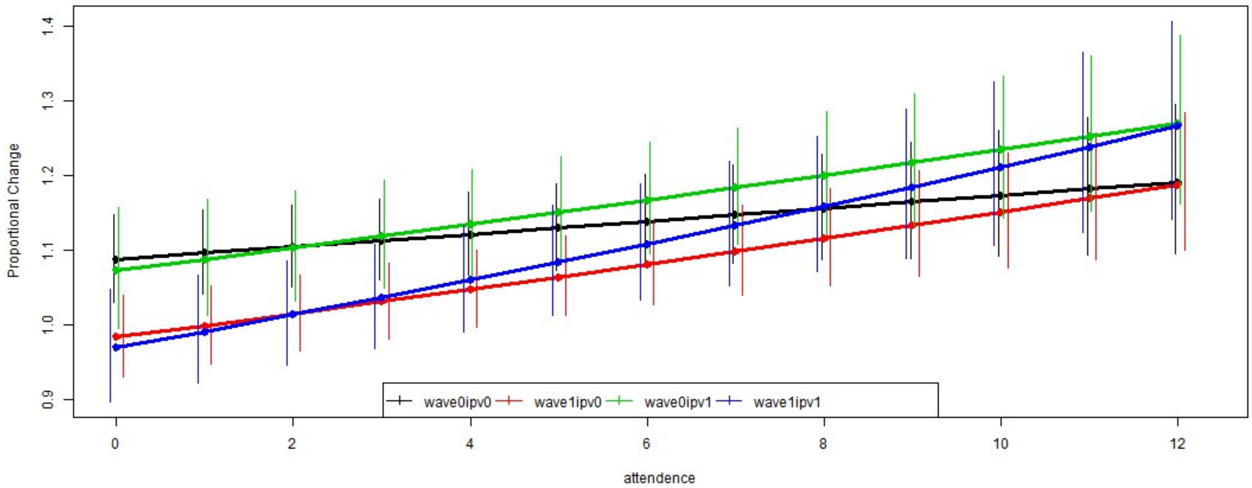
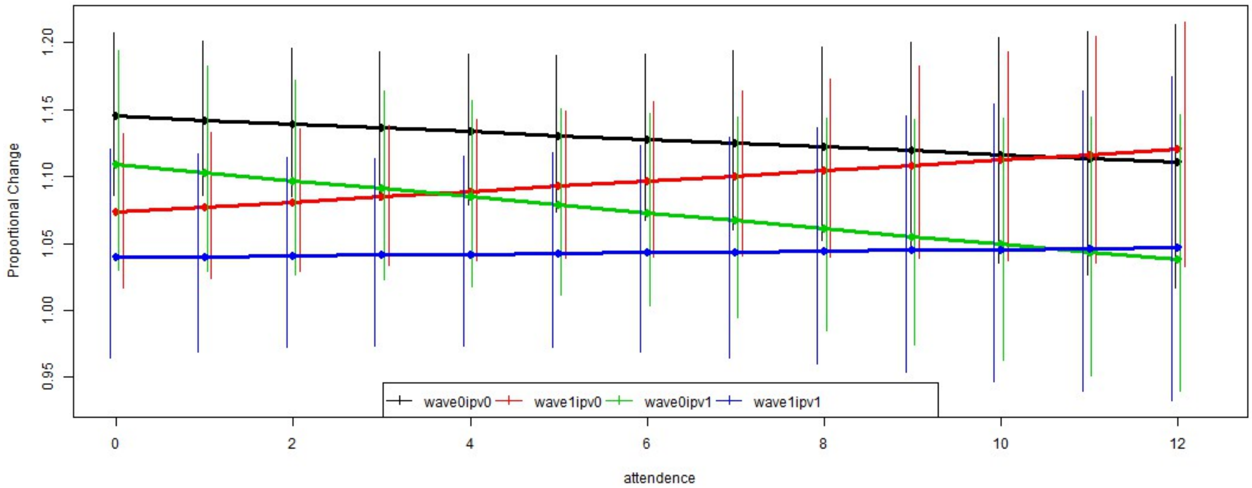


Figure 4.8: Positive Parenting Frequency change from baseline to visit 2




The **positive parenting frequency** score is summarized in the above table and plots. By construction, this score is expected to increase with attendance a trend shown in the first plot $[\exp(\beta_{att:time1}) = 1.0076; 95\%CI = (0.9999 - 1.0154); p - value = 0.0543]$ but very unclear in the second time interval. Between both time periods under consideration (i.e. baseline-to-visit1 and baseline-to-visit2) participants in Khayelitsha seem to have significantly lower relative score changes compared to those in Nyanga i.e. $[\exp(\beta_{wave1:time1}) = 0.9040; 95\%CI = (0.8414 - 0.9984)]$ and $[\exp(\beta_{wave1:time2}) = 0.9374; 95\%CI = (0.8743 - 1.0050)]$.

### 4.9.4 Positive Parenting Problem Score

Table 4.11: Dose-Response Model Results

| | Imputed Data | | | Unimputed Data | | |
|---|---|---|---|---|---|---|
| | Exp(Beta) | 95% CI | P-value | Exp(Beta) | 95% CI | P-value |
| (Intercept) | 4.8903 | ( 4.1528 - 5.7589 ) | < 0.0001 | 4.5449 | ( 4.0618 - 5.0854 ) | < 0.0001 |
| wave1 | 0.5406 | ( 0.4425 - 0.6604 ) | < 0.0001 | 0.5595 | ( 0.4873 - 0.6425 ) | < 0.0001 |
| time1 (control) | 0.5505 | ( 0.4341 - 0.6981 ) | < 0.0001 | 0.5060 | ( 0.4272 - 0.5993 ) | < 0.0001 |
| time2 (control) | 0.3756 | ( 0.2921 - 0.4828 ) | < 0.0001 | 0.2717 | ( 0.2209 - 0.3342 ) | < 0.0001 |
| ipv1 | 0.8711 | ( 0.7030 - 1.0795 ) | 0.2074 | 0.8095 | ( 0.6973 - 0.9396 ) | 0.0056 |
| sex (Male) | 1.3841 | ( 1.2259 - 1.5628 ) | < 0.0001 | 1.4030 | ( 1.2794 - 1.5387 ) | < 0.0001 |
| age (6 - 9 yrs) | 0.9357 | ( 0.8281 - 1.0573 ) | 0.2864 | 1.0146 | ( 0.9253 - 1.1126 ) | 0.7577 |
| att.time1 | 0.9887 | ( 0.9494 - 1.0296 ) | 0.5835 | 0.9581 | ( 0.9259 - 0.9915 ) | 0.0146 |
| att.time2 | 0.9868 | ( 0.9451 - 1.0302 ) | 0.5442 | 0.9698 | ( 0.9308 - 1.0104 ) | 0.1433 |
| wave1:time1 | 2.4621 | ( 1.7920 - 3.3827 ) | < 0.0001 | 2.4939 | ( 1.9862 - 3.1315 ) | < 0.0001 |
| wave1:time2 | 2.3207 | ( 1.6595 - 3.2455 ) | < 0.0001 | 1.9499 | ( 1.4748 - 2.5780 ) | < 0.0001 |
| time1:ipv1 | 1.1109 | ( 0.7804 - 1.5815 ) | 0.5594 | 1.0997 | ( 0.8492 - 1.4241 ) | 0.4715 |
| time2:ipv1 | 1.0379 | ( 0.7122 - 1.5124 ) | 0.8466 | 1.2809 | ( 0.9386 - 1.7481 ) | 0.1192 |
| wave1:att.time1 | 0.9716 | ( 0.9233 - 1.0224 ) | 0.2681 | 1.0017 | ( 0.9613 - 1.0438 ) | 0.9354 |
| ipv1:att.time1 | 0.9360 | ( 0.8857 - 0.9892 ) | 0.0189 | 0.9535 | ( 0.9105 - 0.9985 ) | 0.0436 |
| wave1:att.time2 | 1.0154 | ( 0.9648 - 1.0686 ) | 0.5578 | 1.0680 | ( 1.0186 - 1.1197 ) | 0.0067 |
| ipv1:att.time2 | 1.0433 | ( 0.9882 - 1.1016 ) | 0.1259 | 0.9899 | ( 0.9420 - 1.0403 ) | 0.6886 |

Figure 4.9: Positive Parenting Problem change from baseline to visit 1

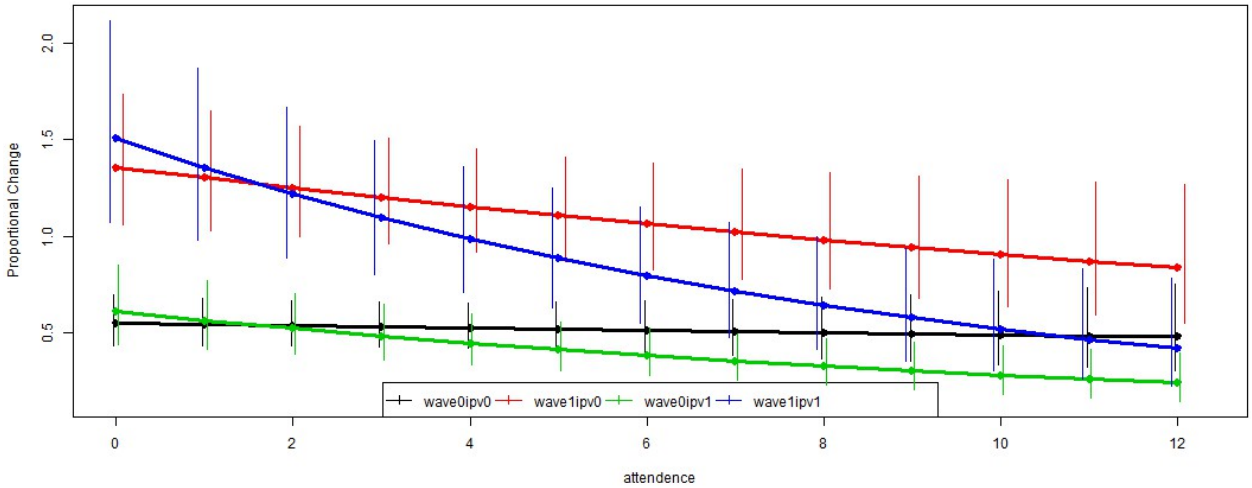
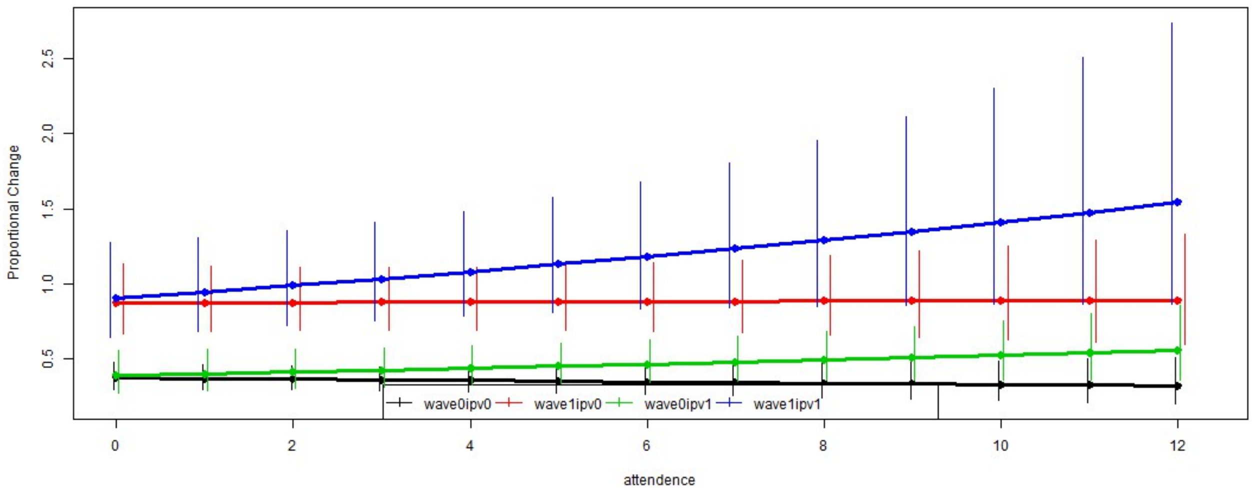



Figure 4.10: Positive Parenting Problem change from baseline to visit 2




The above table and plots summarize the results for the **positive parenting problem** score. The profiles shown in the first plot (baseline-to-visit 1) show an unexpected negative relation between attendance and the score but this relationship is as expected in the baseline-to-visit 2 time period. The first plot also shows some crossing of the profiles, specifically the more negative trending profiles of groups with some ipv tend to cross the flatter profiles from the no-ipv groups. Together with the model results in the table, one can conclude that compared to the no-ipv group, the ipv group had significantly lower **positive parenting problem** scores per session at visit 1 [$\exp(\beta_{ipv1:att:time1}) = 0.9360; 95\%CI = (0.8857 - 0.9892)$].

### 4.9.5 Physical Discipline Score

Table 4.12: Dose-Response Model Results

| | Imputed Data | | | Unimputed Data | | |
|---|---|---|---|---|---|---|
| | Exp(Beta) | 95% CI | P-value | Exp(Beta) | 95% CI | P-value |
| (Intercept) | 3.1460 | ( 2.6740 - 3.7013 ) | < 0.0001 | 3.1415 | ( 2.6857 - 3.6746 ) | < 0.0001 |
| wave1 | 0.8485 | ( 0.7045 - 1.0219 ) | 0.0834 | 0.8466 | ( 0.7078 - 1.0126 ) | 0.0687 |
| time1 (control) | 0.3429 | ( 0.2644 - 0.4447 ) | < 0.0001 | 0.3511 | ( 0.2727 - 0.4522 ) | < 0.0001 |
| time2 (control) | 0.3578 | ( 0.2715 - 0.4714 ) | < 0.0001 | 0.3732 | ( 0.2870 - 0.4854 ) | < 0.0001 |
| ipv1 | 1.4571 | ( 1.1997 - 1.7696 ) | 0.0001 | 1.4443 | ( 1.1985 - 1.7405 ) | 0.0001 |
| sex (Male) | 1.3606 | ( 1.2005 - 1.5421 ) | < 0.0001 | 1.3813 | ( 1.2233 - 1.5599 ) | < 0.0001 |
| age (6 - 9 yrs) | 1.0608 | ( 0.9358 - 1.2024 ) | 0.3561 | 1.0575 | ( 0.9363 - 1.1945 ) | 0.3683 |
| att.time1 | 1.0164 | ( 0.9729 - 1.0618 ) | 0.4659 | 1.0229 | ( 0.9803 - 1.0672 ) | 0.2974 |
| att.time2 | 0.9531 | ( 0.9053 - 1.0033 ) | 0.0667 | 0.9562 | ( 0.9103 - 1.0044 ) | 0.0746 |
| wave1:time1 | 2.2254 | ( 1.6139 - 3.0686 ) | < 0.0001 | 2.2480 | ( 1.6467 - 3.0689 ) | < 0.0001 |
| wave1:time2 | 1.0313 | ( 0.7234 - 1.4703 ) | 0.8648 | 1.1137 | ( 0.7865 - 1.5771 ) | 0.5442 |
| time1:ipv1 | 1.2535 | ( 0.8947 - 1.7561 ) | 0.1892 | 1.2606 | ( 0.9099 - 1.7464 ) | 0.1643 |
| time2:ipv1 | 0.7745 | ( 0.5260 - 1.1404 ) | 0.1954 | 0.7631 | ( 0.5235 - 1.1124 ) | 0.1602 |
| wave1:att.time1 | 0.9138 | ( 0.8660 - 0.9643 ) | 0.0010 | 0.9077 | ( 0.8615 - 0.9564 ) | 0.0003 |
| ipv1:att.time1 | 0.9647 | ( 0.9141 - 1.0181 ) | 0.1911 | 0.9622 | ( 0.9130 - 1.0140 ) | 0.1501 |
| wave1:att.time2 | 0.9914 | ( 0.9327 - 1.0538 ) | 0.7819 | 0.9757 | ( 0.9187 - 1.0364 ) | 0.4248 |
| ipv1:att.time2 | 1.0685 | ( 1.0045 - 1.1366 ) | 0.0356 | 1.0793 | ( 1.0155 - 1.1471 ) | 0.0143 |

Figure 4.11: Physical Discipline change from baseline to visit 1

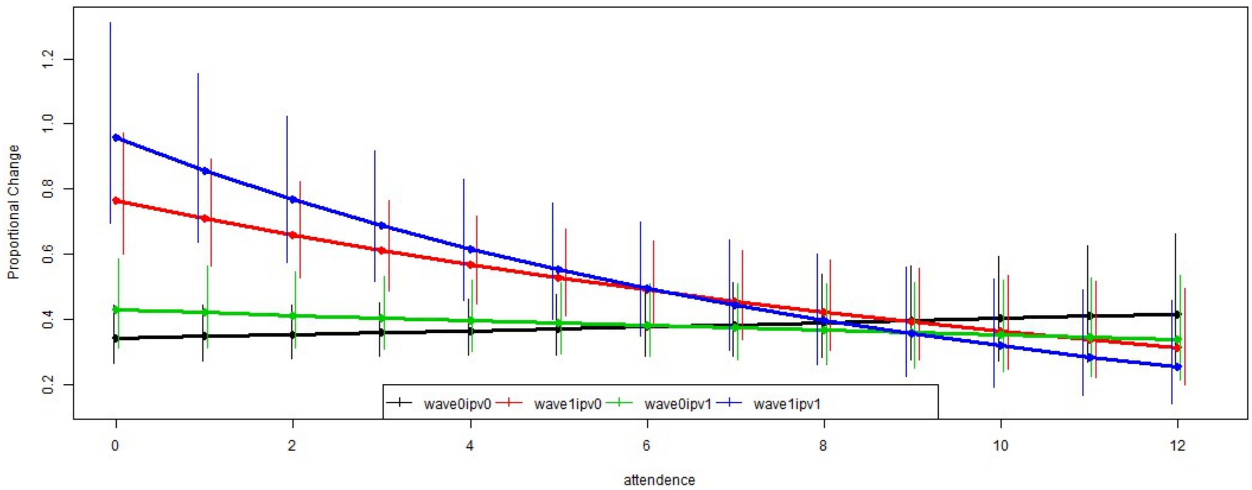
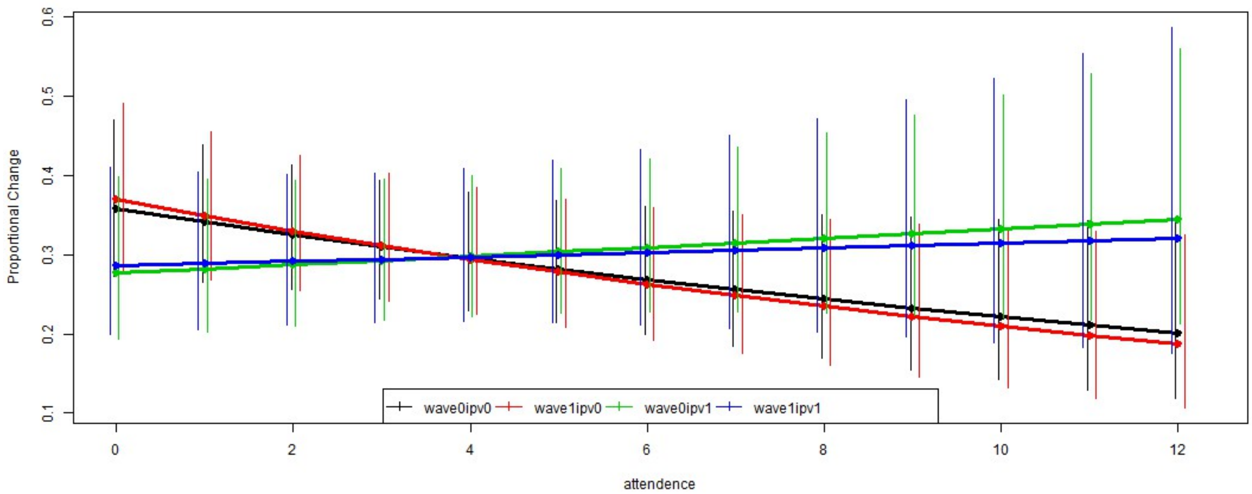


Figure 4.12: Physical Discipline change from baseline to visit 2




By construction, the lower **physical discipline** scores are more desirable. The first plot as well as the table model summary above show that in the baseline to visit 1 interval, there is a significant difference between the two waves in terms of their respective response to dosage. The profiles for groups from Khayelitsha have a significantly lower slope than those from Nyanga which are much flatter, more specifically the slopes for Khayelitsha are about 0.0862-fold lower than Nyanga [$\exp(\beta_{wave1:att:time1}) = 0.9138; 95\%CI = (0.8660 - 0.9643)$]. The second plot shows clear differences between the two ipv groupings, the ipv-exposed group shows increasing trends whilst the no-ipv group shows the desired decreasing trends. One can conclude from the model summary table that the ipv-exposed group has a 0.0685-fold greater slope than the no-ipv group [$\exp(\beta_{ipv1:att:time2}) = 1.0685; 95\%CI = (1.0045 - 1.1366)$].

### 4.9.6 Psychological Discipline Score

Table 4.13: Dose-Response Model Results

| | Imputed Data | | | Unimputed Data | | |
|---|---|---|---|---|---|---|
| | Exp(Beta) | 95% CI | P-value | Exp(Beta) | 95% CI | P-value |
| (Intercept) | 5.3473 | ( 4.7685 - 5.9963 ) | < 0.0001 | 5.3365 | ( 4.7755 - 5.9634 ) | < 0.0001 |
| wave1 | 0.8036 | ( 0.7050 - 0.9160 ) | 0.0011 | 0.8017 | ( 0.7063 - 0.9100 ) | 0.0007 |
| time1 (control) | 0.3888 | ( 0.3235 - 0.4671 ) | < 0.0001 | 0.3897 | ( 0.3256 - 0.4664 ) | < 0.0001 |
| time2 (control) | 0.3014 | ( 0.2458 - 0.3695 ) | < 0.0001 | 0.3118 | ( 0.2560 - 0.3799 ) | < 0.0001 |
| ipv1 | 1.3200 | ( 1.1514 - 1.5132 ) | 0.0001 | 1.3126 | ( 1.1502 - 1.4979 ) | 0.0001 |
| sex (Male) | 1.2265 | ( 1.1215 - 1.3414 ) | < 0.0001 | 1.2425 | ( 1.1389 - 1.3555 ) | < 0.0001 |
| age (6 - 9 yrs) | 1.0975 | ( 1.0042 - 1.1994 ) | 0.0401 | 1.0918 | ( 1.0007 - 1.1912 ) | 0.0486 |
| att.time1 | 1.0163 | ( 0.9851 - 1.0484 ) | 0.3094 | 1.0189 | ( 0.9882 - 1.0506 ) | 0.2300 |
| att.time2 | 1.0090 | ( 0.9744 - 1.0448 ) | 0.6134 | 1.0114 | ( 0.9778 - 1.0463 ) | 0.5109 |
| wave1:time1 | 1.6958 | ( 1.3454 - 2.1374 ) | < 0.0001 | 1.7338 | ( 1.3835 - 2.1729 ) | < 0.0001 |
| wave1:time2 | 1.4904 | ( 1.1504 - 1.9310 ) | 0.0025 | 1.5374 | ( 1.1936 - 1.9801 ) | 0.0009 |
| time1:ipv1 | 1.1937 | ( 0.9355 - 1.5231 ) | 0.1545 | 1.2114 | ( 0.9553 - 1.5361 ) | 0.1140 |
| time2:ipv1 | 0.9254 | ( 0.7000 - 1.2235 ) | 0.5864 | 0.9106 | ( 0.6942 - 1.1945 ) | 0.4989 |
| wave1:att.time1 | 0.9418 | ( 0.9061 - 0.9789 ) | 0.0023 | 0.9361 | ( 0.9011 - 0.9724 ) | 0.0007 |
| ipv1:att.time1 | 0.9695 | ( 0.9322 - 1.0083 ) | 0.1216 | 0.9707 | ( 0.9342 - 1.0086 ) | 0.1283 |
| wave1:att.time2 | 0.9679 | ( 0.9270 - 1.0106 ) | 0.1381 | 0.9609 | ( 0.9209 - 1.0025 ) | 0.0656 |
| ipv1:att.time2 | 1.0004 | ( 0.9577 - 1.0450 ) | 0.9862 | 1.0111 | ( 0.9681 - 1.0560 ) | 0.6200 |

Figure 4.13: Psychological Discipline change from baseline to visit 1

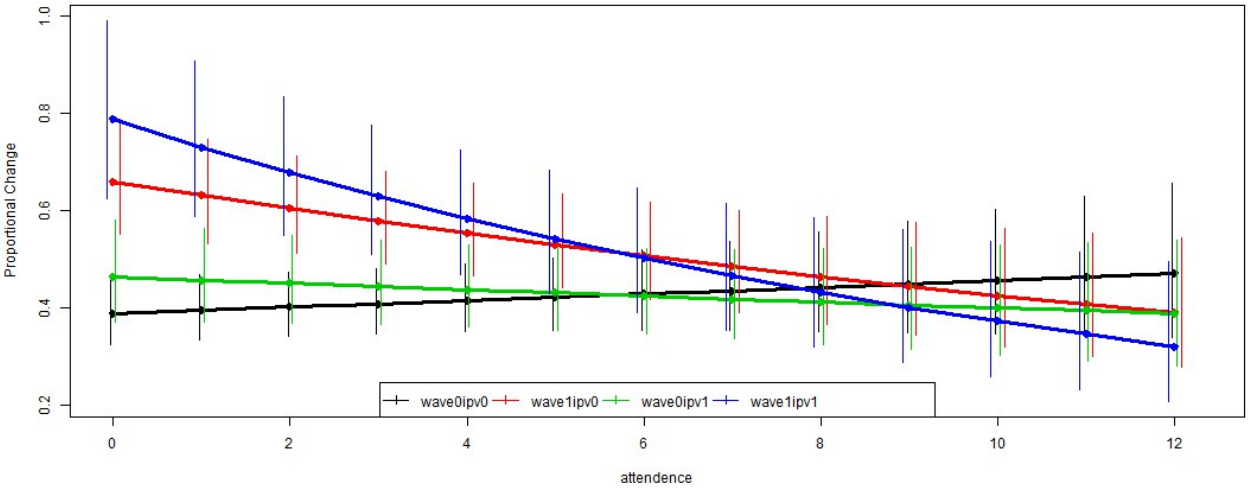
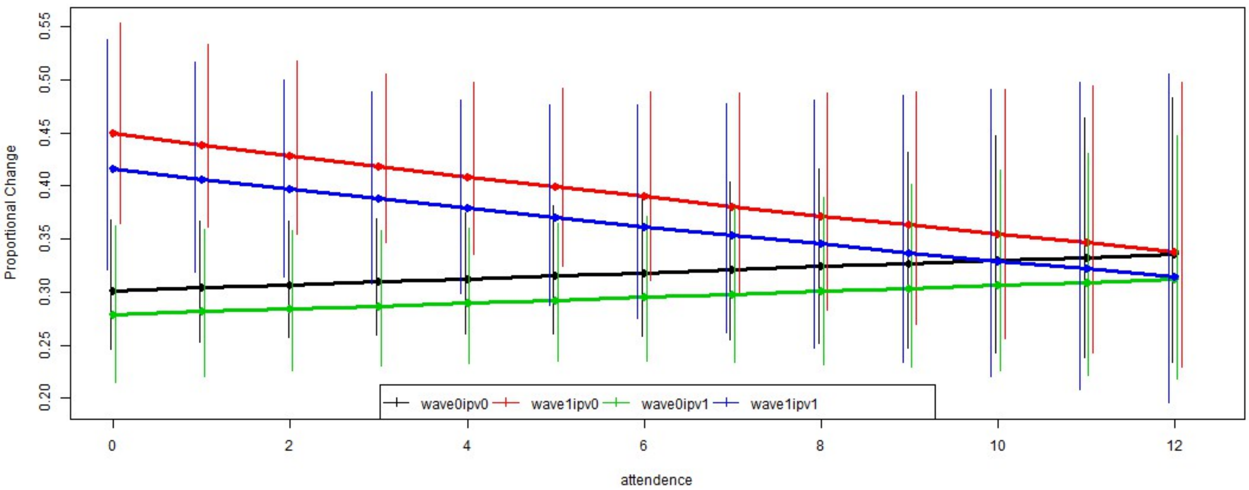



Figure 4.14: Psychological Discipline change from baseline to visit 2




**Psychological discipline**, just like **physical discipline** is more desirable when it decreases. In both plots above, the profiles for the groups in Khayelitsha (wave 1) show the desirable (decreasing) trends whilst profiles for groups in Nyanga shoe the less desirable increasing trends over the two time periods. The table shows that in the baseline-to-visit 1 interval, groups from Khayelitsha had significantly lower responses to the doses (slopes of the profiles) compared to those in Nyanga [$\exp(\beta_{wave1:att:time1}) = 0.9418; 95\%CI = (0.9061 - 0.9789)$]. The second plot also indicates that there is a similar difference in the response-to-dosage between the two waves again in the baseline to visit 2 interval but it is not significant.
[$\exp(\beta_{wave1:att:time2}) = 0.9679; 95\%CI = (0.9270 - 1.0106)$].

### 4.9.7 Parent Positive Behaviour

Table 4.14: Dose-Response Model Results

| | Imputed Data | | | Unimputed Data | | |
|---|---|---|---|---|---|---|
| | Exp(Beta) | 95% CI | P-value | Exp(Beta) | 95% CI | P-value |
| (Intercept) | 18.0898 | ( 17.1069 - 19.1292 ) | < 0.0001 | 18.0578 | ( 17.1168 - 19.0506 ) | < 0.0001 |
| wave1 | 1.1430 | ( 1.0730 - 1.2176 ) | < 0.0001 | 1.1537 | ( 1.084 - 1.2279 ) | < 0.0001 |
| time1 (control) | 0.7814 | ( 0.7153 - 0.8536 ) | < 0.0001 | 0.8174 | ( 0.7515 - 0.8891 ) | < 0.0001 |
| time2 (control) | 0.5143 | ( 0.4566 - 0.5793 ) | < 0.0001 | 0.6333 | ( 0.5726 - 0.7006 ) | < 0.0001 |
| ipv1 | 1.0173 | ( 0.9507 - 1.0885 ) | 0.6197 | 0.9946 | ( 0.93 - 1.0636 ) | 0.8739 |
| sex (Male) | 0.9250 | ( 0.8734 - 0.9797 ) | 0.0078 | 0.954 | ( 0.9146 - 0.9952 ) | 0.0291 |
| age (6 - 9 yrs) | 0.5584 | ( 0.5265 - 0.5924 ) | < 0.0001 | 0.5544 | ( 0.5297 - 0.5802 ) | < 0.0001 |
| att.time1 | 1.0085 | ( 0.9952 - 1.0219 ) | 0.2129 | 1.0062 | ( 0.9938 - 1.0188 ) | 0.3281 |
| att.time2 | 0.9955 | ( 0.9667 - 1.0251 ) | 0.7625 | 0.9939 | ( 0.9766 - 1.0116 ) | 0.498 |
| wave1:time1 | 0.8088 | ( 0.6718 - 0.9737 ) | 0.025 | 0.7771 | ( 0.6917 - 0.873 ) | < 0.0001 |
| wave1:time2 | 0.8470 | ( 0.7075 - 1.0140 ) | 0.0706 | 0.7981 | ( 0.6933 - 0.9187 ) | 0.0017 |
| time1:ipv1 | 0.8465 | ( 0.7235 - 0.9904 ) | 0.0374 | 0.8402 | ( 0.7359 - 0.9593 ) | 0.0101 |
| time2:ipv1 | 0.7524 | ( 0.5701 - 0.9929 ) | 0.0444 | 0.7266 | ( 0.6089 - 0.8671 ) | 0.0004 |
| wave1:att.time1 | 1.0387 | ( 1.0176 - 1.0601 ) | 0.0003 | 1.0367 | ( 1.0208 - 1.0528 ) | < 0.0001 |
| ipv1:att.time1 | 1.0390 | ( 1.0217 - 1.0566 ) | < 0.0001 | 1.0515 | ( 1.0342 - 1.0691 ) | < 0.0001 |
| wave1:att.time2 | 1.0703 | ( 1.0388 - 1.1027 ) | < 0.0001 | 1.0717 | ( 1.0496 - 1.0943 ) | < 0.0001 |
| ipv1:att.time2 | 1.0099 | ( 0.9756 - 1.0454 ) | 0.5755 | 1.0108 | ( 0.986 - 1.0362 ) | 0.3973 |

Figure 4.15: Parent Positive Behaviour change from baseline to visit 1

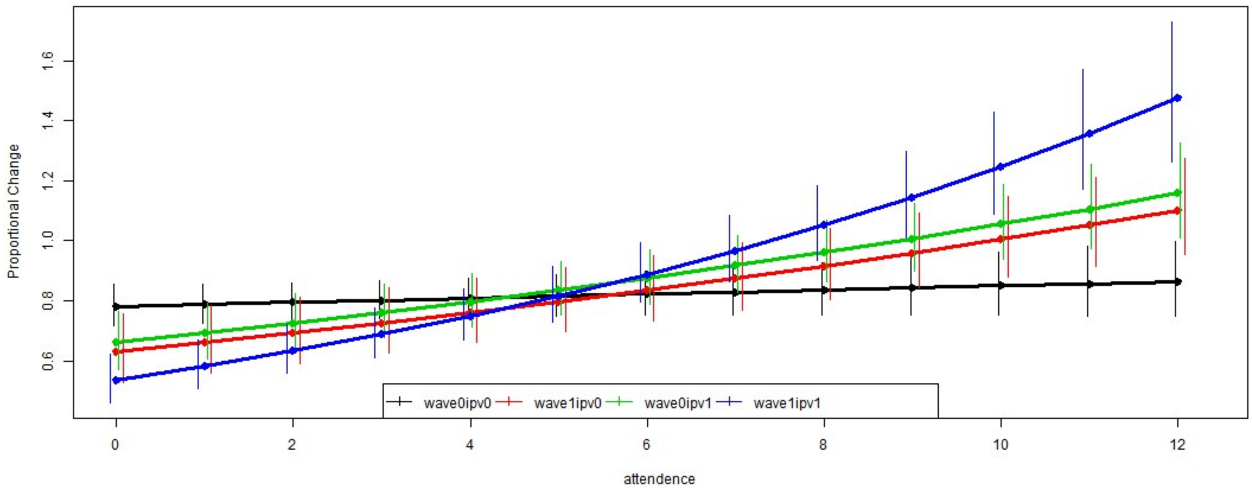
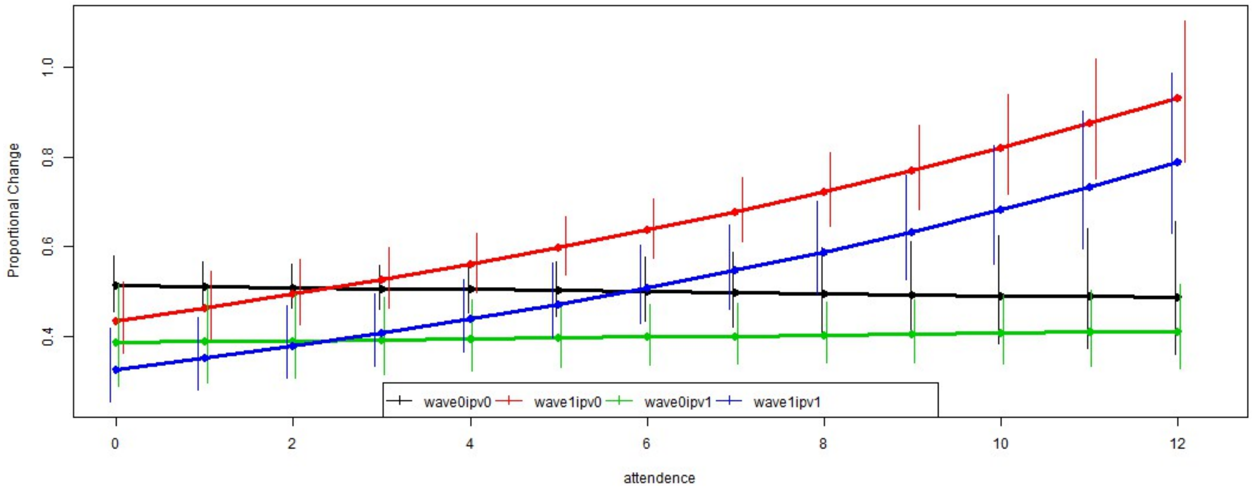


Figure 4.16: Parent Positive Behaviour change from baseline to visit 2




The above plots show a general positive relation (desirable) between the **positive parenting** score and attendance over the two time periods, though a few of the profiles are much flatter. The first plot shows the plots of the profiles crossing in such a way to indicate that both three-way interactions are present. The model summary table above shows that in the baseline to visit 1 interval, there is a significantly different response to dosage (slope) between the groups in Nyanga and Kwayelitsha [$\exp(\beta_{wave1:att:time1}) = 1.0387; 95\%CI = (1.0176 - 1.0601)$] and also between the two ipv groupings [$\exp(\beta_{ipv1:att:time1}) = 1.0390; 95\%CI = (1.0217 - 1.0566)$]. In other words, the programme had a greater (desirable) impact, per session, in Khayelitsha and amongst those who had baseline ipv-exposure when focussing on the relative score changes between the baseline and post-test visits. Only the wave impact is sustained through to the one year follow-up visit [$\exp(\beta_{wave1:att:time2}) = 1.0703; 95\%CI = (1.0388 - 1.1027)$].

### 4.9.8 Parent Negative Behaviour

Table 4.15: Dose-Response Model Results

| | Imputed Data | | | Unimputed Data | | |
|---|---|---|---|---|---|---|
| | Exp(Beta) | 95% CI | P-value | Exp(Beta) | 95% CI | P-value |
| (Intercept) | 3.7863 | ( 3.3839 - 4.2365 ) | < 0.0001 | 3.8406 | ( 3.4393 - 4.2888 ) | < 0.0001 |
| wave1 | 0.7620 | ( 0.6636 - 0.8750 ) | 1.00E-04 | 0.7729 | ( 0.6740 - 0.8864 ) | 0.0002 |
| time1 (control) | 0.8200 | ( 0.6945 - 0.9682 ) | 0.0192 | 0.8512 | ( 0.7230 - 1.0023 ) | 0.0533 |
| time2 (control) | 0.3687 | ( 0.2845 - 0.4778 ) | < 0.0001 | 0.4476 | ( 0.3580 - 0.5597 ) | < 0.0001 |
| ipv1 | 0.8908 | ( 0.7630 - 1.0400 ) | 0.1433 | 0.8957 | ( 0.7689 - 1.0435 ) | 0.1574 |
| sex (Male) | 1.1022 | ( 0.9994 - 1.2155 ) | 0.0514 | 1.1336 | ( 1.0354 - 1.2412 ) | 0.0067 |
| age (6 - 9 yrs) | 0.8568 | ( 0.7778 - 0.9438 ) | 0.0017 | 0.8612 | ( 0.7864 - 0.9431 ) | 0.0013 |
| att.time1 | 0.9774 | ( 0.9504 - 1.0052 ) | 0.1104 | 0.9765 | ( 0.9499 - 1.0039 ) | 0.0925 |
| att.time2 | 1.0524 | ( 1.0094 - 1.0972 ) | 0.0164 | 1.0570 | ( 1.0214 - 1.0939 ) | 0.0015 |
| wave1:time1 | 0.9950 | ( 0.7703 - 1.2852 ) | 0.9692 | 0.9414 | ( 0.7434 - 1.1920 ) | 0.6159 |
| wave1:time2 | 3.1707 | ( 2.2638 - 4.4409 ) | < 0.0001 | 3.3281 | ( 2.5325 - 4.3735 ) | < 0.0001 |
| time1:ipv1 | 0.8769 | ( 0.6543 - 1.1751 ) | 0.3790 | 0.8678 | ( 0.6571 - 1.1459 ) | 0.3174 |
| time2:ipv1 | 0.5264 | ( 0.3112 - 0.8907 ) | 0.0168 | 0.5073 | ( 0.3512 - 0.7329 ) | 0.0003 |
| wave1:att.time1 | 1.0493 | ( 1.0118 - 1.0882 ) | 0.0096 | 1.0525 | ( 1.0166 - 1.0897 ) | 0.0039 |
| ipv1:att.time1 | 1.0257 | ( 0.9834 - 1.0699 ) | 0.2376 | 1.0259 | ( 0.9870 - 1.0664 ) | 0.1948 |
| wave1:att.time2 | 0.8622 | ( 0.8184 - 0.9083 ) | < 0.0001 | 0.8678 | ( 0.8286 - 0.9088 ) | < 0.0001 |
| ipv1:att.time2 | 1.0422 | ( 0.9506 - 1.1425 ) | 0.3785 | 1.0337 | ( 0.9790 - 1.0915 ) | 0.2317 |

Figure 4.17: Parent Negative Behaviour change from baseline to visit 1

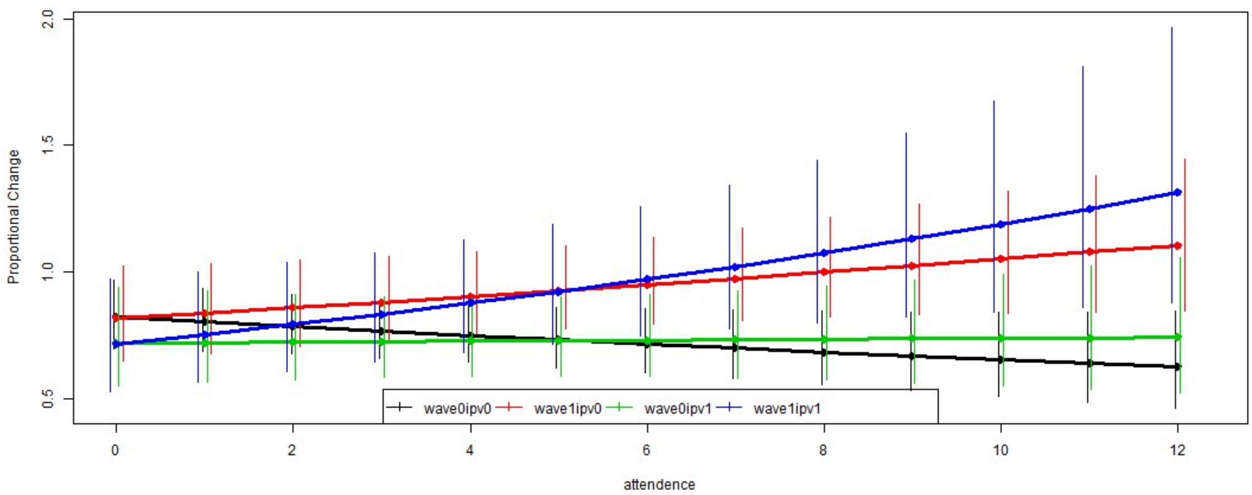
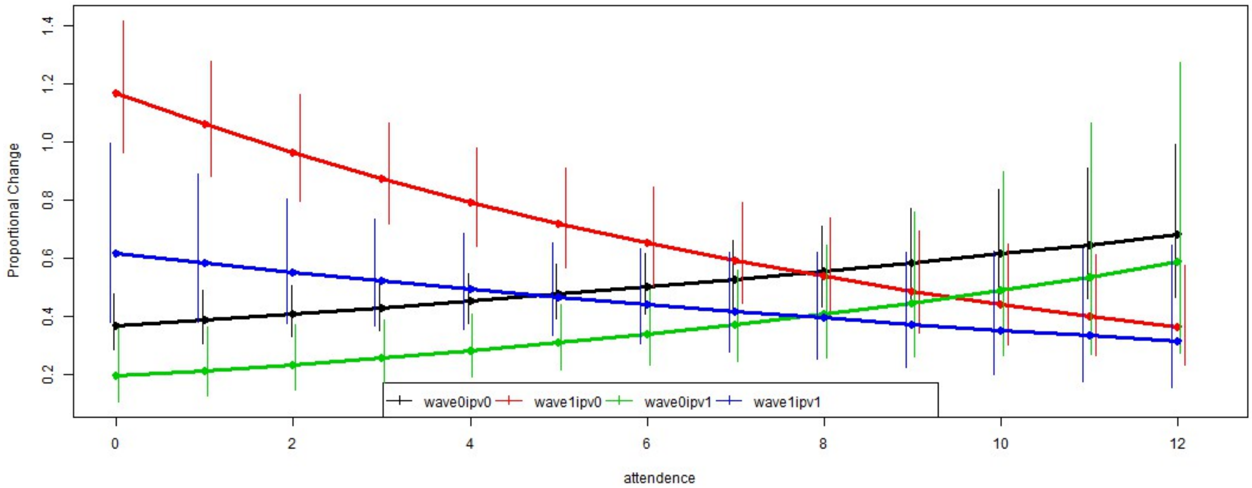


Figure 4.18: Parent Negative Behaviour change from baseline to visit 2




The plots above show quite different profiles especially between the two waves. It is unclear to ascertain a general trend in the way the observed **parent negative behaviour** moves with attendance as the profiles for the two waves contrast each other in the two time periods. The model summary table above also shows that there is a significant difference between the two waves in terms of the response to dosage over the two time intervals with Khayelitsha having a higher (undesirable) response-to-dosage between baseline and visit 1 [$\exp(\beta_{wave1:att:time1}) = 1.0493; 95\%CI = (1.0118 - 1.0699)$] and lower in the other time interval [$\exp(\beta_{wave1:att:time2}) = 0.8622; 95\%CI = (0.8184 - 0.9083)$]. No significant differences were noticed in the response-to-dosage between the ipv groups over the two time intervals under consideration.

### 4.9.9 Child Positive Behaviour

Table 4.16: Dose-Response Model Results

| | Imputed Data | | | Unimputed Data | | |
|---|---|---|---|---|---|---|
| | Exp(Beta) | 95% CI | P-value | Exp(Beta) | 95% CI | P-value |
| (Intercept) | 38.8232 | ( 37.0065 - 40.7292 ) | < 0.0001 | 38.7053 | ( 37.3324 - 40.1287 ) | < 0.0001 |
| wave1 | 0.7880 | ( 0.7524 - 0.8253 ) | < 0.0001 | 0.7884 | ( 0.7533 - 0.8252 ) | < 0.0001 |
| time1 (control) | 0.8709 | ( 0.8209 - 0.9239 ) | < 0.0001 | 0.9139 | ( 0.8671 - 0.9632 ) | 0.0008 |
| time2 (control) | 0.5646 | ( 0.4942 - 0.6450 ) | < 0.0001 | 0.6244 | ( 0.5852 - 0.6662 ) | < 0.0001 |
| ipv1 | 0.8926 | ( 0.8466 - 0.9411 ) | < 0.0001 | 0.9004 | ( 0.8570 - 0.9461 ) | < 0.0001 |
| sex (Male) | 0.9959 | ( 0.9469 - 1.0474 ) | 0.8727 | 1.0228 | ( 0.9929 - 1.0535 ) | 0.1363 |
| age (6 - 9 yrs) | 0.6283 | ( 0.6018 - 0.6560 ) | < 0.0001 | 0.6171 | ( 0.5982 - 0.6367 ) | < 0.0001 |
| att.time1 | 1.0051 | ( 0.9910 - 1.0195 ) | 0.4784 | 1.0036 | ( 0.9956 - 1.0117 ) | 0.3812 |
| att.time2 | 0.9910 | ( 0.9734 - 1.0090 ) | 0.3254 | 0.9902 | ( 0.9791 - 1.0015 ) | 0.0895 |
| wave1:time1 | 0.8519 | ( 0.7743 - 0.9373 ) | 0.0010 | 0.8340 | ( 0.7708 - 0.9025 ) | < 0.0001 |
| wave1:time2 | 1.1376 | ( 1.0130 - 1.2775 ) | 0.0294 | 1.1192 | ( 1.0201 - 1.2279 ) | 0.0173 |
| time1:ipv1 | 0.9886 | ( 0.8930 - 1.0944 ) | 0.8246 | 0.9419 | ( 0.8632 - 1.0278 ) | 0.1791 |
| time2:ipv1 | 0.8228 | ( 0.6998 - 0.9675 ) | 0.0183 | 0.8407 | ( 0.7507 - 0.9414 ) | 0.0026 |
| wave1:att.time1 | 1.0426 | ( 1.0297 - 1.0556 ) | < 0.0001 | 1.0422 | ( 1.0310 - 1.0535 ) | < 0.0001 |
| ipv1:att.time1 | 1.0065 | ( 0.9945 - 1.0187 ) | 0.2873 | 1.0131 | ( 1.0015 - 1.0248 ) | 0.0273 |
| wave1:att.time2 | 1.0099 | ( 0.9820 - 1.0387 ) | 0.4904 | 1.0064 | ( 0.9917 - 1.0213 ) | 0.3947 |
| ipv1:att.time2 | 1.0368 | ( 1.0135 - 1.0606 ) | 0.0019 | 1.0415 | ( 1.0247 - 1.0587 ) | < 0.0001 |

Figure 4.19: Child Positive Behaviour change from baseline to visit 1

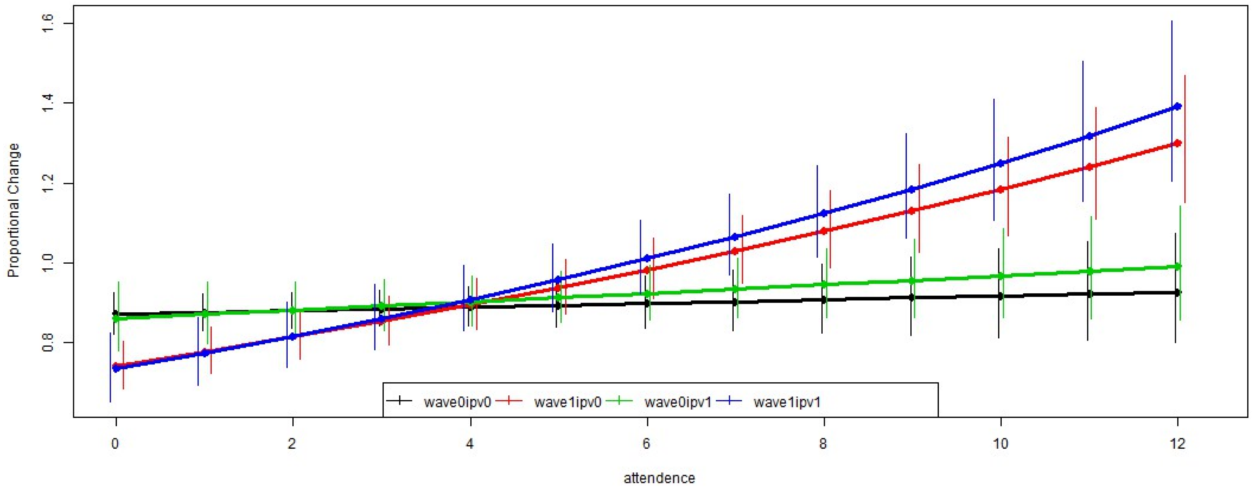
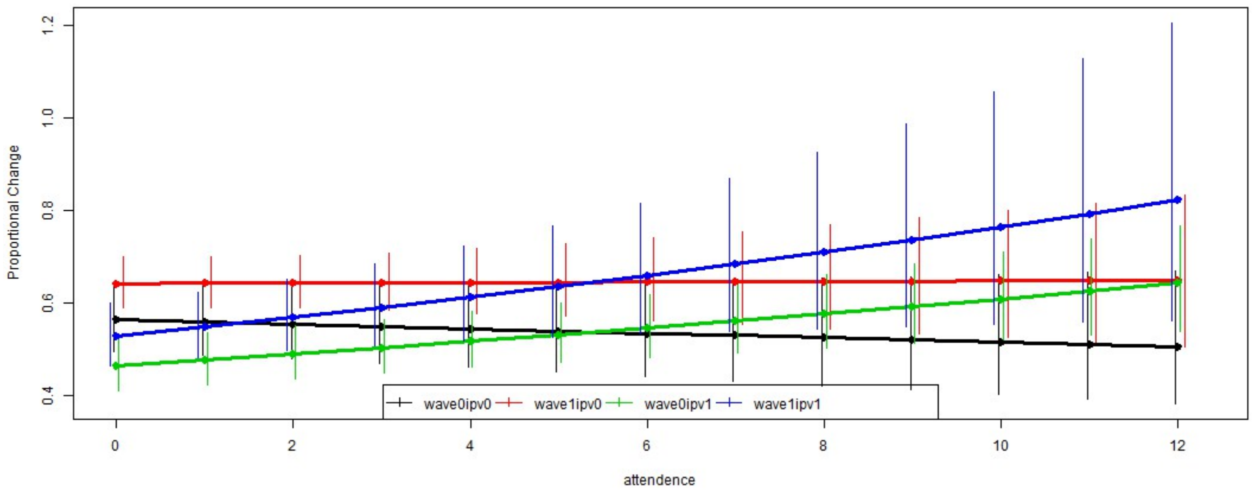



Figure 4.20: Child Positive Behaviour change from baseline to visit 2




The two plots above show a general positive (desired) relation between attendance and the observed **child positive behaviour** score. Between baseline and visit 1, there is a clear difference in the slopes of the profiles in the two waves with Khayelitsha showing higher relative responses to the dosage [$\exp(\beta_{wave1:att:time1}) = 1.0426; 95\%CI = (1.0297 - 1.0556)$]. In other words, the groups in Khayelitsha seem to have 4.26% higher responses to the dosage than those in Nyanga. Between baseline and visit 2, it also seems that the ipv-exposed group has a higher response-to-dosage than the no–ipv group [$\exp(\beta_{ipv1:att:time2}) = 1.0368; 95\%CI = (1.0135 - 1.0606)$].

### 4.9.10 Child Negative Behaviour

Table 4.17: Dose-Response Model Results

| | Imputed Data | | | Unimputed Data | | |
|---|---|---|---|---|---|---|
| | Exp(Beta) | 95% CI | P-value | Exp(Beta) | 95% CI | P-value |
| (Intercept) | 3.06 | ( 2.657 - 3.5241 ) | < 0.0001 | 3.0692 | ( 2.6703 - 3.5277 ) | < 0.0001 |
| wave1 | 0.5553 | ( 0.4575 - 0.6739 ) | < 0.0001 | 0.5553 | ( 0.4577 - 0.6736 ) | < 0.0001 |
| time1 (control) | 0.7303 | ( 0.5920 - 0.9008 ) | 0.0033 | 0.7999 | ( 0.6497 - 0.9848 ) | 0.0354 |
| time2 (control) | 0.2433 | ( 0.1763 - 0.3356 ) | < 0.0001 | 0.3241 | ( 0.2360 - 0.4451 ) | < 0.0001 |
| ipv1 | 0.8294 | ( 0.6771 - 1.0160 ) | 0.0708 | 0.8263 | ( 0.6765 - 1.0092 ) | 0.0615 |
| sex (Male) | 1.0481 | ( 0.9229 - 1.1902 ) | 0.4695 | 1.0779 | ( 0.9497 - 1.2233 ) | 0.2456 |
| age (6 - 9 yrs) | 0.5555 | ( 0.4816 - 0.6408 ) | < 0.0001 | 0.5485 | ( 0.4783 - 0.6292 ) | < 0.0001 |
| att.time1 | 0.9500 | ( 0.9164 - 0.9848 ) | 0.0052 | 0.9496 | ( 0.9166 - 0.9839 ) | 0.0043 |
| att.time2 | 1.0138 | ( 0.9632 - 1.0670 ) | 0.6002 | 1.0141 | ( 0.9657 - 1.0648 ) | 0.5746 |
| wave1:time1 | 1.3596 | ( 0.9860 - 1.8747 ) | 0.0609 | 1.3551 | ( 0.9855 - 1.8635 ) | 0.0615 |
| wave1:time2 | 1.6799 | ( 1.0277 - 2.7460 ) | 0.0385 | 1.5779 | ( 0.9819 - 2.5357 ) | 0.0595 |
| time1:ipv1 | 1.0476 | ( 0.7348 - 1.4934 ) | 0.7974 | 0.9807 | ( 0.6863 - 1.4012 ) | 0.9145 |
| time2:ipv1 | 0.7577 | ( 0.3707 - 1.5487 ) | 0.4468 | 0.7540 | ( 0.4053 - 1.4028 ) | 0.3727 |
| wave1:att.time1 | 1.0449 | ( 0.9980 - 1.0941 ) | 0.0608 | 1.0438 | ( 0.9972 - 1.0926 ) | 0.0656 |
| ipv1:att.time1 | 1.0653 | ( 1.0156 - 1.1174 ) | 0.0095 | 1.0711 | ( 1.0213 - 1.1233 ) | 0.0047 |
| wave1:att.time2 | 0.9926 | ( 0.9205 - 1.0704 ) | 0.8472 | 0.9920 | ( 0.9223 - 1.0670 ) | 0.8295 |
| ipv1:att.time2 | 0.9901 | ( 0.8991 - 1.0902 ) | 0.8389 | 0.9918 | ( 0.9069 - 1.0847 ) | 0.8570 |

Figure 4.21: Child Negative Behaviour change from baseline to visit 1

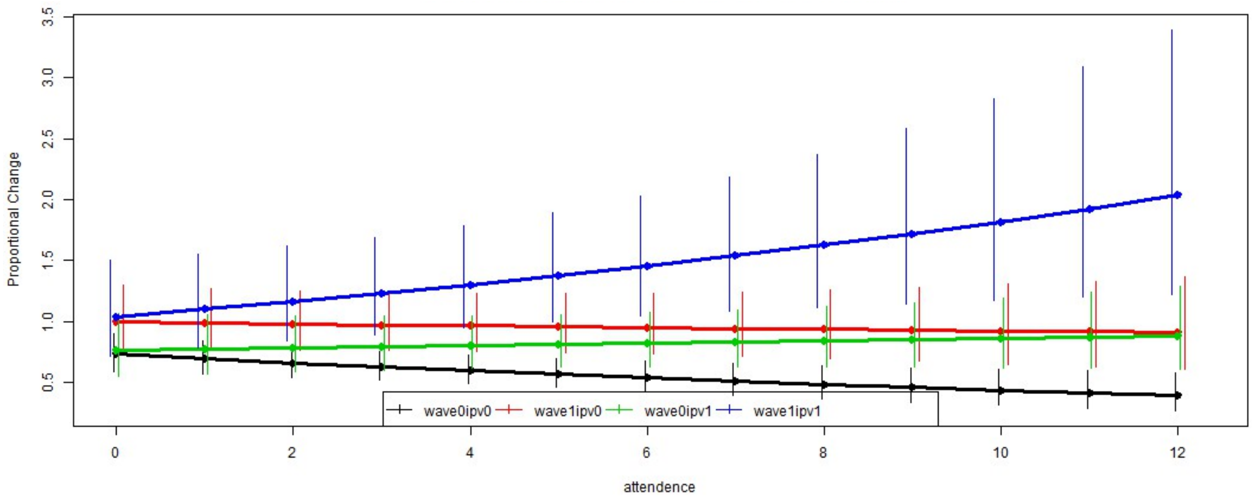
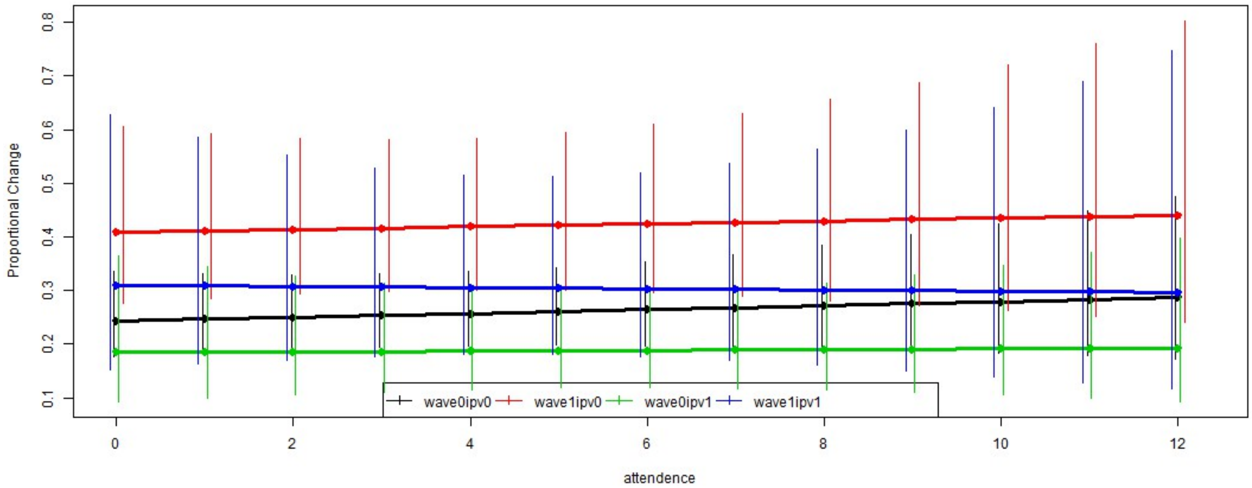


Figure 4.22: Child Negative Behaviour change from baseline to visit 2




There is a generally unclear relationship between attendance and the observed **child negative behaviour** score from the plots since some profiles show positive trends and others show the opposite in the first plot and the profiles in the second plot are flatter. In the baseline to visit 1 interval, the slope of the profiles are significantly different between the two waves $[\exp(\beta_{wave1:att:time1}) = 1.0449; 95\%CI = (0.9980 - 1.0941); p - value = 0.0608]$. The aforementioned actually implies that the participants in Khayelitsha had about 4.49% higher response to dosage compared to those in Nyanga in this time period. The model also shows that the response to dosage for the ipv-exposed group was also higher about 6.53% higher than the no-ipv group $[\exp(\beta_{ipv1:att:time1}) = 1.0653; 95\%CI = (1.0156 - 1.1174)]$. The confidence intervals for the proportional score changes between baseline and visit 2 overlap and the profiles are all flat which implies no significant differences in the slopes.

## 4.10 Model Diagnostics

In order to check that the models that were fitted were appropriate, plots were made for (1) the residuals versus the fitted values and (2) the histograms of the residuals. A random scatter in the residuals vs. fitted values plot shows a good fit whilst some patterns in the aforementioned plot usually indicate the need to transform the response variable or fit a different distribution or change the link function, e.t.c. The histograms show the distributions of the model (within-group) residuals and ideally these should be be small and centred around zero. These plots were used to ascertain whether the models (both the binary-intervention and the dose-response models) were appropriate and where necessary, models were refitted until the plots were satisfactory. The plots in Figures 4.23 and 4.24 below show the aforementioned diagnostics for the dose-response models computed on the unimputed data, the rest of the plots are in Appendix D. Generally the plots show reasonable model fit as there are no clear patterns to suggest otherwise.

Figure 4.23: Plots of residuals vs. fitted values for the dose-response models

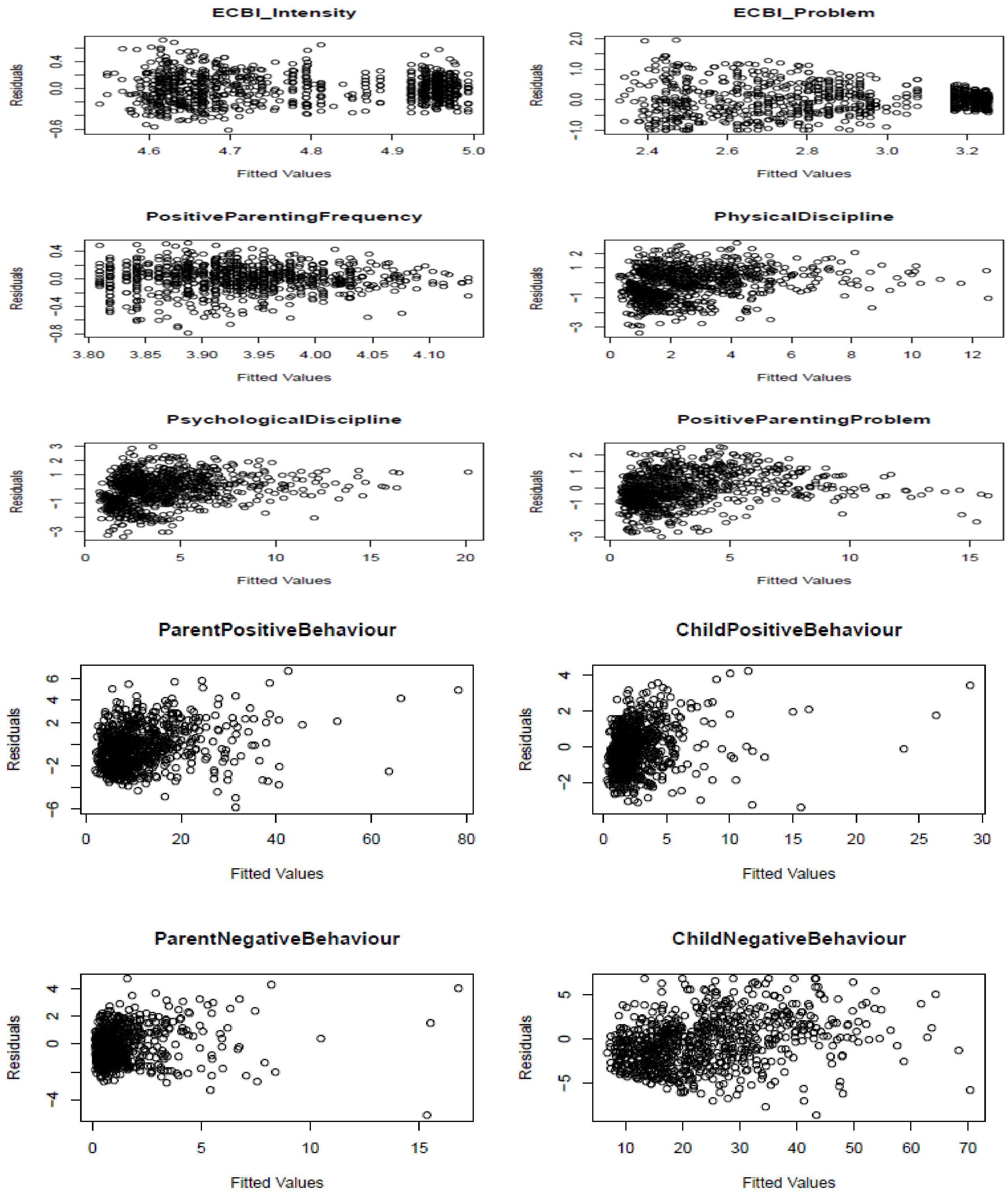
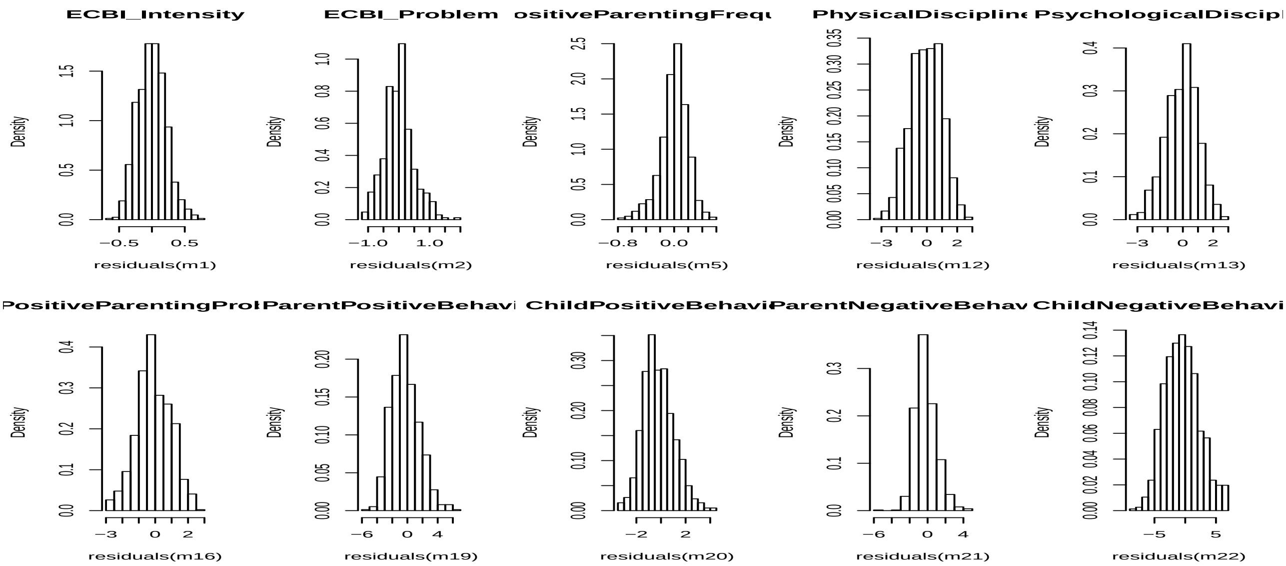


Figure 4.24: Histograms of residuals for the dose-response models




## 4.11 Chapter Summary/ Conclusions

This chapter focused on the theory behind GLMMs, the model fitting process and summarizing the model results. The choice of fitting GLMMs was mainly because of the study design: (1) the fact that repeated measurements were taken on each participant dyad introduced the need to account for within-subject variation and (2) the intervention program was group-based and as such participants' responses were also expected to depend on the groups. One major feature of GLMMs is that they treat responses as continuous or count measurements calculated by summing the individual (observable) item scores.

The model fitting stage, consisted of using the R packages gamlss (Rigby and Stasinopoulos, 2005) and glmmPQL (Venables and Ripley, 2002) to estimate the effects maximizing the penalized quasi-likelihood and penalized likelihood functions respectively. One major setback encountered during estimation is that at times the covariance structure of the random effects had to be simplified (by assuming independence) in order to achieve convergence of the algorithms. For simpler projects, this estimation process can easily be implemented in the multiple imputation framework using the R package **mice** (Van Buuren and Groothuis-Oudshoorn, 2011). However, because of the complexities of both the imputation strategy in the previous chapter and the models assumed in this chapter, models were fitted separately for the 5 data sets and Rubin's rules were programmed explicitly to pool the results.

Discussions on the model results for the binary-intervention and dose response models are included in sections 4.8 and 4.9 respectively. The former models revealed that the intervention program worked particularly well in improving outcomes like (1) **ECBI Intensity**, (2) **BDI**, (3) **Parenting Distress**, (4) **Parenting Stress** , (5) **Observed Parent Positive Behaviour** and (6) **Observed Child Positive Behaviour**. These improvements were more prominent on the per-protocol analysis (indicating that session attendance matters) and in some of the models, the effect of the program wasn't very clear between baseline and visit 2.

The dose-response models were based on the main hypothesis account that each additional group session attended might have an impact on the outcomes of interest. These models included three-way interaction terms that were based on two hypotheses: (1) the response-to-dosage differed between the two waves in the two time intervals and (2) the response-to-dosage differed between the two ipv-groupings in the two time intervals. Most of the models showed that higher attendance significantly improved the scores especially in the baseline to visit 1 interval however these improvements were not as substantial between baseline and visit 2.

# Chapter 5

# Conclusions and further discussions

This chapter summarizes all the findings both in terms of the methodology and also with respect to the study hypotheses. Firstly, it was concluded that the randomization of participants into the control and intervention arms was successful since the profiles in these two were comparable at baseline. Imputation was done on the item level since it was not always the case that if one item is missing then the rest of the items for the score were also missing.It was assumed that data was Missing at Random (MAR) and the multiple imputation model that was implemented closely followed that used in the measurement model. Other possibly more preferable imputation methods include (1) imputing within the different randomized study arms and not treat the study arm as another imputation predictor variable (2) likelihood-based implicit imputation which is inherent in the GLMMs that were fitted and (3) multi-level imputation which specifically accounts for the nested nature of the group-based RCT data. Random forest imputation, a relatively new non-parametric method was chosen over other regression based methods like predictive mean matching (semi-parametric, continuous data), logistic regression (parametric for binary data) as it was relatively faster (in terms of convergence) and it could be used for any data type. Also batch imputation was implemented as a way to solve computer memory problems as the imputation was done on the item level.

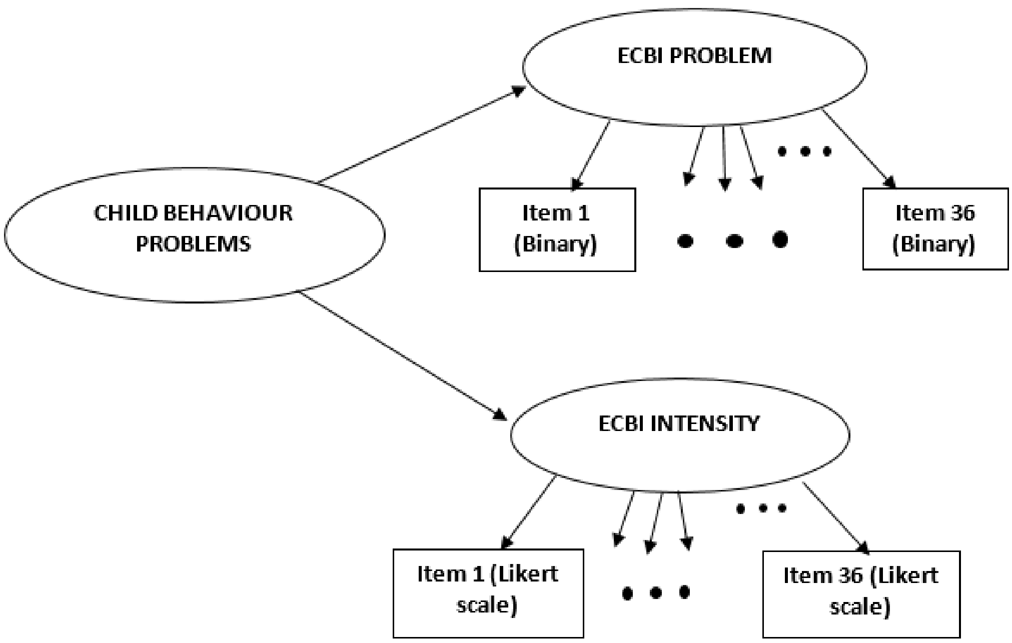
Generalized linear mixed effects models (GLMMs) were fitted and estimates were obtained by optimizing the penalized quasi-likelihood and penalized likelihood functions as a way to control the stability of estimates. In some cases during the model fitting stage, the correlation structure of the random effects was simplified by assuming independence in order to minimize the number of parameters to be estimated and therefore allow for convergence to be reached. The models were fitted on multiple data sets and estimates were pooled using Rubin's rules. Due to the low incidence of missingness, it was expected that models models fitted on the unimputed complete cases would not differ substantially from those fitted on the imputed data regarding the precision of estimates. Model fit was assessed by using residual analysis.

Generally, both the binary-intervention models together with the dose-response models showed some improvements in behaviour between baseline and post test but fewer models showed the significant differences between the baseline and one-year follow-up visits. It was also concluded that attending more sessions was more beneficial as the intervention effect was noticed more in the per-protocol analysis than in the intention to treat analysis. Another conclusion is that the intervention program did help to improve behaviour on most of the outcomes of interest but not all of these improvements were sustained through to the one year follow-up visit. Similar findings can be pointed out about whether the different wave-ipv combinations influenced the participants' response to the program. Most of the models show wave and ipv interacting significantly in the baseline-to-visit1 interval and this generally faded when considering the baseline-to-visit2 interval.

Other sophisticated models like SEM or Rasch models could possibly be more preferable as they treat the outcomes of interest as latent variables rather than assuming that they are directly observable as is assumed with the GLMMs. The input data is in the form of responses to questions aimed to determine the extent/ level of certain behavioural characteristics. These traits are actually latent characteristics and the measured data just gives us pointers towards what these may be. One main criticism of the GLMM approach taken in this chapter is from the fact that, by design the outcomes of interest (responses) were treated as continuous measures calculated by summing up individual (binary or Likert scale) item scores. This assumes that the outcomes of interest are actually (directly) observable. The construction of the summative scores makes also assumes that (1) the items making up the scores are equally weighted i.e. they contribute equally weighted information about the outcome of interest and (2) "Very rarely" - "Never" $\equiv$ "Rarely" - "Very rarely" $\equiv \cdots \equiv$ "Always" - "Very often" for a 7-point Likert scale. There are more sophisticated ways of dealing with these latent variables for example Structural equation modelling (SEM) or Rasch modelling.

The main idea behind SEM (Hair *et al.*, 1998, pp. 627 - 777 ) is that it allows for the creation of latent constructs based on the measured item responses using factor analysis methods (which allow for different weightings of the items) and then analyse these latent constructs using linear models. So for example, instead of summing 36 Likert scale items to obtain **ECBI Intensity** and 36 binary items to obtain **ECBI Problem** that together measure the **Child Behaviour Problems**, you can construct a latent variable, **Child Behaviour Problems** that is measured by the **ECBI Intensity** and **ECBI Problem** and their items and is calculated through a Confirmatory Factor Analysis (CFA).

Figure 5.1: Illustration example of a CFA on Child Behaviour Problems

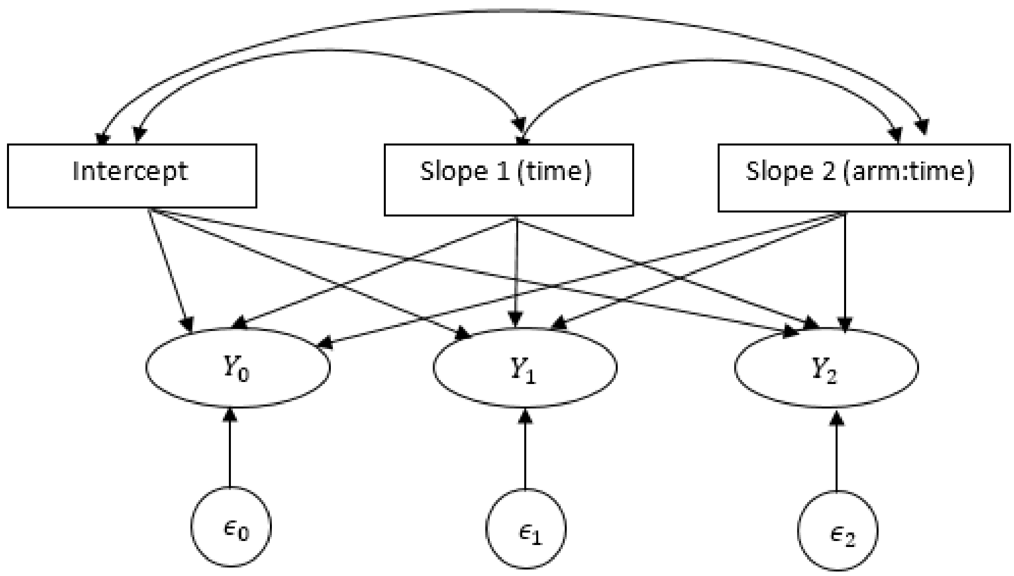



This (as illustrated in Figure 5.1) can be combined with the model time impact:

$$\mathbf{Y} = (\beta_0 + b_0) + (\beta_1 + b_1)\mathbf{time} + (\beta_2 + b_2)\mathbf{arm{:}time} + \mathbf{error} \tag{5.1}$$

where **Y** is the latent variable **Child Behaviour Problem** in the example.

Figure 5.2: Illustration example of a SEM Child Behaviour Problems




Rasch models (Hatzinger, 2008) also work on the concept that the observed data is a function of the unobserved latent traits specifically (1) the 'item difficulty' and (2) 'person ability'. If the observed data is binary in nature then logistic regression would be used to model the odds of getting the 'right response' on an item. The item difficulty is obtained by (subjectively) ranking the items in a questionnaire and person ability is obtained by assessing how 'well' the individual scores relative to the others. If the observed data is not binary (e.g. ordinal) then the Rasch model can be extended by using a multinomial/ ordinal logistic regression.

# Bibliography


Cronbach, L. J. (1951). Coefficient alpha and the internal structure of tests, *Psychometrika* **16**(3): 297–334.
**URL:** *https://doi.org/10.1007/BF02310555*

Dobson, A. J. and Barnett, A. (2008). *An introduction to generalized linear models*, CRC press.

Hair, J. F., Black, W. C., Babin, B. J., Anderson, R. E., Tatham, R. L. *et al. (1998).* Multivariate data analysis*, Vol. 5, Prentice hall Upper Saddle River, NJ.*

*Hatzinger, R. (2008). A glm framework for item response theory models. reissue of 1994 habilitation thesis.*

*Liaw, A. and Wiener, M. (2001). Classification and regression by randomforest,* **23***.*

*McCulloch, C. E. (2003). Generalized linear mixed models,* NSF-CBMS Regional Conference Series in Probability and Statistics **7***: i–84.*
***URL:*** *http://www.jstor.org/stable/4153190*

*Rigby, R. A. and Stasinopoulos, D. M. (2005). Generalized additive models for location, scale and shape,(with discussion),* Applied Statistics **54***: 507–554.*

*Shah, A. D., Bartlett, J. W., Carpenter, J., Nicholas, O. and Hemingway, H. (2014). Comparison of random forest and parametric imputation models for imputing missing data using mice: a caliber study,* American journal of epidemiology **179***(6): 764–774.*

*Stasinopoulos, M., with contributions from Calliope Akantziliotou, B. R., Heller, G., Ospina, R., Motpan, N., McElduff, F., Voudouris, V., Djennad, M., Enea, M. and Ghalanos., A. (2016).* gamlss.dist: Distributions to be Used for GAMLSS Modelling. *R package version 5.0-0.*
***URL:*** *https://CRAN.R-project.org/package=gamlss.dist*

*Van Buuren, S. and Groothuis-Oudshoorn, K. (2011). mice: Multivariate imputation by chained equations in r,* Journal of Statistical Software **45***(3): 1–67.*
***URL:*** *http://www.jstatsoft.org/v45/i03/*

*Venables, W. N. and Ripley, B. D. (2002).* Modern Applied Statistics with S*, fourth edn, Springer, New York. ISBN 0-387-95457-0.*
***URL:*** *http://www.stats.ox.ac.uk/pub/MASS4*

*White, I. R., Royston, P. and Wood, A. M. (2011). Multiple imputation using chained equations: issues and guidance for practice,* Statistics in medicine **30***(4): 377–399.*

# Appendix A

# Binary-Intervention Model Results

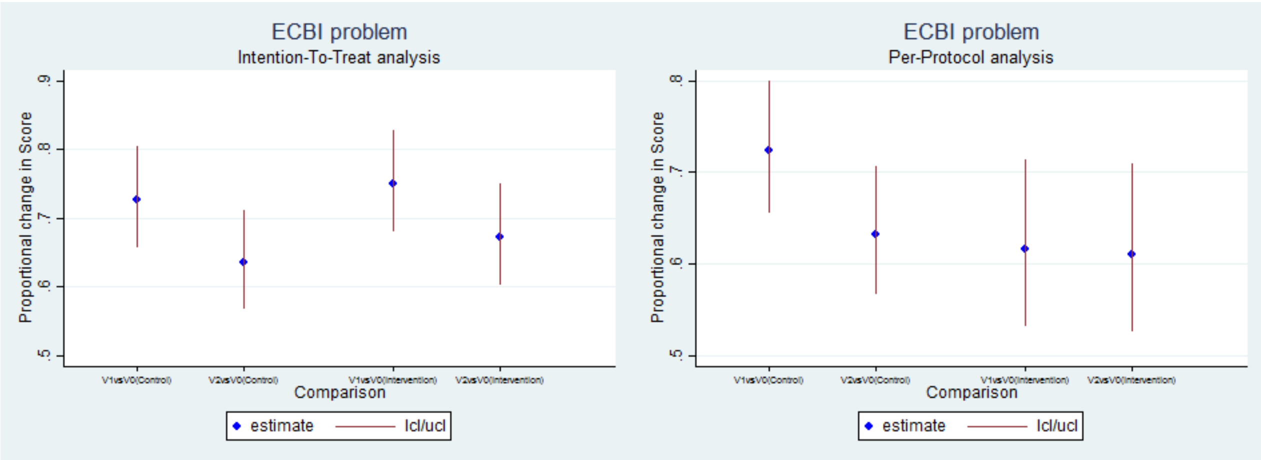
### A.0.1 ECBI Problem Score

Table A.1: (Imputed Data)

| | Intention to Treat | | | Per-Protocol | | |
|---|---|---|---|---|---|---|
| | Exp(Beta) | 95% CI | P-value | Exp(Beta) | 95% CI | P-value |
| (Intercept) | 24.0472 | ( 22.3902 - 25.8268 ) | < 0.0001 | 24.6412 | ( 22.8314 - 26.5945 ) | < 0.0001 |
| wave (Khayelitsha) | 1.0437 | ( 0.9758 - 1.1163 ) | 0.2128 | 1.0269 | ( 0.9534 - 1.1062 ) | 0.4831 |
| sex (Male) | 1.0611 | ( 0.9922 - 1.1348 ) | 0.0835 | 1.0549 | ( 0.9798 - 1.1358 ) | 0.1560 |
| age (6 - 9 yrs) | 0.9574 | ( 0.8946 - 1.0246 ) | 0.2086 | 0.9377 | ( 0.8704 - 1.0101 ) | 0.0902 |
| time1 (control) | 0.7276 | ( 0.6577 - 0.8050 ) | < 0.0001 | 0.7243 | ( 0.6563 - 0.7992 ) | < 0.0001 |
| time2 (control) | 0.6351 | ( 0.5672 - 0.7112 ) | < 0.0001 | 0.6326 | ( 0.5670 - 0.7058 ) | < 0.0001 |
| armtime1 | 1.0189 | ( 0.8983 - 1.1557 ) | 0.7708 | 0.8502 | ( 0.7225 - 1.0005 ) | 0.0507 |
| armtime2 | 1.0407 | ( 0.9013 - 1.2015 ) | 0.5868 | 0.9651 | ( 0.8129 - 1.1460 ) | 0.6855 |
| time1 (intervention) | 0.7503 | ( 0.6801 - 0.8278 ) | < 0.0001 | 0.6158 | ( 0.5313 - 0.7136 ) | < 0.0001 |
| time2 (intervention) | 0.6720 | ( 0.6026 - 0.7493 ) | < 0.0001 | 0.6106 | ( 0.5258 - 0.7090 ) | < 0.0001 |

Table A.2: (Unimputed Data)

| | Intention to Treat | | | Per-Protocol | | |
|---|---|---|---|---|---|---|
| | Exp(Beta) | 95% CI | P-value | Exp(Beta) | 95% CI | P-value |
| (Intercept) | 23.9308 | ( 22.6630 - 25.2695 ) | < 0.0001 | 24.5829 | ( 23.1246 - 26.1333 ) | < 0.0001 |
| wave (Khayelitsha) | 1.0501 | ( 0.9960 - 1.1070 ) | 0.0699 | 1.0164 | ( 0.9563 - 1.0804 ) | 0.5990 |
| sex (Male) | 1.0527 | ( 0.9988 - 1.1096 ) | 0.0556 | 1.0519 | ( 0.9900 - 1.1178 ) | 0.1015 |
| age (6 - 9 yrs) | 0.9716 | ( 0.9215 - 1.0245 ) | 0.2859 | 0.9574 | ( 0.9009 - 1.0175 ) | 0.1604 |
| time1 (control) | 0.7273 | ( 0.6753 - 0.7832 ) | < 0.0001 | 0.7238 | ( 0.6703 - 0.7817 ) | < 0.0001 |
| time2 (control) | 0.5246 | ( 0.4742 - 0.5803 ) | < 0.0001 | 0.5221 | ( 0.4709 - 0.5789 ) | < 0.0001 |
| armtime1 | 0.8897 | ( 0.8043 - 0.9841 ) | 0.0232 | 0.7943 | ( 0.6940 - 0.9091 ) | 0.0009 |
| armtime2 | 0.9833 | ( 0.8578 - 1.1271 ) | 0.8082 | 0.9153 | ( 0.7685 - 1.0902 ) | 0.3203 |
| time1 (intervention) | 0.6470 | ( 0.5952 - 0.7034 ) | < 0.0001 | 0.5749 | ( 0.5079 - 0.6507 ) | < 0.0001 |
| time2 (intervention) | 0.5158 | ( 0.4651 - 0.5721 ) | < 0.0001 | 0.4779 | ( 0.4108 - 0.5560 ) | < 0.0001 |



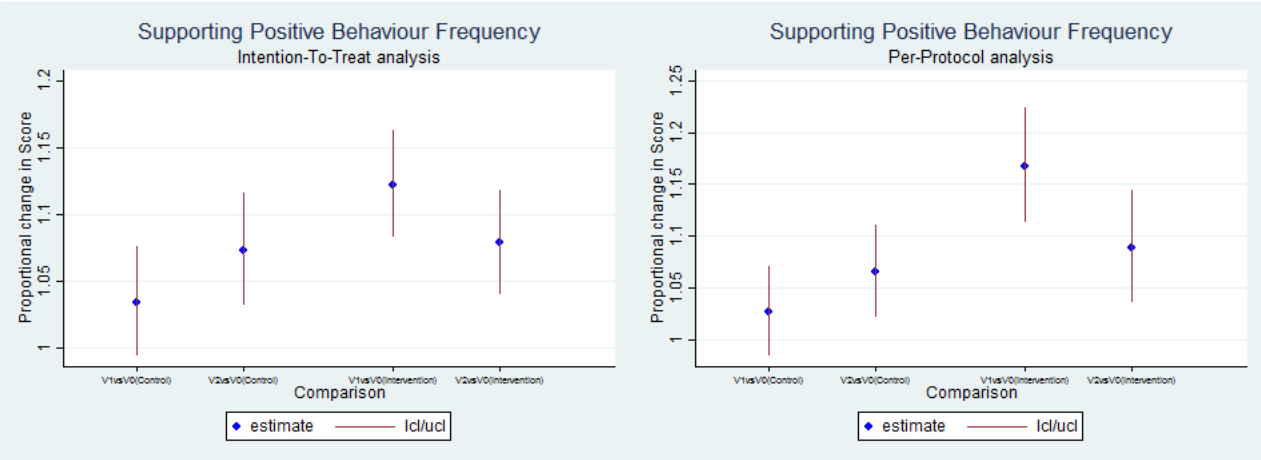

### A.0.2 Supporting Positive Behaviour Frequency Score

Table A.3: (Imputed Data)

| | Intention to Treat | | | Per-Protocol | | |
|---|---|---|---|---|---|---|
| | Exp(Beta) | 95% CI | P-value | Exp(Beta) | 95% CI | P-value |
| (Intercept) | 26.2574 | ( 25.4570 - 27.0829 ) | < 0.0001 | 26.2611 | ( 25.3513 - 27.2035 ) | < 0.0001 |
| wave (Khayelitsha) | 1.0166 | ( 0.9903 - 1.0436 ) | 0.2181 | 1.0165 | ( 0.9865 - 1.0474 ) | 0.2853 |
| sex (Male) | 0.9617 | ( 0.9367 - 0.9873 ) | 0.0036 | 0.9634 | ( 0.9350 - 0.9926 ) | 0.0144 |
| age (6 - 9 yrs) | 1.0216 | ( 0.9953 - 1.0486 ) | 0.1082 | 1.0357 | ( 1.0052 - 1.0671 ) | 0.0214 |
| time1 (control) | 1.0340 | ( 0.9940 - 1.0755 ) | 0.0964 | 1.0262 | ( 0.9842 - 1.0700 ) | 0.2256 |
| time2 (control) | 1.0731 | ( 1.0322 - 1.1157 ) | 0.0004 | 1.0651 | ( 1.0219 - 1.1100 ) | 0.0028 |
| armtime1 | 1.0825 | ( 1.0361 - 1.1309 ) | 0.0004 | 1.1376 | ( 1.0812 - 1.1969 ) | < 0.0001 |
| armtime2 | 1.0034 | ( 0.9604 - 1.0484 ) | 0.8790 | 1.0221 | ( 0.9699 - 1.0772 ) | 0.4135 |
| time1 (intervention) | 1.1221 | ( 1.0827 - 1.1629 ) | < 0.0001 | 1.1674 | ( 1.1132 - 1.2242 ) | < 0.0001 |
| time2 (intervention) | 1.0784 | ( 1.0402 - 1.1180 ) | < 0.0001 | 1.0887 | ( 1.0355 - 1.1445 ) | 0.0009 |

Table A.4: (Unimputed Data)

| | Intention to Treat | | | Per-Protocol | | |
|---|---|---|---|---|---|---|
| | Exp(Beta) | 95% CI | P-value | Exp(Beta) | 95% CI | P-value |
| (Intercept) | 26.2368 | ( 25.4210 - 27.0787 ) | < 0.0001 | 26.2622 | ( 25.3374 - 27.2207 ) | < 0.0001 |
| wave (Khayelitsha) | 1.0175 | ( 0.9903 - 1.0454 ) | 0.2081 | 1.0165 | ( 0.9856 - 1.0483 ) | 0.2971 |
| sex (Male) | 0.9603 | ( 0.9350 - 0.9863 ) | 0.0031 | 0.9619 | ( 0.9329 - 0.9918 ) | 0.0131 |
| age (6 - 9 yrs) | 1.0240 | ( 0.9968 - 1.0520 ) | 0.0846 | 1.0373 | ( 1.0060 - 1.0695 ) | 0.0193 |
| time1 (control) | 1.0341 | ( 0.9935 - 1.0764 ) | 0.1008 | 1.0263 | ( 0.9837 - 1.0709 ) | 0.2293 |
| time2 (control) | 1.0754 | ( 1.0333 - 1.1193 ) | 0.0004 | 1.0671 | ( 1.0228 - 1.1132 ) | 0.0028 |
| armtime1 | 1.0828 | ( 1.0354 - 1.1322 ) | 0.0005 | 1.1386 | ( 1.0811 - 1.1990 ) | < 0.0001 |
| armtime2 | 0.9997 | ( 0.9548 - 1.0466 ) | 0.9891 | 1.0237 | ( 0.9687 - 1.0818 ) | 0.4047 |
| time1 (intervention) | 1.1197 | ( 1.0771 - 1.1640 ) | < 0.0001 | 1.1686 | ( 1.1134 - 1.2265 ) | < 0.0001 |
| time2 (intervention) | 1.0751 | ( 1.0326 - 1.1193 ) | 0.0005 | 1.0923 | ( 1.0369 - 1.1507 ) | 0.0009 |



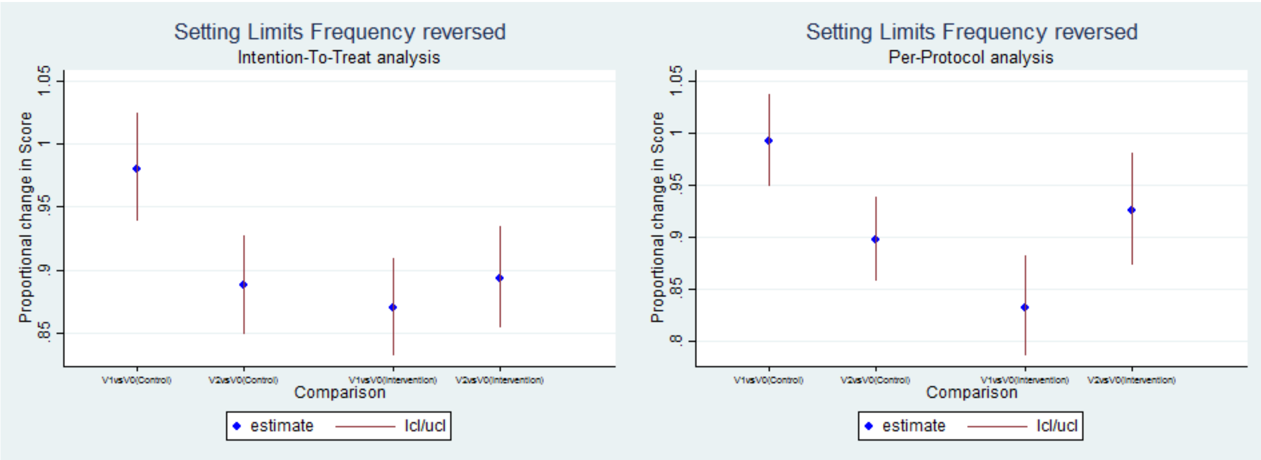

### A.0.3 Setting Limits Frequency Reversed Score

Table A.5: (Imputed Data)

| | Intention to Treat | | | Per-Protocol | | |
|---|---|---|---|---|---|---|
| | Exp(Beta) | 95% CI | P-value | Exp(Beta) | 95% CI | P-value |
| (Intercept) | 26.2643 | ( 25.3828 - 27.1763 ) | < 0.0001 | 25.7578 | ( 24.7961 - 26.7569 ) | < 0.0001 |
| wave (Khayelitsha) | 0.9609 | ( 0.9328 - 0.9898 ) | 0.0083 | 0.9779 | ( 0.9459 - 1.0110 ) | 0.1885 |
| sex (Male) | 1.0070 | ( 0.9778 - 1.0371 ) | 0.6405 | 1.0084 | ( 0.9755 - 1.0423 ) | 0.6222 |
| age (6 - 9 yrs) | 0.9396 | ( 0.9122 - 0.9678 ) | < 0.0001 | 0.9497 | ( 0.9188 - 0.9816 ) | 0.0022 |
| time1 (control) | 0.9804 | ( 0.9387 - 1.0239 ) | 0.3711 | 0.9915 | ( 0.9481 - 1.0370 ) | 0.7107 |
| time2 (control) | 0.8877 | ( 0.8496 - 0.9274 ) | < 0.0001 | 0.8967 | ( 0.8571 - 0.9382 ) | < 0.0001 |
| armtime1 | 0.8872 | ( 0.8435 - 0.9332 ) | < 0.0001 | 0.8376 | ( 0.7891 - 0.8890 ) | < 0.0001 |
| armtime2 | 1.0072 | ( 0.9570 - 1.0600 ) | 0.7832 | 1.0320 | ( 0.9713 - 1.0965 ) | 0.3086 |
| time1 (intervention) | 0.8699 | ( 0.8322 - 0.9093 ) | < 0.0001 | 0.8323 | ( 0.7859 - 0.8814 ) | < 0.0001 |
| time2 (intervention) | 0.8935 | ( 0.8546 - 0.9341 ) | < 0.0001 | 0.9249 | ( 0.8728 - 0.9800 ) | 0.0082 |

Table A.6: (Unimputed Data)

| | Intention to Treat | | | Per-Protocol | | |
|---|---|---|---|---|---|---|
| | Exp(Beta) | 95% CI | P-value | Exp(Beta) | 95% CI | P-value |
| (Intercept) | 26.3187 | ( 25.4318 - 27.2365 ) | < 0.0001 | 25.7797 | ( 24.8180 - 26.7785 ) | < 0.0001 |
| wave (Khayelitsha) | 0.9586 | ( 0.9302 - 0.9879 ) | 0.0060 | 0.9769 | ( 0.9446 - 1.0103 ) | 0.1733 |
| sex (Male) | 1.0075 | ( 0.9780 - 1.0379 ) | 0.6219 | 1.0087 | ( 0.9757 - 1.0429 ) | 0.6087 |
| age (6 - 9 yrs) | 0.9371 | ( 0.9094 - 0.9656 ) | < 0.0001 | 0.9495 | ( 0.9184 - 0.9817 ) | 0.0025 |
| time1 (control) | 0.9809 | ( 0.9393 - 1.0244 ) | 0.3835 | 0.9918 | ( 0.9485 - 1.0372 ) | 0.7190 |
| time2 (control) | 0.8853 | ( 0.8467 - 0.9256 ) | < 0.0001 | 0.8943 | ( 0.8543 - 0.9362 ) | < 0.0001 |
| armtime1 | 0.8851 | ( 0.8412 - 0.9313 ) | < 0.0001 | 0.8357 | ( 0.7871 - 0.8873 ) | < 0.0001 |
| armtime2 | 1.0057 | ( 0.9544 - 1.0597 ) | 0.8327 | 1.0318 | ( 0.9696 - 1.0980 ) | 0.3245 |
| time1 (intervention) | 0.8681 | ( 0.8300 - 0.9079 ) | < 0.0001 | 0.8308 | ( 0.7839 - 0.8806 ) | < 0.0001 |
| time2 (intervention) | 0.8902 | ( 0.8507 - 0.9316 ) | < 0.0001 | 0.9226 | ( 0.8692 - 0.9794 ) | 0.0085 |



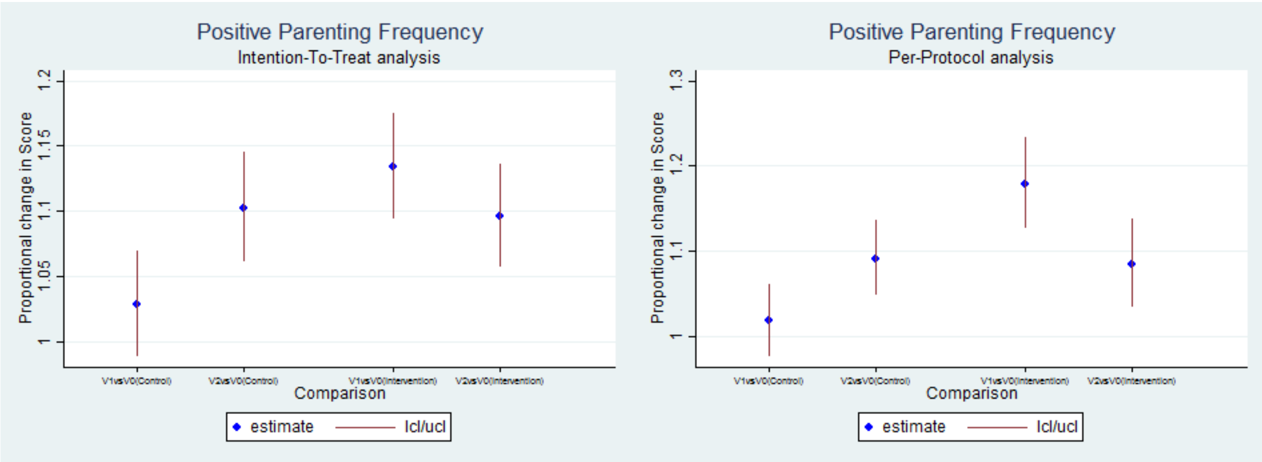

## A.0.4 Positive Parenting Frequency Score

Table A.7: (Imputed Data)

| | Intention to Treat | | | Per-Protocol | | |
|---|---|---|---|---|---|---|
| | Exp(Beta) | 95% CI | P-value | Exp(Beta) | 95% CI | P-value |
| (Intercept) | 47.4957 | ( 46.0463 - 48.9907 ) | < 0.0001 | 48.0720 | ( 46.4101 - 49.7935 ) | < 0.0001 |
| wave (Khayelitsha) | 1.0287 | ( 1.0024 - 1.0556 ) | 0.0325 | 1.0220 | ( 0.9922 - 1.0528 ) | 0.1497 |
| sex (Male) | 0.9771 | ( 0.9525 - 1.0024 ) | 0.0759 | 0.9781 | ( 0.9498 - 1.0072 ) | 0.1393 |
| age (6 - 9 yrs) | 1.0410 | ( 1.0145 - 1.0682 ) | 0.0022 | 1.0430 | ( 1.0127 - 1.0741 ) | 0.0050 |
| time1 (control) | 1.0284 | ( 0.9885 - 1.0698 ) | 0.1653 | 1.0178 | ( 0.9761 - 1.0613 ) | 0.4078 |
| time2 (control) | 1.1027 | ( 1.0615 - 1.1456 ) | < 0.0001 | 1.0913 | ( 1.0480 - 1.1363 ) | < 0.0001 |
| armtime1 | 1.1016 | ( 1.0547 - 1.1506 ) | < 0.0001 | 1.1591 | ( 1.1022 - 1.2190 ) | < 0.0001 |
| armtime2 | 0.9927 | ( 0.9511 - 1.0360 ) | 0.7354 | 0.9923 | ( 0.9418 - 1.0454 ) | 0.7713 |
| time1 (intervention) | 1.1338 | ( 1.0943 - 1.1747 ) | < 0.0001 | 1.1790 | ( 1.1274 - 1.2330 ) | < 0.0001 |
| time2 (intervention) | 1.0960 | ( 1.0573 - 1.1362 ) | < 0.0001 | 1.0851 | ( 1.0347 - 1.1380 ) | 0.0008 |

Table A.8: (Unimputed Data)

| | Intention to Treat | | | Per-Protocol | | |
|---|---|---|---|---|---|---|
| | Exp(Beta) | 95% CI | P-value | Exp(Beta) | 95% CI | P-value |
| (Intercept) | 47.4300 | ( 45.9460 - 48.9620 ) | < 0.0001 | 48.0489 | ( 46.3571 - 49.8024 ) | < 0.0001 |
| wave (Khayelitsha) | 1.0308 | ( 1.0035 - 1.0589 ) | 0.0268 | 1.0239 | ( 0.9930 - 1.0557 ) | 0.1301 |
| sex (Male) | 0.9760 | ( 0.9505 - 1.0022 ) | 0.0724 | 0.9768 | ( 0.9476 - 1.0069 ) | 0.1291 |
| age (6 - 9 yrs) | 1.0432 | ( 1.0157 - 1.0715 ) | 0.0021 | 1.0436 | ( 1.0124 - 1.0758 ) | 0.0061 |
| time1 (control) | 1.0289 | ( 0.9882 - 1.0712 ) | 0.1662 | 1.0184 | ( 0.9760 - 1.0626 ) | 0.4007 |
| time2 (control) | 1.1059 | ( 1.0632 - 1.1503 ) | < 0.0001 | 1.0944 | ( 1.0498 - 1.1409 ) | < 0.0001 |
| armtime1 | 1.1015 | ( 1.0534 - 1.1518 ) | < 0.0001 | 1.1597 | ( 1.1015 - 1.2209 ) | < 0.0001 |
| armtime2 | 0.9897 | ( 0.9461 - 1.0352 ) | 0.6498 | 0.9917 | ( 0.9388 - 1.0476 ) | 0.7658 |
| time1 (intervention) | 1.1345 | ( 1.0943 - 1.1763 ) | < 0.0001 | 1.1802 | ( 1.1276 - 1.2352 ) | < 0.0001 |
| time2 (intervention) | 1.0972 | ( 1.0569 - 1.1389 ) | < 0.0001 | 1.0892 | ( 1.0368 - 1.1443 ) | 0.0007 |



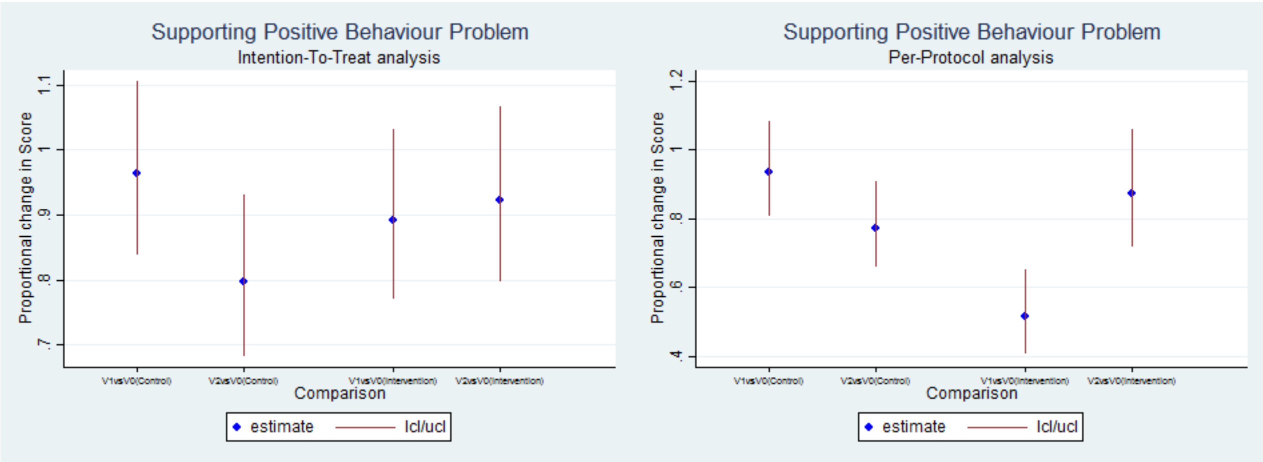

### A.0.5 Supporting Positive Behaviour Problem Score

Table A.9: (Imputed Data)

| | Intention to Treat | | | Per-Protocol | | |
|---|---|---|---|---|---|---|
| | Exp(Beta) | 95% CI | P-value | Exp(Beta) | 95% CI | P-value |
| (Intercept) | 1.9258 | ( 1.7196 - 2.1567 ) | < 0.0001 | 1.9578 | ( 1.7185 - 2.2305 ) | < 0.0001 |
| wave (Khayelitsha) | 0.9661 | ( 0.8764 - 1.0651 ) | 0.4886 | 1.0033 | ( 0.8959 - 1.1237 ) | 0.9539 |
| sex (Male) | 1.3782 | ( 1.2484 - 1.5216 ) | < 0.0001 | 1.4115 | ( 1.2581 - 1.5837 ) | < 0.0001 |
| age (6 - 9 yrs) | 0.9252 | ( 0.8387 - 1.0206 ) | 0.1206 | 0.8908 | ( 0.7948 - 0.9984 ) | 0.0469 |
| time1 (control) | 0.9630 | ( 0.8380 - 1.1066 ) | 0.5945 | 0.9347 | ( 0.8083 - 1.0808 ) | 0.3619 |
| time2 (control) | 0.7968 | ( 0.6818 - 0.9311 ) | 0.0043 | 0.7734 | ( 0.6582 - 0.9087 ) | 0.0018 |
| armtime1 | 0.9283 | ( 0.7852 - 1.0974 ) | 0.3836 | 0.5478 | ( 0.4278 - 0.7013 ) | < 0.0001 |
| armtime2 | 1.1387 | ( 0.9543 - 1.3587 ) | 0.1495 | 1.1364 | ( 0.9115 - 1.4168 ) | 0.2557 |
| time1 (intervention) | 0.8918 | ( 0.7705 - 1.0321 ) | 0.1245 | 0.5152 | ( 0.4065 - 0.6530 ) | < 0.0001 |
| time2 (intervention) | 0.9224 | ( 0.7979 - 1.0663 ) | 0.2749 | 0.8734 | ( 0.7195 - 1.0601 ) | 0.1708 |

Table A.10: (Unimputed Data)

| | Intention to Treat | | | Per-Protocol | | |
|---|---|---|---|---|---|---|
| | Exp(Beta) | 95% CI | P-value | Exp(Beta) | 95% CI | P-value |
| (Intercept) | 1.8793 | ( 1.6637 - 2.1227 ) | < 0.0001 | 1.8769 | ( 1.6343 - 2.1556 ) | < 0.0001 |
| wave (Khayelitsha) | 0.8859 | ( 0.7927 - 0.9900 ) | 0.0325 | 0.9264 | ( 0.8152 - 1.0528 ) | 0.2415 |
| sex (Male) | 1.4507 | ( 1.2962 - 1.6235 ) | < 0.0001 | 1.4825 | ( 1.3019 - 1.6880 ) | < 0.0001 |
| age (6 - 9 yrs) | 1.0019 | ( 0.8971 - 1.1190 ) | 0.9728 | 1.0028 | ( 0.8837 - 1.1380 ) | 0.9653 |
| time1 (control) | 0.8210 | ( 0.7082 - 0.9517 ) | 0.0089 | 0.8043 | ( 0.6896 - 0.9381 ) | 0.0055 |
| time2 (control) | 0.4825 | ( 0.4014 - 0.5800 ) | < 0.0001 | 0.4727 | ( 0.3913 - 0.5710 ) | < 0.0001 |
| armtime1 | 0.7997 | ( 0.6603 - 0.9687 ) | 0.0223 | 0.4782 | ( 0.3619 - 0.6320 ) | < 0.0001 |
| armtime2 | 1.0097 | ( 0.7973 - 1.2787 ) | 0.9360 | 0.9757 | ( 0.7279 - 1.3080 ) | 0.8695 |
| time1 (intervention) | 0.6615 | ( 0.5602 - 0.7812 ) | < 0.0001 | 0.4221 | ( 0.3238 - 0.5502 ) | < 0.0001 |
| time2 (intervention) | 0.5135 | ( 0.4262 - 0.6185 ) | < 0.0001 | 0.4728 | ( 0.3653 - 0.6120 ) | < 0.0001 |



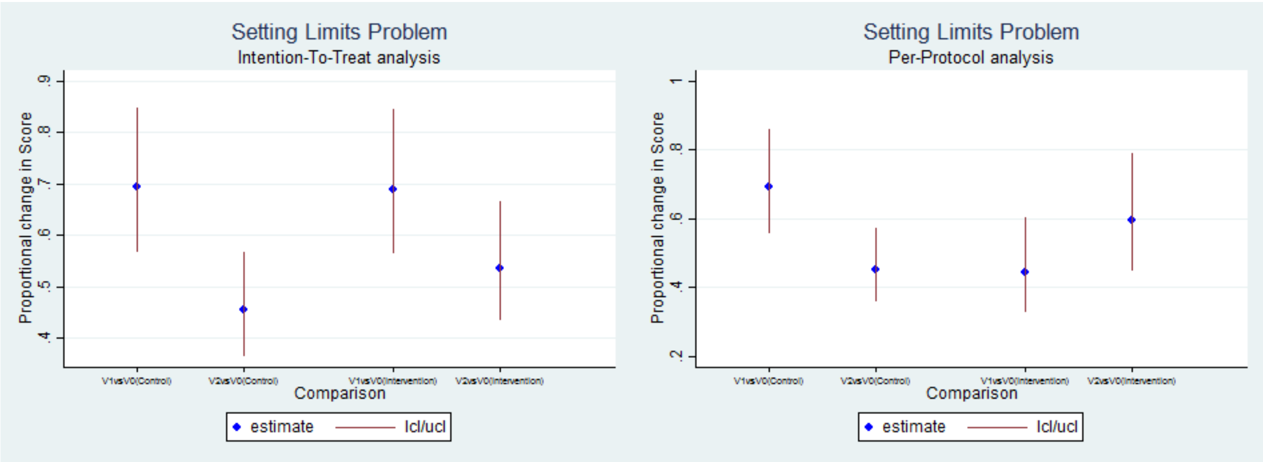

### A.0.6 Setting Limits Problem Score

Table A.11: (Imputed Data)

| | Intention to Treat | | | Per-Protocol | | |
|---|---|---|---|---|---|---|
| | Exp(Beta) | 95% CI | P-value | Exp(Beta) | 95% CI | P-value |
| (Intercept) | 2.2191 | ( 1.8986 - 2.5936 ) | < 0.0001 | 2.1240 | ( 1.7641 - 2.5573 ) | < 0.0001 |
| wave (Khayelitsha) | 0.8847 | ( 0.7684 - 1.0187 ) | 0.0886 | 0.9059 | ( 0.7657 - 1.0719 ) | 0.2497 |
| sex (Male) | 1.3371 | ( 1.1612 - 1.5397 ) | 0.0001 | 1.4177 | ( 1.1983 - 1.6772 ) | < 0.0001 |
| age (6 - 9 yrs) | 0.9490 | ( 0.8245 - 1.0923 ) | 0.4660 | 0.9490 | ( 0.8031 - 1.1214 ) | 0.5387 |
| time1 (control) | 0.6939 | ( 0.5675 - 0.8485 ) | 0.0004 | 0.6928 | ( 0.5575 - 0.8608 ) | 0.0009 |
| time2 (control) | 0.4549 | ( 0.3646 - 0.5676 ) | < 0.0001 | 0.4533 | ( 0.3580 - 0.5741 ) | < 0.0001 |
| armtime1 | 0.9967 | ( 0.7847 - 1.2659 ) | 0.9782 | 0.6348 | ( 0.4579 - 0.8801 ) | 0.0064 |
| armtime2 | 1.1682 | ( 0.8960 - 1.5231 ) | 0.2507 | 1.3092 | ( 0.9497 - 1.8048 ) | 0.1000 |
| time1 (intervention) | 0.6898 | ( 0.5635 - 0.8443 ) | 0.0003 | 0.4438 | ( 0.3266 - 0.6030 ) | < 0.0001 |
| time2 (intervention) | 0.5367 | ( 0.4335 - 0.6644 ) | < 0.0001 | 0.5946 | ( 0.4468 - 0.7913 ) | 0.0004 |

Table A.12: (Unimputed Data)

| | Intention to Treat | | | Per-Protocol | | |
|---|---|---|---|---|---|---|
| | Exp(Beta) | 95% CI | P-value | Exp(Beta) | 95% CI | P-value |
| (Intercept) | 2.3052 | ( 2.0647 - 2.5739 ) | < 0.0001 | 2.1114 | ( 1.8539 - 2.4045 ) | < 0.0001 |
| wave (Khayelitsha) | 0.7887 | ( 0.7090 - 0.8774 ) | < 0.0001 | 0.8171 | ( 0.7219 - 0.9250 ) | 0.0015 |
| sex (Male) | 1.3291 | ( 1.1954 - 1.4778 ) | < 0.0001 | 1.4200 | ( 1.2540 - 1.6080 ) | < 0.0001 |
| age (6 - 9 yrs) | 0.9935 | ( 0.8938 - 1.1042 ) | 0.9031 | 1.0511 | ( 0.9301 - 1.1878 ) | 0.4250 |
| time1 (control) | 0.7155 | ( 0.6230 - 0.8216 ) | < 0.0001 | 0.7140 | ( 0.6167 - 0.8265 ) | < 0.0001 |
| time2 (control) | 0.2803 | ( 0.2277 - 0.3451 ) | < 0.0001 | 0.2802 | ( 0.2263 - 0.3468 ) | < 0.0001 |
| armtime1 | 0.8121 | ( 0.6770 - 0.9742 ) | 0.0253 | 0.5556 | ( 0.4281 - 0.7210 ) | < 0.0001 |
| armtime2 | 1.1663 | ( 0.8931 - 1.5231 ) | 0.2589 | 1.4066 | ( 1.0312 - 1.9187 ) | 0.0317 |
| time1 (intervention) | 0.5782 | ( 0.4904 - 0.6818 ) | < 0.0001 | 0.4033 | ( 0.3124 - 0.5207 ) | < 0.0001 |
| time2 (intervention) | 0.3502 | ( 0.2859 - 0.4291 ) | < 0.0001 | 0.4024 | ( 0.3093 - 0.5235 ) | < 0.0001 |



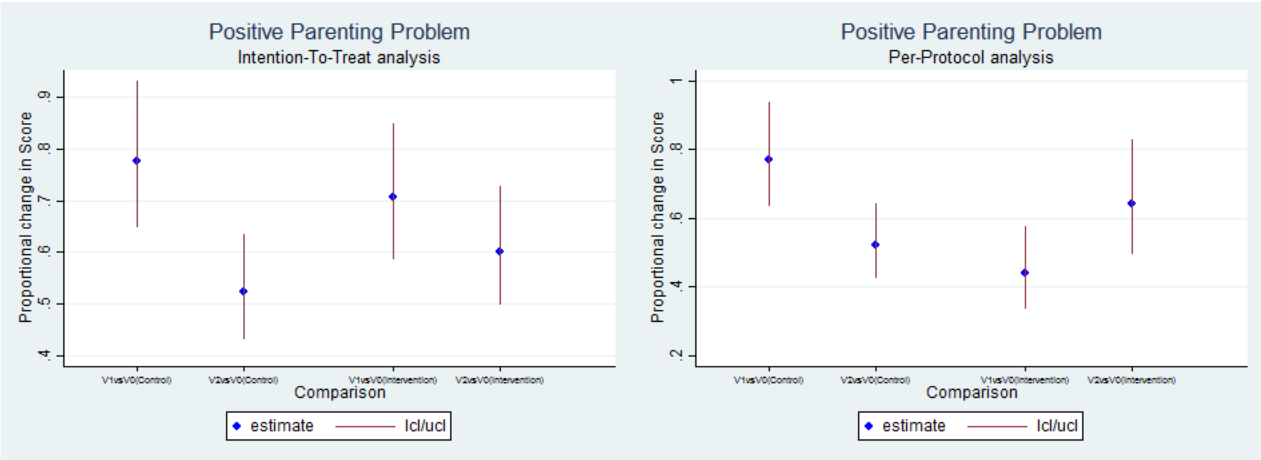

### A.0.7 Positive Parenting Problem Score

Table A.13: (Imputed Data)

| | Intention to Treat | | | Per-Protocol | | |
|---|---|---|---|---|---|---|
| | Exp(Beta) | 95% CI | P-value | Exp(Beta) | 95% CI | P-value |
| (Intercept) | 3.8189 | ( 3.3100 - 4.4059 ) | < 0.0001 | 3.8115 | ( 3.2198 - 4.5119 ) | < 0.0001 |
| wave (Khayelitsha) | 0.9212 | ( 0.8116 - 1.0457 ) | 0.2044 | 0.9697 | ( 0.8342 - 1.1272 ) | 0.6887 |
| sex (Male) | 1.3693 | ( 1.2071 - 1.5532 ) | < 0.0001 | 1.4068 | ( 1.2110 - 1.6343 ) | < 0.0001 |
| age (6 - 9 yrs) | 0.9396 | ( 0.8280 - 1.0663 ) | 0.3345 | 0.9096 | ( 0.7832 - 1.0565 ) | 0.2150 |
| time1 (control) | 0.7769 | ( 0.6479 - 0.9316 ) | 0.0064 | 0.7709 | ( 0.6330 - 0.9387 ) | 0.0096 |
| time2 (control) | 0.5233 | ( 0.4318 - 0.6341 ) | < 0.0001 | 0.5217 | ( 0.4243 - 0.6415 ) | < 0.0001 |
| armtime1 | 0.9121 | ( 0.7360 - 1.1302 ) | 0.4003 | 0.5660 | ( 0.4247 - 0.7544 ) | 0.0001 |
| armtime2 | 1.1237 | ( 0.8951 - 1.4106 ) | 0.3150 | 1.2276 | ( 0.9283 - 1.6233 ) | 0.1503 |
| time1 (intervention) | 0.7055 | ( 0.5861 - 0.8493 ) | 0.0002 | 0.4409 | ( 0.3363 - 0.5779 ) | < 0.0001 |
| time2 (intervention) | 0.6010 | ( 0.4972 - 0.7264 ) | < 0.0001 | 0.6405 | ( 0.4965 - 0.8264 ) | 0.0006 |

Table A.14: (Unimputed Data)

| | Intention to Treat | | | Per-Protocol | | |
|---|---|---|---|---|---|---|
| | Exp(Beta) | 95% CI | P-value | Exp(Beta) | 95% CI | P-value |
| (Intercept) | 3.7366 | ( 3.3610 - 4.1541 ) | < 0.0001 | 3.5804 | ( 3.1653 - 4.0499 ) | < 0.0001 |
| wave (Khayelitsha) | 0.8268 | ( 0.7495 - 0.9120 ) | 0.0002 | 0.8845 | ( 0.7888 - 0.9919 ) | 0.0363 |
| sex (Male) | 1.3980 | ( 1.2678 - 1.5415 ) | < 0.0001 | 1.4444 | ( 1.2882 - 1.6196 ) | < 0.0001 |
| age (6 - 9 yrs) | 1.0053 | ( 0.9119 - 1.1083 ) | 0.9155 | 1.0325 | ( 0.9219 - 1.1565 ) | 0.5800 |
| time1 (control) | 0.7400 | ( 0.6487 - 0.8441 ) | < 0.0001 | 0.7309 | ( 0.6353 - 0.8408 ) | < 0.0001 |
| time2 (control) | 0.3590 | ( 0.3048 - 0.4227 ) | < 0.0001 | 0.3547 | ( 0.2991 - 0.4208 ) | < 0.0001 |
| armtime1 | 0.8036 | ( 0.6807 - 0.9486 ) | 0.0100 | 0.5264 | ( 0.4186 - 0.6620 ) | < 0.0001 |
| armtime2 | 1.0719 | ( 0.8707 - 1.3195 ) | 0.5131 | 1.1545 | ( 0.8984 - 1.4836 ) | 0.2622 |
| time1 (intervention) | 0.5908 | ( 0.5076 - 0.6878 ) | < 0.0001 | 0.3996 | ( 0.3202 - 0.4986 ) | < 0.0001 |
| time2 (intervention) | 0.4042 | ( 0.3409 - 0.4791 ) | < 0.0001 | 0.4231 | ( 0.3386 - 0.5288 ) | < 0.0001 |



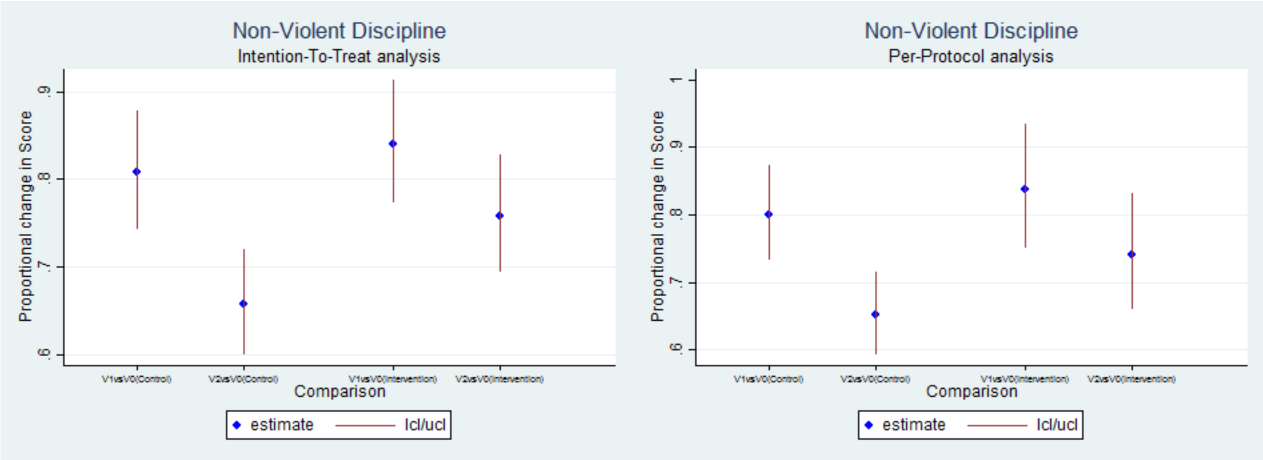

## A.0.8 Non-Violent Discipline Score

Table A.15: (Imputed Data)

| | Intention to Treat | | | Per-Protocol | | |
|---|---|---|---|---|---|---|
| | Exp(Beta) | 95% CI | P-value | Exp(Beta) | 95% CI | P-value |
| (Intercept) | 5.5821 | ( 5.2275 - 5.9608 ) | < 0.0001 | 5.7164 | ( 5.3070 - 6.1573 ) | < 0.0001 |
| wave (Khayelitsha) | 1.0822 | ( 1.0211 - 1.1470 ) | 0.0077 | 1.0273 | ( 0.9605 - 1.0987 ) | 0.4322 |
| sex (Male) | 1.0294 | ( 0.9721 - 1.0900 ) | 0.3215 | 1.0494 | ( 0.9825 - 1.1210 ) | 0.1516 |
| age (6 - 9 yrs) | 1.1954 | ( 1.1270 - 1.2679 ) | < 0.0001 | 1.2010 | ( 1.1238 - 1.2834 ) | < 0.0001 |
| time1 (control) | 0.8075 | ( 0.7423 - 0.8785 ) | < 0.0001 | 0.8005 | ( 0.7332 - 0.8739 ) | < 0.0001 |
| time2 (control) | 0.6568 | ( 0.5996 - 0.7194 ) | < 0.0001 | 0.6510 | ( 0.5924 - 0.7155 ) | < 0.0001 |
| armtime1 | 1.0360 | ( 0.9377 - 1.1446 ) | 0.4872 | 1.0338 | ( 0.9164 - 1.1663 ) | 0.5890 |
| armtime2 | 1.1533 | ( 1.0336 - 1.2867 ) | 0.0107 | 1.1366 | ( 0.9981 - 1.2942 ) | 0.0534 |
| time1 (intervention) | 0.8404 | ( 0.7733 - 0.9134 ) | < 0.0001 | 0.8374 | ( 0.7504 - 0.9345 ) | 0.0015 |
| time2 (intervention) | 0.7574 | ( 0.6935 - 0.8273 ) | < 0.0001 | 0.7399 | ( 0.6591 - 0.8307 ) | < 0.0001 |

Table A.16: (Unimputed Data)

| | Intention to Treat | | | Per-Protocol | | |
|---|---|---|---|---|---|---|
| | Exp(Beta) | 95% CI | P-value | Exp(Beta) | 95% CI | P-value |
| (Intercept) | 5.5777 | ( 5.2210 - 5.9588 ) | < 0.0001 | 5.7168 | ( 5.3073 - 6.1580 ) | < 0.0001 |
| wave (Khayelitsha) | 1.0811 | ( 1.0195 - 1.1465 ) | 0.0094 | 1.0233 | ( 0.9568 - 1.0944 ) | 0.5015 |
| sex (Male) | 1.0314 | ( 0.9732 - 1.0931 ) | 0.2975 | 1.0517 | ( 0.9839 - 1.1242 ) | 0.1388 |
| age (6 - 9 yrs) | 1.2000 | ( 1.1318 - 1.2723 ) | < 0.0001 | 1.2045 | ( 1.1268 - 1.2877 ) | < 0.0001 |
| time1 (control) | 0.8097 | ( 0.7442 - 0.8810 ) | < 0.0001 | 0.8028 | ( 0.7352 - 0.8766 ) | < 0.0001 |
| time2 (control) | 0.6580 | ( 0.5992 - 0.7225 ) | < 0.0001 | 0.6520 | ( 0.5918 - 0.7183 ) | < 0.0001 |
| armtime1 | 1.0397 | ( 0.9397 - 1.1502 ) | 0.4508 | 1.0346 | ( 0.9169 - 1.1673 ) | 0.5816 |
| armtime2 | 1.1669 | ( 1.0432 - 1.3053 ) | 0.0071 | 1.1457 | ( 1.0015 - 1.3108 ) | 0.0481 |
| time1 (intervention) | 0.8466 | ( 0.7779 - 0.9215 ) | 0.0001 | 0.8419 | ( 0.7541 - 0.9401 ) | 0.0023 |
| time2 (intervention) | 0.7676 | ( 0.7024 - 0.8390 ) | < 0.0001 | 0.7469 | ( 0.6635 - 0.8408 ) | < 0.0001 |



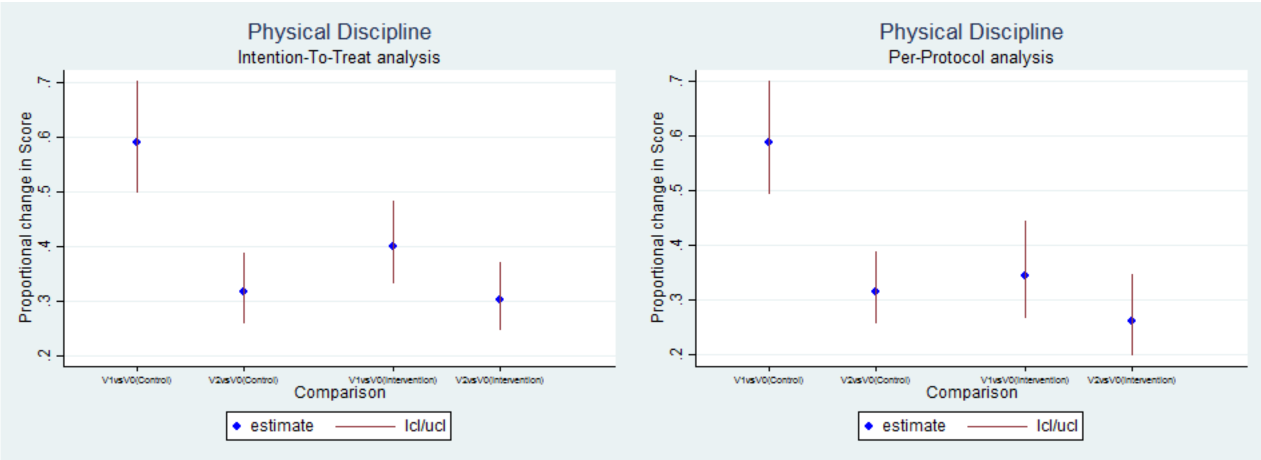

### A.0.9 Physical Discipline Score

Table A.17: (Imputed Data)

| | Intention to Treat | | | Per-Protocol | | |
|---|---|---|---|---|---|---|
| | Exp(Beta) | 95% CI | P-value | Exp(Beta) | 95% CI | P-value |
| (Intercept) | 3.2980 | ( 2.8746 - 3.7839 ) | < 0.0001 | 3.3440 | ( 2.8760 - 3.8882 ) | < 0.0001 |
| wave (Khayelitsha) | 0.9848 | ( 0.8676 - 1.1178 ) | 0.8122 | 0.9406 | ( 0.8164 - 1.0836 ) | 0.3966 |
| sex (Male) | 1.3559 | ( 1.1953 - 1.5381 ) | < 0.0001 | 1.3248 | ( 1.1509 - 1.5249 ) | 0.0001 |
| age (6 - 9 yrs) | 1.0968 | ( 0.9668 - 1.2443 ) | 0.1513 | 1.1526 | ( 1.0017 - 1.3263 ) | 0.0473 |
| time1 (control) | 0.5909 | ( 0.4965 - 0.7033 ) | < 0.0001 | 0.5872 | ( 0.4921 - 0.7006 ) | < 0.0001 |
| time2 (control) | 0.3172 | ( 0.2587 - 0.3888 ) | < 0.0001 | 0.3147 | ( 0.2562 - 0.3866 ) | < 0.0001 |
| armtime1 | 0.6772 | ( 0.5430 - 0.8446 ) | 0.0005 | 0.5856 | ( 0.4446 - 0.7714 ) | 0.0001 |
| armtime2 | 0.9530 | ( 0.7379 - 1.2308 ) | 0.7121 | 0.8293 | ( 0.6018 - 1.1427 ) | 0.2525 |
| time1 (intervention) | 0.4002 | ( 0.3312 - 0.4835 ) | < 0.0001 | 0.3439 | ( 0.2660 - 0.4446 ) | < 0.0001 |
| time2 (intervention) | 0.3023 | ( 0.2464 - 0.3708 ) | < 0.0001 | 0.2610 | ( 0.1967 - 0.3464 ) | < 0.0001 |

Table A.18: (Unimputed Data)

| | Intention to Treat | | | Per-Protocol | | |
|---|---|---|---|---|---|---|
| | Exp(Beta) | 95% CI | P-value | Exp(Beta) | 95% CI | P-value |
| (Intercept) | 3.2863 | ( 2.8758 - 3.7555 ) | < 0.0001 | 3.3331 | ( 2.8807 - 3.8566 ) | < 0.0001 |
| wave (Khayelitsha) | 0.9888 | ( 0.8736 - 1.1192 ) | 0.8589 | 0.9498 | ( 0.8271 - 1.0907 ) | 0.4660 |
| sex (Male) | 1.3703 | ( 1.2117 - 1.5497 ) | < 0.0001 | 1.3348 | ( 1.1639 - 1.5310 ) | < 0.0001 |
| age (6 - 9 yrs) | 1.0907 | ( 0.9641 - 1.2339 ) | 0.1682 | 1.1394 | ( 0.9936 - 1.3065 ) | 0.0623 |
| time1 (control) | 0.5935 | ( 0.5010 - 0.7030 ) | < 0.0001 | 0.5886 | ( 0.4958 - 0.6988 ) | < 0.0001 |
| time2 (control) | 0.3398 | ( 0.2790 - 0.4139 ) | < 0.0001 | 0.3371 | ( 0.2763 - 0.4113 ) | < 0.0001 |
| armtime1 | 0.7230 | ( 0.5821 - 0.8981 ) | 0.0035 | 0.6024 | ( 0.4599 - 0.7892 ) | 0.0003 |
| armtime2 | 0.9583 | ( 0.7451 - 1.2325 ) | 0.7401 | 0.8222 | ( 0.6023 - 1.1224 ) | 0.2182 |
| time1 (intervention) | 0.4291 | ( 0.3563 - 0.5168 ) | < 0.0001 | 0.3546 | ( 0.2757 - 0.4562 ) | < 0.0001 |
| time2 (intervention) | 0.3256 | ( 0.2662 - 0.3983 ) | < 0.0001 | 0.2771 | ( 0.2099 - 0.3659 ) | < 0.0001 |



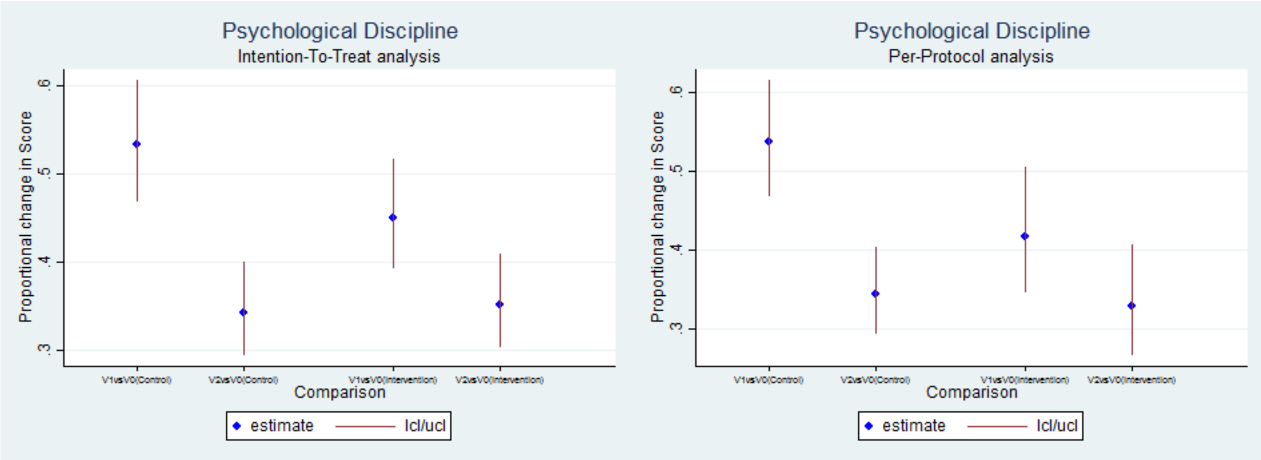

### A.0.10 Psychological Discipline Score

Table A.19: (Imputed Data)

| | Intention to Treat | | | Per-Protocol | | |
|---|---|---|---|---|---|---|
| | Exp(Beta) | 95% CI | P-value | Exp(Beta) | 95% CI | P-value |
| (Intercept) | 5.3492 | ( 4.8421 - 5.9094 ) | < 0.0001 | 5.4648 | ( 4.8758 - 6.1250 ) | < 0.0001 |
| wave (Khayelitsha) | 0.9547 | ( 0.8705 - 1.0469 ) | 0.3244 | 0.8729 | ( 0.7833 - 0.9727 ) | 0.0139 |
| sex (Male) | 1.2339 | ( 1.1257 - 1.3525 ) | < 0.0001 | 1.2479 | ( 1.1218 - 1.3881 ) | < 0.0001 |
| age (6 - 9 yrs) | 1.1210 | ( 1.0236 - 1.2277 ) | 0.0138 | 1.1174 | ( 1.0055 - 1.2418 ) | 0.0393 |
| time1 (control) | 0.5332 | ( 0.4681 - 0.6072 ) | < 0.0001 | 0.5375 | ( 0.4686 - 0.6165 ) | < 0.0001 |
| time2 (control) | 0.3425 | ( 0.2938 - 0.3994 ) | < 0.0001 | 0.3449 | ( 0.2940 - 0.4046 ) | < 0.0001 |
| armtime1 | 0.8440 | ( 0.7161 - 0.9948 ) | 0.0432 | 0.7787 | ( 0.6340 - 0.9564 ) | 0.0171 |
| armtime2 | 1.0263 | ( 0.8492 - 1.2402 ) | 0.7885 | 0.9375 | ( 0.7389 - 1.1895 ) | 0.5952 |
| time1 (intervention) | 0.4500 | ( 0.3921 - 0.5165 ) | < 0.0001 | 0.4182 | ( 0.3458 - 0.5056 ) | < 0.0001 |
| time2 (intervention) | 0.3515 | ( 0.3022 - 0.4089 ) | < 0.0001 | 0.3294 | ( 0.2662 - 0.4075 ) | < 0.0001 |

Table A.20: (Unimputed Data)

| | Intention to Treat | | | Per-Protocol | | |
|---|---|---|---|---|---|---|
| | Exp(Beta) | 95% CI | P-value | Exp(Beta) | 95% CI | P-value |
| (Intercept) | 5.3548 | ( 4.8579 - 5.9026 ) | < 0.0001 | 5.4530 | ( 4.8756 - 6.0988 ) | < 0.0001 |
| wave (Khayelitsha) | 0.9482 | ( 0.8663 - 1.0378 ) | 0.2483 | 0.8701 | ( 0.7827 - 0.9674 ) | 0.0103 |
| sex (Male) | 1.2464 | ( 1.1394 - 1.3635 ) | < 0.0001 | 1.2646 | ( 1.1384 - 1.4049 ) | < 0.0001 |
| age (6 - 9 yrs) | 1.1130 | ( 1.0174 - 1.2175 ) | 0.0198 | 1.1068 | ( 0.9970 - 1.2288 ) | 0.0576 |
| time1 (control) | 0.5366 | ( 0.4724 - 0.6096 ) | < 0.0001 | 0.5402 | ( 0.4719 - 0.6184 ) | < 0.0001 |
| time2 (control) | 0.3557 | ( 0.3074 - 0.4115 ) | < 0.0001 | 0.3575 | ( 0.3069 - 0.4166 ) | < 0.0001 |
| armtime1 | 0.8642 | ( 0.7354 - 1.0157 ) | 0.0770 | 0.7880 | ( 0.6429 - 0.9659 ) | 0.0222 |
| armtime2 | 1.0516 | ( 0.8744 - 1.2646 ) | 0.5933 | 0.9751 | ( 0.7739 - 1.2287 ) | 0.8311 |
| time1 (intervention) | 0.4638 | ( 0.4052 - 0.5308 ) | < 0.0001 | 0.4253 | ( 0.3527 - 0.5129 ) | < 0.0001 |
| time2 (intervention) | 0.3740 | ( 0.3234 - 0.4325 ) | < 0.0001 | 0.3527 | ( 0.2874 - 0.4327 ) | < 0.0001 |



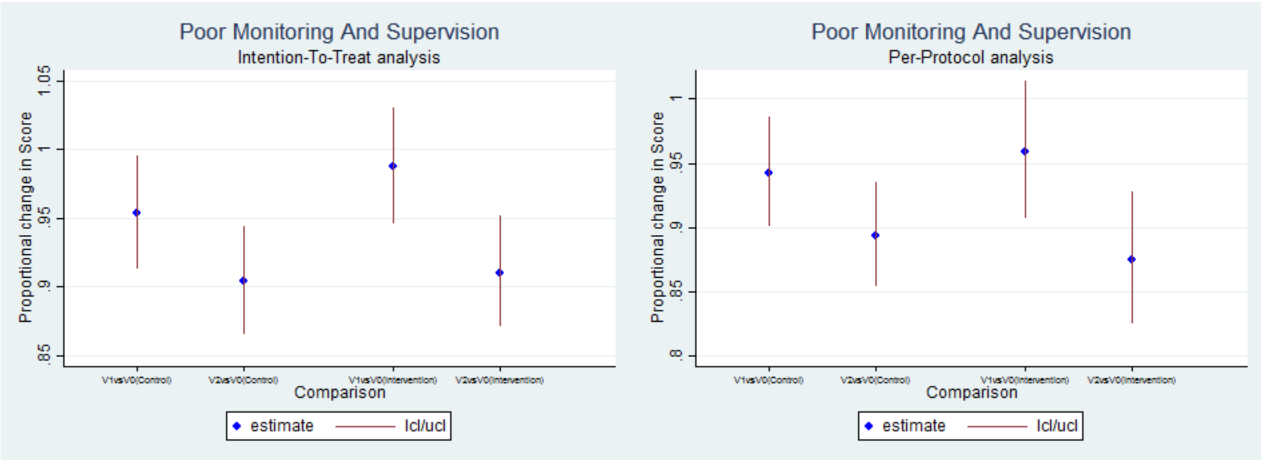

### A.0.11 Poor Monitoring And Supervision Score

Table A.21: (Imputed Data)

| | Intention to Treat | | | Per-Protocol | | |
|---|---|---|---|---|---|---|
| | Exp(Beta) | 95% CI | P-value | Exp(Beta) | 95% CI | P-value |
| (Intercept) | 20.7684 | ( 20.0797 - 21.4807 ) | < 0.0001 | 21.2159 | ( 20.4210 - 22.0418 ) | < 0.0001 |
| wave (Khayelitsha) | 1.0044 | ( 0.9753 - 1.0344 ) | 0.7704 | 0.9909 | ( 0.9582 - 1.0246 ) | 0.5915 |
| sex (Male) | 1.0282 | ( 0.9988 - 1.0585 ) | 0.0606 | 1.0170 | ( 0.9837 - 1.0515 ) | 0.3206 |
| age (6 - 9 yrs) | 1.0899 | ( 1.0585 - 1.1222 ) | < 0.0001 | 1.1024 | ( 1.0662 - 1.1399 ) | < 0.0001 |
| time1 (control) | 0.9533 | ( 0.9134 - 0.9950 ) | 0.0284 | 0.9424 | ( 0.9012 - 0.9856 ) | 0.0095 |
| time2 (control) | 0.9039 | ( 0.8653 - 0.9442 ) | < 0.0001 | 0.8936 | ( 0.8538 - 0.9353 ) | < 0.0001 |
| armtime1 | 1.0359 | ( 0.9858 - 1.0885 ) | 0.1635 | 1.0178 | ( 0.9586 - 1.0806 ) | 0.5644 |
| armtime2 | 1.0071 | ( 0.9559 - 1.0611 ) | 0.7895 | 0.9788 | ( 0.9183 - 1.0434 ) | 0.5115 |
| time1 (intervention) | 0.9875 | ( 0.9463 - 1.0305 ) | 0.5620 | 0.9592 | ( 0.9071 - 1.0143 ) | 0.1439 |
| time2 (intervention) | 0.9103 | ( 0.8709 - 0.9516 ) | < 0.0001 | 0.8748 | ( 0.8247 - 0.9279 ) | < 0.0001 |

Table A.22: (Unimputed Data)

| | Intention to Treat | | | Per-Protocol | | |
|---|---|---|---|---|---|---|
| | Exp(Beta) | 95% CI | P-value | Exp(Beta) | 95% CI | P-value |
| (Intercept) | 20.7268 | ( 20.0280 - 21.4500 ) | < 0.0001 | 21.1548 | ( 20.3551 - 21.9859 ) | < 0.0001 |
| wave (Khayelitsha) | 1.0034 | ( 0.9738 - 1.0338 ) | 0.8261 | 0.9907 | ( 0.9574 - 1.0251 ) | 0.5915 |
| sex (Male) | 1.0301 | ( 1.0001 - 1.0610 ) | 0.0500 | 1.0182 | ( 0.9844 - 1.0533 ) | 0.2958 |
| age (6 - 9 yrs) | 1.0943 | ( 1.0622 - 1.1274 ) | < 0.0001 | 1.1065 | ( 1.0697 - 1.1446 ) | < 0.0001 |
| time1 (control) | 0.9555 | ( 0.9151 - 0.9977 ) | 0.0392 | 0.9451 | ( 0.9033 - 0.9888 ) | 0.0147 |
| time2 (control) | 0.9066 | ( 0.8669 - 0.9481 ) | < 0.0001 | 0.8965 | ( 0.8556 - 0.9393 ) | < 0.0001 |
| armtime1 | 1.0391 | ( 0.9881 - 1.0929 ) | 0.1359 | 1.0150 | ( 0.9553 - 1.0785 ) | 0.6303 |
| armtime2 | 1.0132 | ( 0.9608 - 1.0685 ) | 0.6289 | 0.9852 | ( 0.9235 - 1.0511 ) | 0.6527 |
| time1 (intervention) | 0.9967 | ( < 0.0001 - Inf ) | 1.0000 | 0.9594 | ( 0.9067 - 1.0152 ) | 0.1513 |
| time2 (intervention) | 0.9237 | ( < 0.0001 - Inf ) | 1.0000 | 0.8833 | ( 0.8321 - 0.9377 ) | 0.0001 |



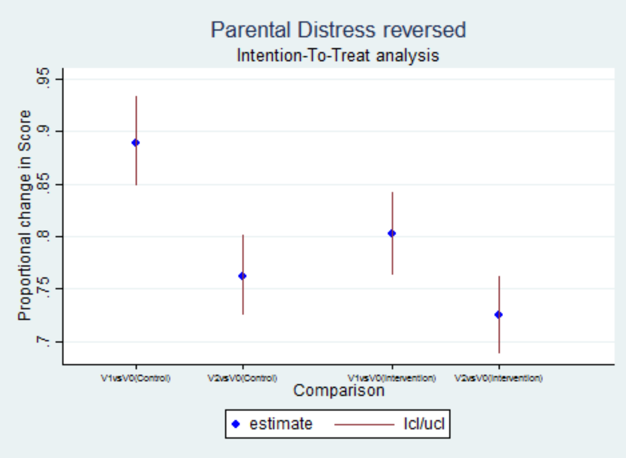

### A.0.12 Parental Distress Reversed Score

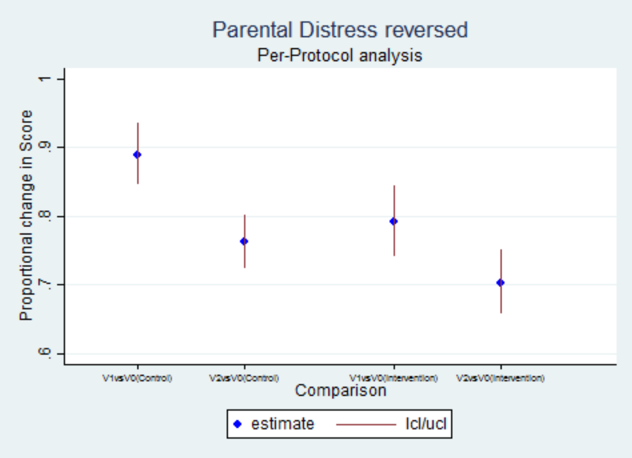
Table A.23: (Imputed Data)

| | Intention to Treat | | | Per-Protocol | | |
|---|---|---|---|---|---|---|
| | Exp(Beta) | 95% CI | P-value | Exp(Beta) | 95% CI | P-value |
| (Intercept) | 24.6793 | ( 23.7727 - 25.6204 ) | < 0.0001 | 24.5036 | ( 23.5021 - 25.5477 ) | < 0.0001 |
| wave (Khayelitsha) | 1.0460 | ( 1.0118 - 1.0815 ) | 0.0081 | 1.0491 | ( 1.0104 - 1.0892 ) | 0.0125 |
| sex (Male) | 1.0139 | ( 0.9811 - 1.0479 ) | 0.4107 | 1.0360 | ( 0.9983 - 1.0753 ) | 0.0618 |
| age (6 - 9 yrs) | 1.0400 | ( 1.0060 - 1.0750 ) | 0.0206 | 1.0266 | ( 0.9891 - 1.0656 ) | 0.1670 |
| time1 (control) | 0.8893 | ( 0.8481 - 0.9326 ) | < 0.0001 | 0.8897 | ( 0.8471 - 0.9345 ) | < 0.0001 |
| time2 (control) | 0.7622 | ( 0.7252 - 0.8010 ) | < 0.0001 | 0.7626 | ( 0.7245 - 0.8026 ) | < 0.0001 |
| armtime1 | 0.9001 | ( 0.8504 - 0.9526 ) | 0.0003 | 0.8822 | ( 0.8238 - 0.9448 ) | 0.0003 |
| armtime2 | 0.9504 | ( 0.8951 - 1.0090 ) | 0.0958 | 0.9204 | ( 0.8559 - 0.9898 ) | 0.0253 |
| time1 (intervention) | 0.8019 | ( 0.7633 - 0.8423 ) | < 0.0001 | 0.7916 | ( 0.7420 - 0.8445 ) | < 0.0001 |
| time2 (intervention) | 0.7245 | ( 0.6883 - 0.7625 ) | < 0.0001 | 0.7029 | ( 0.6573 - 0.7517 ) | < 0.0001 |

Table A.24: (Unimputed Data)

| | Intention to Treat | | | Per-Protocol | | |
|---|---|---|---|---|---|---|
| | Exp(Beta) | 95% CI | P-value | Exp(Beta) | 95% CI | P-value |
| (Intercept) | 24.5117 | ( 23.6154 - 25.4421 ) | < 0.0001 | 24.3715 | ( 23.3790 - 25.4061 ) | < 0.0001 |
| wave (Khayelitsha) | 1.0525 | ( 1.0181 - 1.0881 ) | 0.0027 | 1.0558 | ( 1.0170 - 1.0961 ) | 0.0047 |
| sex (Male) | 1.0191 | ( 0.9860 - 1.0532 ) | 0.2618 | 1.0411 | ( 1.0030 - 1.0806 ) | 0.0345 |
| age (6 - 9 yrs) | 1.0397 | ( 1.0059 - 1.0748 ) | 0.0215 | 1.0234 | ( 0.9861 - 1.0622 ) | 0.2218 |
| time1 (control) | 0.8917 | ( 0.8506 - 0.9347 ) | < 0.0001 | 0.8922 | ( 0.8498 - 0.9367 ) | < 0.0001 |
| time2 (control) | 0.7716 | ( 0.7341 - 0.8110 ) | < 0.0001 | 0.7722 | ( 0.7336 - 0.8129 ) | < 0.0001 |
| armtime1 | 0.9048 | ( 0.8548 - 0.9578 ) | 0.0006 | 0.8841 | ( 0.8259 - 0.9465 ) | 0.0004 |
| armtime2 | 0.9513 | ( 0.8953 - 1.0109 ) | 0.1075 | 0.9139 | ( 0.8490 - 0.9836 ) | 0.0168 |
| time1 (intervention) | 0.8081 | ( 0.7691 - 0.8492 ) | < 0.0001 | 0.7955 | ( 0.7457 - 0.8485 ) | < 0.0001 |
| time2 (intervention) | 0.7335 | ( 0.6967 - 0.7721 ) | < 0.0001 | 0.7057 | ( 0.6591 - 0.7557 ) | < 0.0001 |



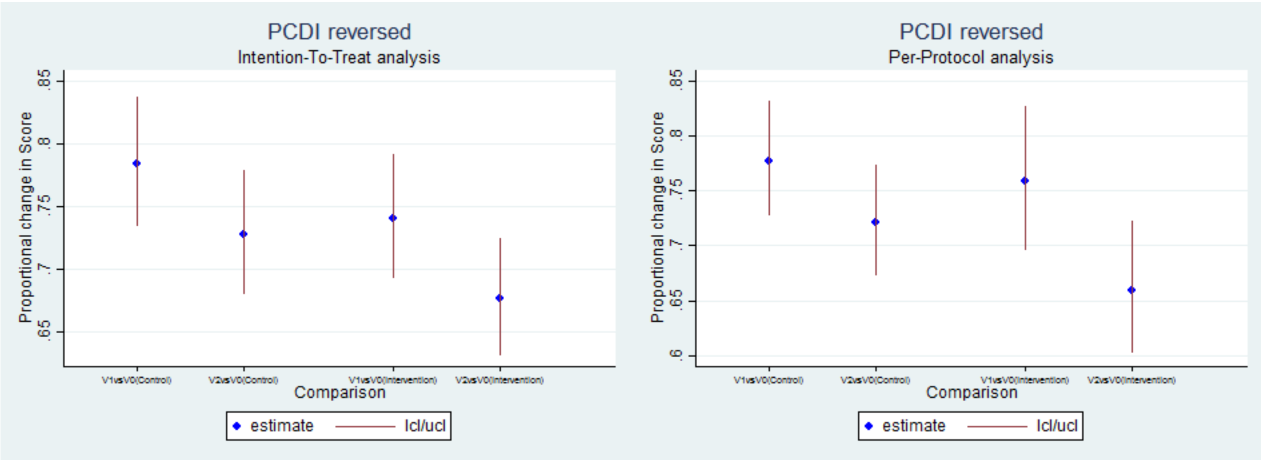

### A.0.13 Parent-Child Dysfunctional Interaction Reverse Score

Table A.25: (Imputed Data)

| | Intention to Treat | | | Per-Protocol | | |
|---|---|---|---|---|---|---|
| | Exp(Beta) | 95% CI | P-value | Exp(Beta) | 95% CI | P-value |
| (Intercept) | 15.6520 | ( 14.8746 - 16.4699 ) | < 0.0001 | 15.5925 | ( 14.7349 - 16.5001 ) | < 0.0001 |
| wave (Khayelitsha) | 1.1027 | ( 1.0537 - 1.1540 ) | < 0.0001 | 1.0919 | ( 1.0379 - 1.1488 ) | 0.0007 |
| sex (Male) | 1.0357 | ( 0.9893 - 1.0842 ) | 0.1338 | 1.0748 | ( 1.0209 - 1.1316 ) | 0.0060 |
| age (6 - 9 yrs) | 1.0783 | ( 1.0307 - 1.1281 ) | 0.0011 | 1.0733 | ( 1.0207 - 1.1285 ) | 0.0058 |
| time1 (control) | 0.7833 | ( 0.7337 - 0.8364 ) | < 0.0001 | 0.7774 | ( 0.7269 - 0.8315 ) | < 0.0001 |
| time2 (control) | 0.7274 | ( 0.6797 - 0.7784 ) | < 0.0001 | 0.7217 | ( 0.6733 - 0.7736 ) | < 0.0001 |
| armtime1 | 0.9444 | ( 0.8729 - 1.0218 ) | 0.1546 | 0.9761 | ( 0.8892 - 1.0715 ) | 0.6109 |
| armtime2 | 0.9292 | ( 0.8554 - 1.0093 ) | 0.0818 | 0.9135 | ( 0.8270 - 1.0090 ) | 0.0745 |
| time1 (intervention) | 0.7398 | ( 0.6917 - 0.7912 ) | < 0.0001 | 0.7588 | ( 0.6961 - 0.8272 ) | < 0.0001 |
| time2 (intervention) | 0.6759 | ( 0.6304 - 0.7246 ) | < 0.0001 | 0.6593 | ( 0.6021 - 0.7219 ) | < 0.0001 |

Table A.26: (Unimputed Data)

| | Intention to Treat | | | Per-Protocol | | |
|---|---|---|---|---|---|---|
| | Exp(Beta) | 95% CI | P-value | Exp(Beta) | 95% CI | P-value |
| (Intercept) | 15.5786 | ( 14.7967 - 16.4018 ) | < 0.0001 | 15.5823 | ( 14.7170 - 16.4985 ) | < 0.0001 |
| wave (Khayelitsha) | 1.1146 | ( 1.0647 - 1.1668 ) | < 0.0001 | 1.1043 | ( 1.0493 - 1.1623 ) | 0.0002 |
| sex (Male) | 1.0375 | ( 0.9915 - 1.0856 ) | 0.1124 | 1.0770 | ( 1.0236 - 1.1333 ) | 0.0044 |
| age (6 - 9 yrs) | 1.0794 | ( 1.0313 - 1.1298 ) | 0.0011 | 1.0697 | ( 1.0168 - 1.1254 ) | 0.0095 |
| time1 (control) | 0.7685 | ( 0.7195 - 0.8209 ) | < 0.0001 | 0.7620 | ( 0.7121 - 0.8154 ) | < 0.0001 |
| time2 (control) | 0.7311 | ( 0.6831 - 0.7825 ) | < 0.0001 | 0.7247 | ( 0.6760 - 0.7770 ) | < 0.0001 |
| armtime1 | 0.9703 | ( 0.8960 - 1.0507 ) | 0.4579 | 1.0000 | ( 0.9105 - 1.0983 ) | 0.9997 |
| armtime2 | 0.9271 | ( 0.8530 - 1.0076 ) | 0.0753 | 0.8993 | ( 0.8135 - 0.9941 ) | 0.0384 |
| time1 (intervention) | 0.7457 | ( 0.6968 - 0.7980 ) | < 0.0001 | 0.7620 | ( 0.6989 - 0.8308 ) | < 0.0001 |
| time2 (intervention) | 0.6778 | ( 0.6319 - 0.7271 ) | < 0.0001 | 0.6517 | ( 0.5944 - 0.7146 ) | < 0.0001 |



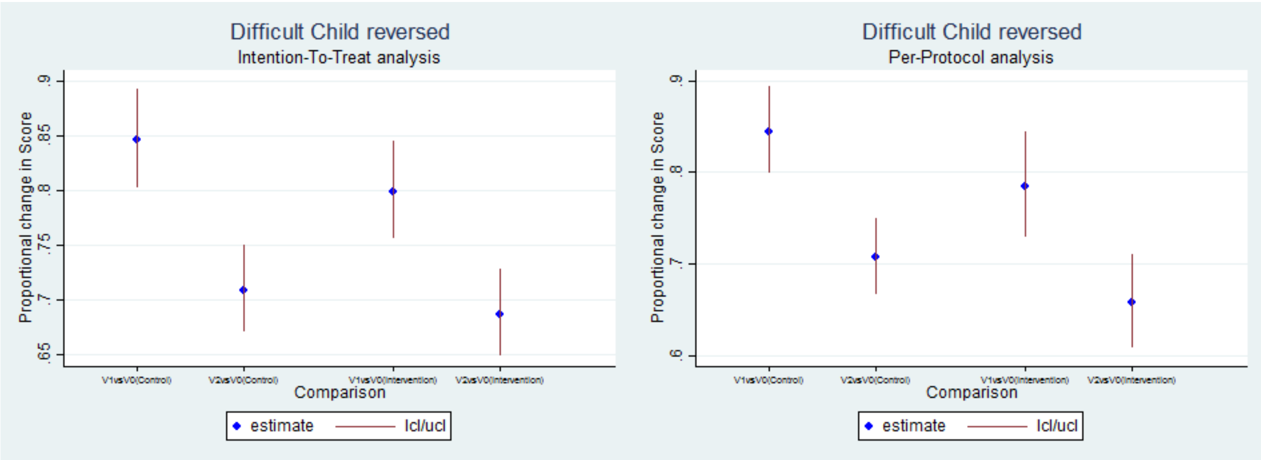

### A.0.14 Difficult Child Reverse Score

Table A.27: (Imputed Data)

| | Intention to Treat | | | Per-Protocol | | |
|---|---|---|---|---|---|---|
| | Exp(Beta) | 95% CI | P-value | Exp(Beta) | 95% CI | P-value |
| (Intercept) | 21.9547 | ( 21.0568 - 22.8909 ) | < 0.0001 | 22.1208 | ( 21.1039 - 23.1866 ) | < 0.0001 |
| wave (Khayelitsha) | 1.0480 | ( 1.0093 - 1.0882 ) | 0.0146 | 1.0382 | ( 0.9945 - 1.0838 ) | 0.0878 |
| sex (Male) | 0.9801 | ( 0.9446 - 1.0170 ) | 0.2862 | 0.9914 | ( 0.9506 - 1.0341 ) | 0.6889 |
| age (6 - 9 yrs) | 1.0474 | ( 1.0087 - 1.0875 ) | 0.0158 | 1.0316 | ( 0.9888 - 1.0762 ) | 0.1505 |
| time1 (control) | 0.8463 | ( 0.8021 - 0.8929 ) | < 0.0001 | 0.8446 | ( 0.7988 - 0.8932 ) | < 0.0001 |
| time2 (control) | 0.7085 | ( 0.6698 - 0.7494 ) | < 0.0001 | 0.7072 | ( 0.6672 - 0.7496 ) | < 0.0001 |
| armtime1 | 0.9389 | ( 0.8804 - 1.0012 ) | 0.0543 | 0.9173 | ( 0.8486 - 0.9915 ) | 0.0296 |
| armtime2 | 0.9691 | ( 0.9055 - 1.0373 ) | 0.3659 | 0.9302 | ( 0.8564 - 1.0103 ) | 0.0861 |
| time1 (intervention) | 0.7988 | ( 0.7557 - 0.8444 ) | < 0.0001 | 0.7846 | ( 0.7292 - 0.8442 ) | < 0.0001 |
| time2 (intervention) | 0.6866 | ( 0.6483 - 0.7272 ) | < 0.0001 | 0.6578 | ( 0.6091 - 0.7104 ) | < 0.0001 |

Table A.28: (Unimputed Data)

| | Intention to Treat | | | Per-Protocol | | |
|---|---|---|---|---|---|---|
| | Exp(Beta) | 95% CI | P-value | Exp(Beta) | 95% CI | P-value |
| (Intercept) | 21.9621 | ( 21.0662 - 22.8961 ) | < 0.0001 | 22.2065 | ( 21.1928 - 23.2687 ) | < 0.0001 |
| wave (Khayelitsha) | 1.0518 | ( 1.0131 - 1.0920 ) | 0.0085 | 1.0415 | ( 0.9981 - 1.0868 ) | 0.0619 |
| sex (Male) | 0.9783 | ( 0.9426 - 1.0152 ) | 0.2458 | 0.9885 | ( 0.9476 - 1.0311 ) | 0.5915 |
| age (6 - 9 yrs) | 1.0455 | ( 1.0072 - 1.0853 ) | 0.0198 | 1.0268 | ( 0.9844 - 1.0711 ) | 0.2194 |
| time1 (control) | 0.8442 | ( 0.8005 - 0.8904 ) | < 0.0001 | 0.8423 | ( 0.7969 - 0.8903 ) | < 0.0001 |
| time2 (control) | 0.7128 | ( 0.6735 - 0.7544 ) | < 0.0001 | 0.7112 | ( 0.6707 - 0.7543 ) | < 0.0001 |
| armtime1 | 0.9423 | ( 0.8839 - 1.0046 ) | 0.0691 | 0.9145 | ( 0.8469 - 0.9876 ) | 0.0231 |
| armtime2 | 0.9766 | ( 0.9113 - 1.0467 ) | 0.5034 | 0.9253 | ( 0.8504 - 1.0068 ) | 0.0722 |
| time1 (intervention) | 0.8006 | ( 0.7574 - 0.8464 ) | < 0.0001 | 0.7790 | ( 0.7244 - 0.8378 ) | < 0.0001 |
| time2 (intervention) | 0.6960 | ( 0.6565 - 0.7379 ) | < 0.0001 | 0.6586 | ( 0.6089 - 0.7124 ) | < 0.0001 |



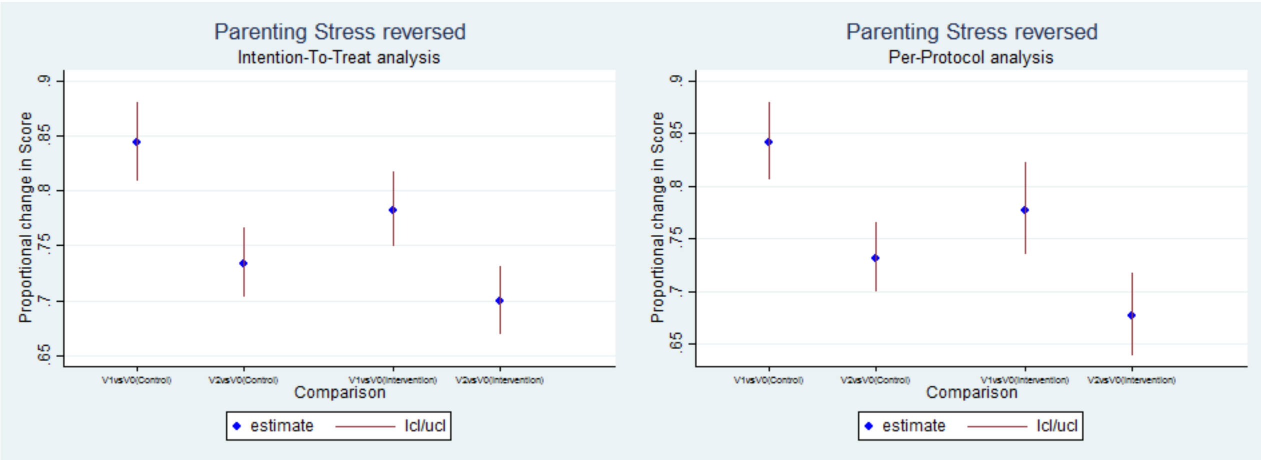

### A.0.15 Parenting Stress Reverse Score

Table A.29: (Imputed Data)

| | Intention to Treat | | | Per-Protocol | | |
|---|---|---|---|---|---|---|
| | Exp(Beta) | 95% CI | P-value | Exp(Beta) | 95% CI | P-value |
| (Intercept) | 62.3178 | ( 60.3027 - 64.4003 ) | < 0.0001 | 62.3329 | ( 60.0739 - 64.6768 ) | < 0.0001 |
| wave (Khayelitsha) | 1.0583 | ( 1.0280 - 1.0895 ) | 0.0001 | 1.0517 | ( 1.0175 - 1.0869 ) | 0.0028 |
| sex (Male) | 1.0061 | ( 0.9776 - 1.0355 ) | 0.6787 | 1.0284 | ( 0.9953 - 1.0625 ) | 0.0930 |
| age (6 - 9 yrs) | 1.0480 | ( 1.0181 - 1.0787 ) | 0.0015 | 1.0353 | ( 1.0022 - 1.0696 ) | 0.0364 |
| time1 (control) | 0.8438 | ( 0.8091 - 0.8800 ) | < 0.0001 | 0.8418 | ( 0.8058 - 0.8793 ) | < 0.0001 |
| time2 (control) | 0.7336 | ( 0.7028 - 0.7658 ) | < 0.0001 | 0.7319 | ( 0.7000 - 0.7653 ) | < 0.0001 |
| armtime1 | 0.9226 | ( 0.8778 - 0.9697 ) | 0.0015 | 0.9123 | ( 0.8595 - 0.9684 ) | 0.0026 |
| armtime2 | 0.9538 | ( 0.9062 - 1.0039 ) | 0.0702 | 0.9257 | ( 0.8704 - 0.9845 ) | 0.0140 |
| time1 (intervention) | 0.7824 | ( 0.7492 - 0.8169 ) | < 0.0001 | 0.7775 | ( 0.7348 - 0.8227 ) | < 0.0001 |
| time2 (intervention) | 0.6997 | ( 0.6694 - 0.7314 ) | < 0.0001 | 0.6773 | ( 0.6392 - 0.7176 ) | < 0.0001 |

Table A.30: (Unimputed Data)

| | Intention to Treat | | | Per-Protocol | | |
|---|---|---|---|---|---|---|
| | Exp(Beta) | 95% CI | P-value | Exp(Beta) | 95% CI | P-value |
| (Intercept) | 62.1862 | ( 60.1998 - 64.2383 ) | < 0.0001 | 62.4087 | ( 60.1723 - 64.7281 ) | < 0.0001 |
| wave (Khayelitsha) | 1.0639 | ( 1.0337 - 1.0950 ) | < 0.0001 | 1.0557 | ( 1.0218 - 1.0908 ) | 0.0012 |
| sex (Male) | 1.0076 | ( 0.9793 - 1.0367 ) | 0.6042 | 1.0306 | ( 0.9978 - 1.0645 ) | 0.0688 |
| age (6 - 9 yrs) | 1.0444 | ( 1.0148 - 1.0748 ) | 0.0032 | 1.0271 | ( 0.9943 - 1.0609 ) | 0.1066 |
| time1 (control) | 0.8405 | ( 0.8065 - 0.8759 ) | < 0.0001 | 0.8384 | ( 0.8031 - 0.8752 ) | < 0.0001 |
| time2 (control) | 0.7393 | ( 0.7084 - 0.7717 ) | < 0.0001 | 0.7376 | ( 0.7055 - 0.7711 ) | < 0.0001 |
| armtime1 | 0.9276 | ( 0.8832 - 0.9742 ) | 0.0028 | 0.9121 | ( 0.8600 - 0.9674 ) | 0.0023 |
| armtime2 | 0.9576 | ( 0.9098 - 1.0079 ) | 0.0976 | 0.9194 | ( 0.8642 - 0.9781 ) | 0.0081 |
| time1 (intervention) | 0.7849 | ( 0.7518 - 0.8194 ) | < 0.0001 | 0.7747 | ( 0.7325 - 0.8194 ) | < 0.0001 |
| time2 (intervention) | 0.7075 | ( 0.6770 - 0.7394 ) | < 0.0001 | 0.6772 | ( 0.6389 - 0.7178 ) | < 0.0001 |



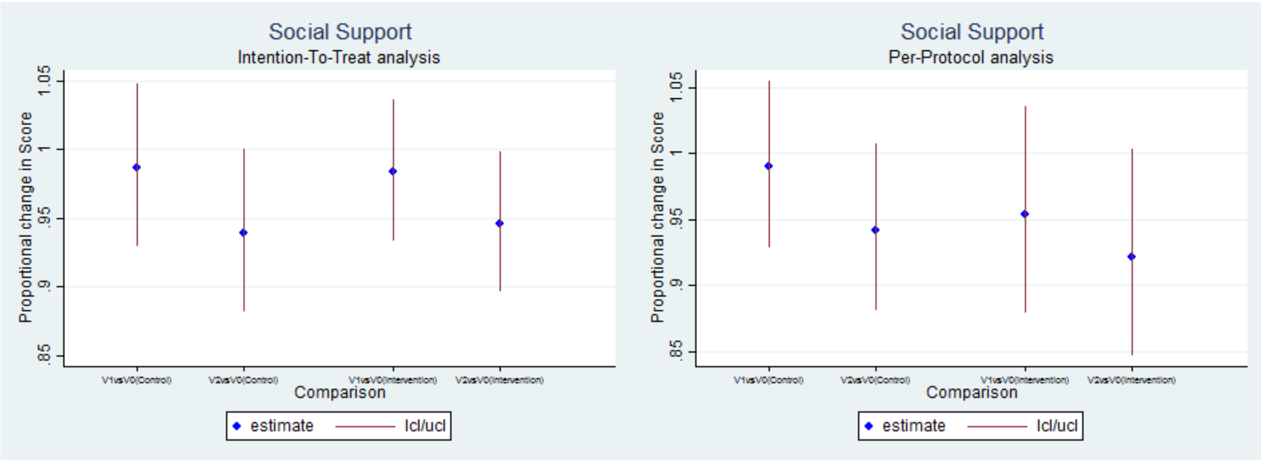

### A.0.16 Social Support Score

Table A.31: (Imputed Data)

| | Intention to Treat | | | Per-Protocol | | |
|---|---|---|---|---|---|---|
| | Exp(Beta) | 95% CI | P-value | Exp(Beta) | 95% CI | P-value |
| (Intercept) | 20.4552 | ( 19.5096 - 21.4465 ) | < 0.0001 | 19.9416 | ( 18.8663 - 21.0782 ) | < 0.0001 |
| wave (Khayelitsha) | 1.0449 | ( 1.0024 - 1.0892 ) | 0.0381 | 1.0602 | ( 1.0100 - 1.1128 ) | 0.0182 |
| sex (Male) | 1.0007 | ( 0.9607 - 1.0424 ) | 0.9725 | 0.9979 | ( 0.9512 - 1.0470 ) | 0.9331 |
| age (6 - 9 yrs) | 0.9937 | ( 0.9532 - 1.0359 ) | 0.7642 | 1.0303 | ( 0.9816 - 1.0814 ) | 0.2269 |
| time1 (control) | 0.9870 | ( 0.9297 - 1.0478 ) | 0.6669 | 0.9898 | ( 0.9284 - 1.0552 ) | 0.7529 |
| time2 (control) | 0.9393 | ( 0.8823 - 1.0000 ) | 0.0501 | 0.9422 | ( 0.8814 - 1.0073 ) | 0.0806 |
| armtime1 | 0.9974 | ( 0.9305 - 1.0692 ) | 0.9418 | 0.9639 | ( 0.8831 - 1.0520 ) | 0.4099 |
| armtime2 | 1.0101 | ( 0.9383 - 1.0874 ) | 0.7895 | 0.9782 | ( 0.8916 - 1.0732 ) | 0.6413 |
| time1 (intervention) | 0.9834 | ( 0.9330 - 1.0365 ) | 0.5320 | 0.9540 | ( 0.8787 - 1.0358 ) | 0.2619 |
| time2 (intervention) | 0.9460 | ( 0.8963 - 0.9984 ) | 0.0436 | 0.9217 | ( 0.8467 - 1.0033 ) | 0.0596 |

Table A.32: (Unimputed Data)

| | Intention to Treat | | | Per-Protocol | | |
|---|---|---|---|---|---|---|
| | Exp(Beta) | 95% CI | P-value | Exp(Beta) | 95% CI | P-value |
| (Intercept) | 20.3962 | ( 19.4233 - 21.4177 ) | < 0.0001 | 19.8941 | ( 18.7937 - 21.0591 ) | < 0.0001 |
| wave (Khayelitsha) | 1.0442 | ( 1.0003 - 1.0900 ) | 0.0482 | 1.0573 | ( 1.0057 - 1.1116 ) | 0.0294 |
| sex (Male) | 1.0005 | ( 0.9588 - 1.0439 ) | 0.9834 | 0.9966 | ( 0.9482 - 1.0475 ) | 0.8934 |
| age (6 - 9 yrs) | 0.9971 | ( 0.9552 - 1.0407 ) | 0.8927 | 1.0345 | ( 0.9843 - 1.0872 ) | 0.1807 |
| time1 (control) | 0.9909 | ( 0.9323 - 1.0532 ) | 0.7685 | 0.9940 | ( 0.9313 - 1.0609 ) | 0.8569 |
| time2 (control) | 0.9350 | ( 0.8762 - 0.9978 ) | 0.0428 | 0.9381 | ( 0.8755 - 1.0050 ) | 0.0691 |
| armtime1 | 0.9949 | ( 0.9258 - 1.0691 ) | 0.8881 | 0.9599 | ( 0.8777 - 1.0498 ) | 0.3695 |
| armtime2 | 1.0088 | ( 0.9335 - 1.0901 ) | 0.8250 | 0.9764 | ( 0.8860 - 1.0761 ) | 0.6298 |
| time1 (intervention) | 0.9858 | ( 0.9262 - 1.0493 ) | 0.6527 | 0.9542 | ( 0.8770 - 1.0381 ) | 0.2746 |
| time2 (intervention) | 0.9432 | ( 0.8839 - 1.0066 ) | 0.0780 | 0.9160 | ( 0.8376 - 1.0017 ) | 0.0544 |



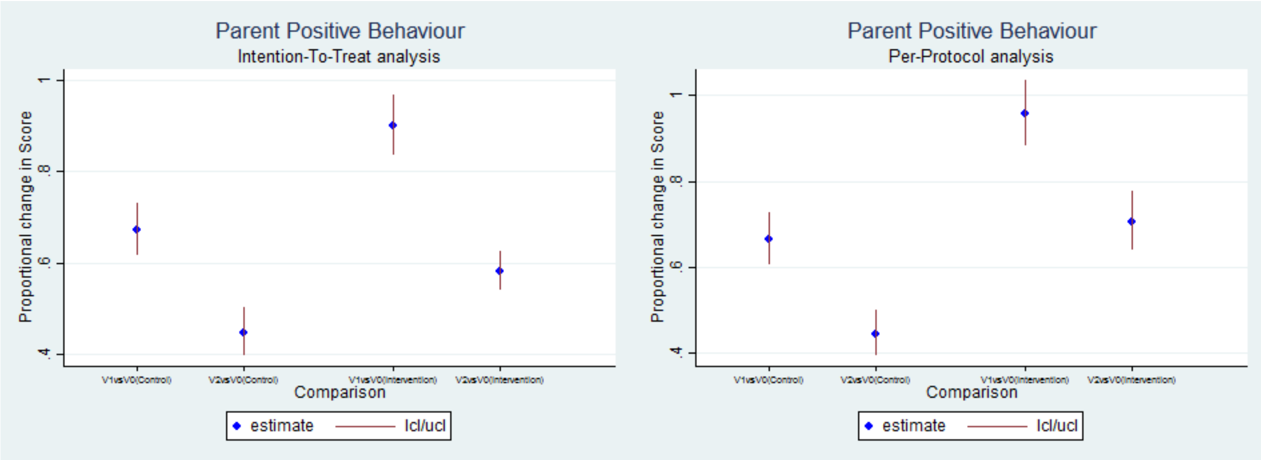

### A.0.17 Parent Positive Behaviour Score

Table A.33: (Imputed Data)

| | Intention to Treat | | | Per-Protocol | | |
|---|---|---|---|---|---|---|
| | Exp(Beta) | 95% CI | P-value | Exp(Beta) | 95% CI | P-value |
| (Intercept) | 18.6914 | ( 17.8432 - 19.5799 ) | < 0.0001 | 18.8286 | ( 17.9346 - 19.7671 ) | < 0.0001 |
| wave (Khayelitsha) | 1.0870 | ( 1.0264 - 1.1512 ) | 0.0044 | 1.0592 | ( 0.9968 - 1.1255 ) | 0.0636 |
| sex (Male) | 0.9070 | ( 0.8576 - 0.9593 ) | 0.0006 | 0.9356 | ( 0.8764 - 0.9987 ) | 0.0458 |
| age (6 - 9 yrs) | 0.5683 | ( 0.5351 - 0.6035 ) | < 0.0001 | 0.5713 | ( 0.5381 - 0.6065 ) | < 0.0001 |
| time1 (control) | 0.6712 | ( 0.6167 - 0.7306 ) | < 0.0001 | 0.6634 | ( 0.6066 - 0.7256 ) | < 0.0001 |
| time2 (control) | 0.4478 | ( 0.3985 - 0.5031 ) | < 0.0001 | 0.4426 | ( 0.3931 - 0.4983 ) | < 0.0001 |
| armtime1 | 1.3185 | ( 1.2033 - 1.4448 ) | < 0.0001 | 1.4368 | ( 1.2949 - 1.5943 ) | < 0.0001 |
| armtime2 | 1.2042 | ( 1.0903 - 1.3299 ) | 0.0002 | 1.3448 | ( 1.1354 - 1.5929 ) | 0.0006 |
| time1 (intervention) | 0.8991 | ( 0.8360 - 0.9671 ) | 0.0042 | 0.9566 | ( 0.8828 - 1.0367 ) | 0.2796 |
| time2 (intervention) | 0.5824 | ( 0.5421 - 0.6256 ) | < 0.0001 | 0.7054 | ( 0.6397 - 0.7778 ) | < 0.0001 |

Table A.34: (Unimputed Data)

| | Intention to Treat | | | Per-Protocol | | |
|---|---|---|---|---|---|---|
| | Exp(Beta) | 95% CI | P-value | Exp(Beta) | 95% CI | P-value |
| (Intercept) | 18.5934 | ( 17.8165 - 19.4042 ) | < 0.0001 | 18.6692 | ( 17.7985 - 19.5825 ) | < 0.0001 |
| wave (Khayelitsha) | 1.0780 | ( 1.0332 - 1.1247 ) | 0.0005 | 1.0541 | ( 1.0043 - 1.1064 ) | 0.0330 |
| sex (Male) | 0.9367 | ( 0.8984 - 0.9766 ) | 0.0021 | 0.9718 | ( 0.9264 - 1.0194 ) | 0.2410 |
| age (6 - 9 yrs) | 0.5659 | ( 0.5409 - 0.5921 ) | < 0.0001 | 0.5691 | ( 0.5408 - 0.5990 ) | < 0.0001 |
| time1 (control) | 0.6764 | ( 0.6345 - 0.7211 ) | < 0.0001 | 0.6695 | ( 0.6266 - 0.7153 ) | < 0.0001 |
| time2 (control) | 0.5316 | ( 0.4912 - 0.5753 ) | < 0.0001 | 0.5259 | ( 0.4851 - 0.5702 ) | < 0.0001 |
| armtime1 | 1.3914 | ( 1.2940 - 1.4962 ) | < 0.0001 | 1.4916 | ( 1.3744 - 1.6188 ) | < 0.0001 |
| armtime2 | 1.2410 | ( 1.1301 - 1.3628 ) | < 0.0001 | 1.3554 | ( 1.2236 - 1.5014 ) | < 0.0001 |
| time1 (intervention) | 0.9604 | ( 0.9088 - 1.0150 ) | 0.1521 | 0.9964 | ( 0.9299 - 1.0676 ) | 0.9182 |
| time2 (intervention) | 0.7189 | ( 0.6729 - 0.7680 ) | < 0.0001 | 0.8667 | ( 0.8007 - 0.9381 ) | 0.0004 |



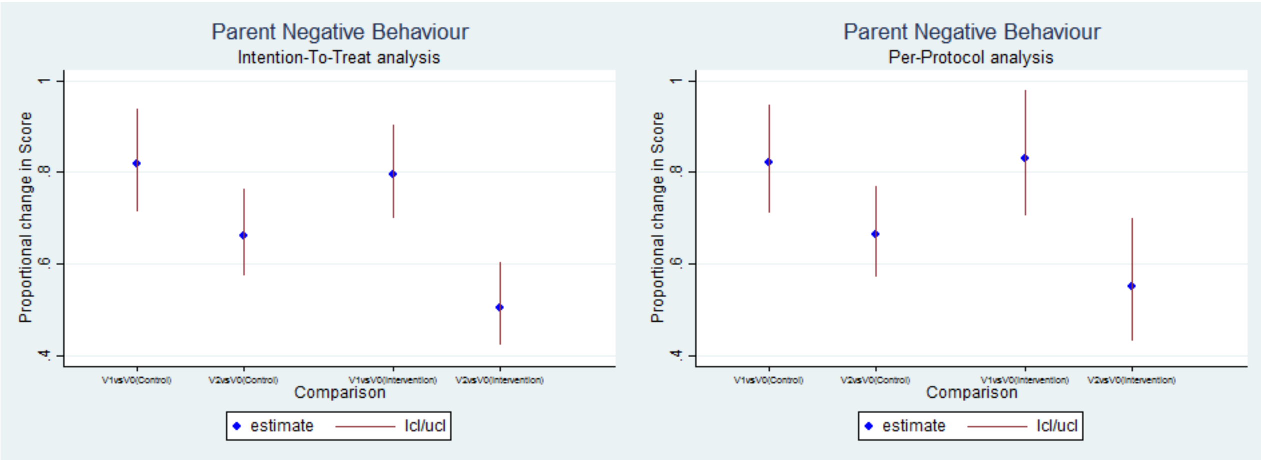

### A.0.18 Parent Negative Behaviour Score

Table A.35: (Imputed Data)

| | Intention to Treat | | | Per-Protocol | | |
|---|---|---|---|---|---|---|
| | Exp(Beta) | 95% CI | P-value | Exp(Beta) | 95% CI | P-value |
| (Intercept) | 3.3354 | ( 3.0250 - 3.6776 ) | < 0.0001 | 3.2939 | ( 2.9396 - 3.6908 ) | < 0.0001 |
| wave (Khayelitsha) | 0.9540 | ( 0.8712 - 1.0447 ) | 0.3094 | 0.9643 | ( 0.8679 - 1.0714 ) | 0.4987 |
| sex (Male) | 1.0873 | ( 0.9834 - 1.2022 ) | 0.1025 | 1.0054 | ( 0.8984 - 1.1252 ) | 0.9251 |
| age (6 - 9 yrs) | 0.8439 | ( 0.7672 - 0.9281 ) | 0.0005 | 0.8942 | ( 0.8023 - 0.9966 ) | 0.0433 |
| time1 (control) | 0.8196 | ( 0.7151 - 0.9394 ) | 0.0043 | 0.8215 | ( 0.7122 - 0.9475 ) | 0.0069 |
| time2 (control) | 0.6621 | ( 0.5735 - 0.7643 ) | < 0.0001 | 0.6636 | ( 0.5722 - 0.7695 ) | < 0.0001 |
| armtime1 | 0.9691 | ( 0.8331 - 1.1273 ) | 0.6839 | 1.0084 | ( 0.8405 - 1.2099 ) | 0.9281 |
| armtime2 | 0.7519 | ( 0.6035 - 0.9369 ) | 0.0111 | 0.8273 | ( 0.6280 - 1.0899 ) | 0.1778 |
| time1 (intervention) | 0.7956 | ( 0.7006 - 0.9035 ) | 0.0004 | 0.8305 | ( 0.7043 - 0.9793 ) | 0.0272 |
| time2 (intervention) | 0.5061 | ( 0.4239 - 0.6043 ) | < 0.0001 | 0.5499 | ( 0.4315 - 0.7008 ) | < 0.0001 |

Table A.36: (Unimputed Data)

| | Intention to Treat | | | Per-Protocol | | |
|---|---|---|---|---|---|---|
| | Exp(Beta) | 95% CI | P-value | Exp(Beta) | 95% CI | P-value |
| (Intercept) | 3.4222 | ( 3.1178 - 3.7562 ) | < 0.0001 | 3.4214 | ( 3.0817 - 3.7985 ) | < 0.0001 |
| wave (Khayelitsha) | 0.9525 | ( 0.8710 - 1.0415 ) | 0.2856 | 0.9717 | ( 0.8774 - 1.0760 ) | 0.5810 |
| sex (Male) | 1.1290 | ( 1.0324 - 1.2345 ) | 0.0078 | 1.0430 | ( 0.9422 - 1.1545 ) | 0.4167 |
| age (6 - 9 yrs) | 0.8490 | ( 0.7760 - 0.9288 ) | 0.0004 | 0.8994 | ( 0.8125 - 0.9955 ) | 0.0407 |
| time1 (control) | 0.8011 | ( 0.7071 - 0.9075 ) | 0.0005 | 0.7975 | ( 0.7000 - 0.9086 ) | 0.0007 |
| time2 (control) | 0.7944 | ( 0.6923 - 0.9115 ) | 0.0010 | 0.7898 | ( 0.6848 - 0.9110 ) | 0.0012 |
| armtime1 | 1.0353 | ( 0.8909 - 1.2031 ) | 0.6510 | 1.0407 | ( 0.8700 - 1.2449 ) | 0.6623 |
| armtime2 | 0.8330 | ( 0.6923 - 1.0022 ) | 0.0528 | 0.8002 | ( 0.6387 - 1.0024 ) | 0.0525 |
| time1 (intervention) | 0.8297 | ( 0.7321 - 0.9404 ) | 0.0035 | 0.8339 | ( 0.7083 - 0.9818 ) | 0.0293 |
| time2 (intervention) | 0.6614 | ( 0.5665 - 0.7723 ) | < 0.0001 | 0.6322 | ( 0.5152 - 0.7759 ) | < 0.0001 |



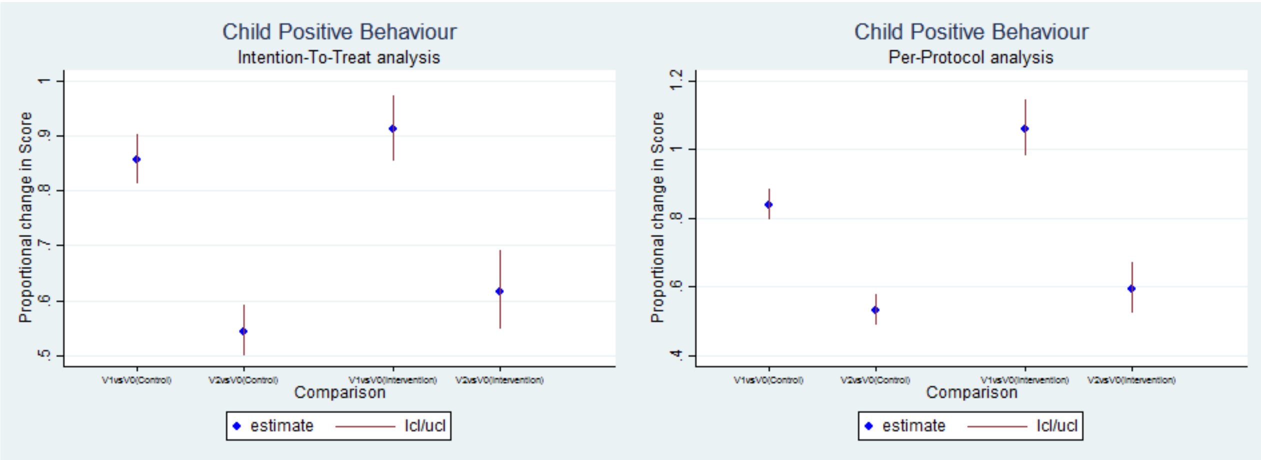

### A.0.19 Child Positive Behaviour Score

Table A.37: (Imputed Data)

| | Intention to Treat | | | Per-Protocol | | |
|---|---|---|---|---|---|---|
| | Exp(Beta) | 95% CI | P-value | Exp(Beta) | 95% CI | P-value |
| (Intercept) | 37.2088 | ( 35.6819 - 38.8010 ) | < 0.0001 | 37.8476 | ( 36.0591 - 39.7248 ) | < 0.0001 |
| wave (Khayelitsha) | 0.8155 | ( 0.7874 - 0.8445 ) | < 0.0001 | 0.8279 | ( 0.7929 - 0.8644 ) | < 0.0001 |
| sex (Male) | 0.9971 | ( 0.9483 - 1.0483 ) | 0.9080 | 0.9818 | ( 0.9399 - 1.0256 ) | 0.4104 |
| age (6 - 9 yrs) | 0.6306 | ( 0.6026 - 0.6600 ) | < 0.0001 | 0.6526 | ( 0.6197 - 0.6871 ) | < 0.0001 |
| time1 (control) | 0.8570 | ( 0.8135 - 0.9028 ) | < 0.0001 | 0.8394 | ( 0.7953 - 0.8860 ) | < 0.0001 |
| time2 (control) | 0.5440 | ( 0.5001 - 0.5918 ) | < 0.0001 | 0.5329 | ( 0.4909 - 0.5784 ) | < 0.0001 |
| armtime1 | 1.0644 | ( 0.9717 - 1.1660 ) | 0.1796 | 1.2623 | ( 1.1496 - 1.3859 ) | < 0.0001 |
| armtime2 | 1.1254 | ( 1.0410 - 1.2166 ) | 0.0030 | 1.1079 | ( 1.0147 - 1.2096 ) | 0.0222 |
| time1 (intervention) | 0.9113 | ( 0.8541 - 0.9723 ) | 0.0050 | 1.0595 | ( 0.9814 - 1.1439 ) | 0.1388 |
| time2 (intervention) | 0.6163 | ( 0.5496 - 0.6911 ) | < 0.0001 | 0.5927 | ( 0.5239 - 0.6706 ) | < 0.0001 |

Table A.38: (Unimputed Data)

| | Intention to Treat | | | Per-Protocol | | |
|---|---|---|---|---|---|---|
| | Exp(Beta) | 95% CI | P-value | Exp(Beta) | 95% CI | P-value |
| (Intercept) | 37.2343 | ( 36.1351 - 38.3669 ) | < 0.0001 | 38.0943 | ( 36.8430 - 39.3880 ) | < 0.0001 |
| wave (Khayelitsha) | 0.8049 | ( 0.7810 - 0.8294 ) | < 0.0001 | 0.8197 | ( 0.7926 - 0.8476 ) | < 0.0001 |
| sex (Male) | 1.0297 | ( 1.0000 - 1.0602 ) | 0.0500 | 1.0103 | ( 0.9776 - 1.0442 ) | 0.5413 |
| age (6 - 9 yrs) | 0.6212 | ( 0.6023 - 0.6407 ) | < 0.0001 | 0.6463 | ( 0.6246 - 0.6687 ) | < 0.0001 |
| time1 (control) | 0.8663 | ( 0.8320 - 0.9020 ) | < 0.0001 | 0.8474 | ( 0.8126 - 0.8838 ) | < 0.0001 |
| time2 (control) | 0.5909 | ( 0.5611 - 0.6223 ) | < 0.0001 | 0.5780 | ( 0.5481 - 0.6094 ) | < 0.0001 |
| armtime1 | 1.1098 | ( 1.0578 - 1.1644 ) | < 0.0001 | 1.2700 | ( 1.2033 - 1.3405 ) | < 0.0001 |
| armtime2 | 1.1675 | ( 1.0943 - 1.2456 ) | < 0.0001 | 1.1270 | ( 1.0432 - 1.2176 ) | 0.0024 |
| time1 (intervention) | 0.9614 | ( 0.9234 - 1.0011 ) | 0.0563 | 1.0763 | ( 1.0252 - 1.1300 ) | 0.0031 |
| time2 (intervention) | 0.6966 | ( 0.6628 - 0.7321 ) | < 0.0001 | 0.6518 | ( 0.6101 - 0.6964 ) | < 0.0001 |



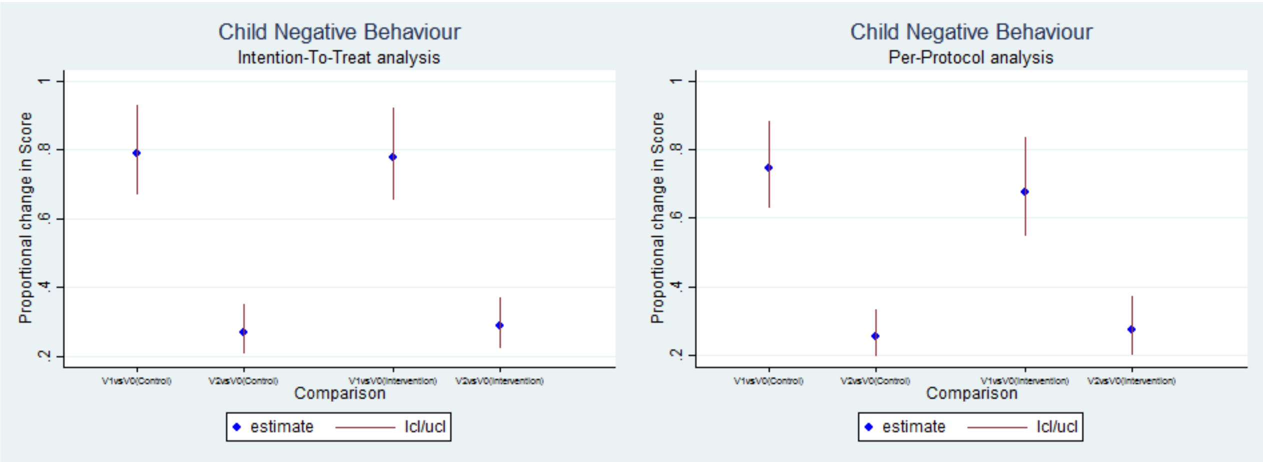

## A.0.20 Child Negative Behaviour Score

Table A.39: (Unimputed Data)

| | Intention to Treat | | | Per-Protocol | | |
|---|---|---|---|---|---|---|
| | Exp(Beta) | 95% CI | P-value | Exp(Beta) | 95% CI | P-value |
| (Intercept) | 2.7109 | ( 2.3975 - 3.0653 ) | < 0.0001 | 3.1377 | ( 2.7553 - 3.5732 ) | < 0.0001 |
| wave (Khayelitsha) | 0.6920 | ( 0.6060 - 0.7901 ) | < 0.0001 | 0.7046 | ( 0.6079 - 0.8168 ) | < 0.0001 |
| sex (Male) | 1.0429 | ( 0.9196 - 1.1827 ) | 0.5127 | 0.9717 | ( 0.8461 - 1.1159 ) | 0.6843 |
| age (6 - 9 yrs) | 0.5536 | ( 0.4808 - 0.6375 ) | < 0.0001 | 0.5621 | ( 0.4812 - 0.6566 ) | < 0.0001 |
| time1 (control) | 0.7878 | ( 0.6683 - 0.9287 ) | 0.0045 | 0.7452 | ( 0.6297 - 0.8818 ) | 0.0006 |
| time2 (control) | 0.2699 | ( 0.2078 - 0.3505 ) | < 0.0001 | 0.2553 | ( 0.1960 - 0.3324 ) | < 0.0001 |
| armtime1 | 0.9910 | ( 0.8084 - 1.2149 ) | 0.9310 | 0.9069 | ( 0.7167 - 1.1476 ) | 0.4159 |
| armtime2 | 1.0719 | ( 0.7611 - 1.5095 ) | 0.6911 | 1.0718 | ( 0.7265 - 1.5811 ) | 0.7267 |
| time1 (intervention) | 0.7775 | ( 0.6553 - 0.9224 ) | 0.0039 | 0.6755 | ( 0.5471 - 0.8340 ) | 0.0003 |
| time2 (intervention) | 0.2884 | ( 0.2248 - 0.3699 ) | < 0.0001 | 0.2736 | ( 0.2006 - 0.3732 ) | < 0.0001 |

Table A.40: (Unimputed Data)

| | Intention to Treat | | | Per-Protocol | | |
|---|---|---|---|---|---|---|
| | Exp(Beta) | 95% CI | P-value | Exp(Beta) | 95% CI | P-value |
| (Intercept) | 2.7180 | ( 2.4112 - 3.0639 ) | < 0.0001 | 3.1701 | ( 2.7878 - 3.6048 ) | < 0.0001 |
| wave (Khayelitsha) | 0.6810 | ( 0.5976 - 0.7761 ) | < 0.0001 | 0.7013 | ( 0.6076 - 0.8095 ) | < 0.0001 |
| sex (Male) | 1.0778 | ( 0.9518 - 1.2206 ) | 0.2374 | 0.9979 | ( 0.8688 - 1.1462 ) | 0.9762 |
| age (6 - 9 yrs) | 0.5472 | ( 0.4778 - 0.6266 ) | < 0.0001 | 0.5635 | ( 0.4858 - 0.6537 ) | < 0.0001 |
| time1 (control) | 0.8055 | ( 0.6833 - 0.9496 ) | 0.0100 | 0.7595 | ( 0.6416 - 0.8991 ) | 0.0014 |
| time2 (control) | 0.3315 | ( 0.2556 - 0.4301 ) | < 0.0001 | 0.3129 | ( 0.2405 - 0.4070 ) | < 0.0001 |
| armtime1 | 1.1003 | ( 0.9019 - 1.3425 ) | 0.3462 | 0.9511 | ( 0.7517 - 1.2035 ) | 0.6764 |
| armtime2 | 1.2009 | ( 0.8558 - 1.6852 ) | 0.2895 | 1.0848 | ( 0.7370 - 1.5967 ) | 0.6799 |
| time1 (intervention) | 0.8816 | ( 0.7466 - 1.0409 ) | 0.1369 | 0.7076 | ( 0.5733 - 0.8732 ) | 0.0013 |
| time2 (intervention) | 0.3974 | ( 0.3098 - 0.5098 ) | < 0.0001 | 0.3400 | ( 0.2490 - 0.4644 ) | < 0.0001 |


Child Negative Behaviour
Intention-To-Treat analysis
Proportional change in Score
.2
.4
.6
.8
1
V1vsV0(Control)
V2vsV0(Control)
V1vsV0(Intervention)
V2vsV0(Intervention)
Comparison
estimate
lcl/ucl
Child Negative Behaviour
Per-Protocol analysis
Proportional change in Score
.2
.4
.6
.8
1
V1vsV0(Control)
V2vsV0(Control)
V1vsV0(Intervention)
V2vsV0(Intervention)
Comparison
estimate
lcl/ucl

# Appendix B

# Analysis Plan

## Randomised Control Trial to Assess the Impact of the Sinovuyo Caring Families Programme: Statistical Analysis Plan for the Immediate Endpoint Analysis

**12 August 2015**




**Authors**: Dr R. Kassanjee, A/Prof F. Little, A/Prof C. Ward & I. Wessels (We acknowledge valuable input from Dr Carl Lombard)

# Overview

## Study Design

A randomised controlled trial is being conducted to assess the effects of the **Sinovuyo Caring Families Programme (SCFP)** – a group-based parent skills training intervention for primary caregivers of children aged 2 to 9 years. The trial aims to measure intervention effects at immediate post-test and 12-month follow-up on four sets of primary endpoints and several secondary endpoints. The four sets of primary endpoints describe different aspects (and measures) of parent and child behaviour: (1) reported child behaviour problems, (2) reported positive parenting, (3) reported harsh parenting, and (4) observed parent and child behaviour. Furthermore, the trial aims to test the processes and theoretical model upon which the intervention was based, and for whom the intervention worked, and to examine programme feasibility including implementation fidelity, exposure and adherence, participant engagement, and satisfaction.

Through targeted sampling and referrals from local agencies, 380 child-caregiver dyads were recruited and screened for trial eligibility. Inclusion criteria for adults included: age 18 years or older; being the primary caregiver of the child participant aged 2 to 9 years; living in the same house as the child for at least 4 nights per week; providing informed consent to participate in the full study including both the intervention (for the intervention group) and at all data collection points; and reporting 15 or more problem behaviours on the Eyberg Child Behaviour Inventory (ECBI) problem scale. Of 330 eligible participant pairs, 310 completed the baseline survey and 296 were randomised into one of the two study arms (the Intervention Arm and the Control Arm). Block randomisation was applied, to ensure a balanced study design with respect to age (2-5 and 6-9 year olds) and sex (boys and girls) of the children.

All primary caregivers received an information pack with the contact details of social and health services that may assist participants with challenges that they may have. In the intervention arm, primary caregivers were invited to participate in the SCFP, a 12-session group-based parent skills training programme implemented by *Clowns Without Borders South Africa*. In the control arm, primary caregivers were only invited to access standard of care services. Both groups were also invited to a finance and health workshop.

The SCFP was conducted in two independent 'waves' – Wave 1 (intervention conducted May to August 2014) took place in Khayelitsha, while Wave 2 (September to December 2014) took place in Nyanga. Within each Wave, each participant that was allocated to the Intervention Arm was also assigned to a Group, based on the caregiver's geographic proximity to intervention sites or availability in terms of time. 11 groups were formed, of 11-17 participants each, and the same group was overseen by the same facilitators throughout the programme. When absent from a group session, or when requiring additional assistance with a concept presented in the session, participants received a home visit from their own group facilitator. Four facilitators (two facilitator pairs) delivered the programme.

Data on primary and secondary endpoints are collected via parental self-report questionnaires and observational assessments at baseline, immediate post-test, and at 12-month follow-up. Process evaluation data is collected via qualitative interviews and focus groups with intervention participants and group leaders, as well as through video-recording programme sessions.

## Scope of Immediate Endpoint Analysis

The objective of this immediate endpoint analysis is to investigate the effectiveness of the SCFP, with respect to the four sets of primary endpoints and a number of secondary endpoints.

Only measurements obtained at baseline and immediately after completion of the 12-session programme (collected within 11 weeks of completion of the programme in Wave 1, and 13 weeks in Wave 2) will be included in this analysis (12 month follow-up data is still being collected).

The analysis **aims to estimate the difference between the two arms in changes from baseline measurements to post-test measurements**. Each of the endpoints will be considered in turn. For each outcome, in the primary analysis, an intention-to-treat approach will be used, and Arm, Wave, Sex (of the child) and Age Group (for the child) will be included in the model. In a secondary analysis, an alternative per-protocol approach will be adopted (participants must have attended a specified minimum number of sessions).

The analysis methodology presented below naturally extends to accommodate the data that will become available after the 12 month follow-up assessments. While the objective of this analysis is to assess the immediate impact of the intervention, the 12-month follow-up data will provide insights into whether this impact strengthens or weakens when considering a later time point. Since the programme is no longer being run, the analysis of this immediate post-test data will not impact the study in any way.

# Analysis Methodology

While the data collectors have been blinded to the group allocation, the statistician will not be blinded so that the data on the allocation of SCFP participants to groups and on the number of sessions attended (for the per-protocol analysis) can be utilised. The integrity of the analysis will be maintained by the formulation of this analysis plan a priori.

## Construction of Outcome Variables

The raw data contains hundreds of fields. For this immediate analysis, a select number of composite scores will therefore be constructed and analysed. Each score is the sum of measurements for several individual items, which are either all Likert scales or all have dichotomous 0/1 responses (some items will be reverse coded before summation, as appropriate). The scores provide sets of primary and secondary outcomes of interest.

Each outcome has a baseline measurement (captured as part of either the screening or subsequent baseline interview/observation process) and post-test measurement (captured as part of the post-test interview/observation process within 13 weeks after the possible intervention). Each outcome will be considered separately in the primary analysis.

The construction of the outcomes is described in Appendix A. The outcomes fall into 8 groups: (1) child behaviour problems, (2) positive parenting, (3) harsh parenting, (4) observed parenting and child behaviour (all primary outcome groups); and (5) monitoring and supervision, (6) depression, (7) parenting stress, and (8) social support (all secondary outcome groups). More specifically, the following scores will be constructed:

| | | |
|---|---|---|
| **Primary Outcomes** | | **Child Behaviour Problems** |
| | **1** | ECBI Frequency |
| | **2** | ECBI Intensity |
| | | **Positive Parenting** |
| | **3** | Supporting Positive Behaviour *Frequency* (subscore) |
| | **4** | Setting Limits *Frequency* (subscore) |
| | **5** | Positive Parenting *Frequency* |
| | **6** | Supporting Positive Behaviour *Problem* (subscore) |
| | **7** | Setting Limits *Problem* (subscore) |
| | **8** | Positive Parenting *Problem* |
| | | **Harsh Parenting** |
| | **9** | Non-Violent Discipline |
| | **10** | Harsh Discipline |
| | **11** | Neglect |
| | | **Observed Parenting and Child Behaviour** |
| | **12** | Child Positive Behaviour |
| | **13** | Child Negative Behaviour |
| | **14** | Parent Positive Behaviour |
| | **15** | Parent Negative Behaviour |
| | **16** | Child Positive Impression |
| | **17** | Parent Positive Impression |

| | | |
|---|---|---|
| **Secondary Outcomes** | | **Monitoring and Supervision** |
| | **18** | Poor Monitoring and Supervision |
| | | **Depression** |
| | **19** | Beck Depression Inventory |
| | | **Parenting Stress** |
| | **20** | Parental Distress (subscore) |
| | **21** | Parent-Child Dysfunctional Interaction (subscore) |
| | **22** | Difficult Child (subscore) |
| | **23** | Parenting Stress |
| | | **Social Support** |
| | **24** | Social Support |

As a confirmatory item reliability analysis, Cronbach's alpha statistics will be calculated for each score. Investigations into the construction of scores will be limited to the following aspects:

- For the individual items used to create the Harsh Parenting scores, a number of behaviours (items) may have very little endorsement (due to a lack of applicability in the rural South African context, caregivers' unwillingness to admit severely abusive practices, or simply because they do not occur). Based on an initial exploration of individual items (frequency tables), items/scores deemed to provide little insight (that is, endorsed at very low frequencies) will be excluded.
- For the Supporting Positive Behaviour Frequency score, a new item has been introduced into the questionnaire for this study (assessing the frequency of having a family meal together). The internal consistency of this item, with existing items, will be assessed by measuring Cronbach's alpha (with and without the item). If there is consistency, the item will be included in the construction of the score. Similarly, a new item for the Supporting Positive Behaviour Problem score (assessing the difficulty of having a family meal together) will be evaluated and possibly included.
- To better understand the interpretation of the Parenting Stress score (Score 23), which is the sum of the Scores 20, 21 and 22, the correlations amongst the three individual scores will be investigated.

The scores will be constructed from individual items within the data imputation process (discussed below).

## Preliminary Data Exploration

Baseline characteristics about the child/caregiver/household (such as HIV status of child and caregiver, age of caregiver, measures of intimate partner violence in household) and baseline outcome variables will be described.

Continuous (or pseudo continuous) variables will be summarised by the observed mean and standard deviation of measurements, and categorical variables by the proportion of observations in each category (together with the number of observations available – that is, with non-missing values). For continuous measurements with underlying skew distributions, the median and inter-quartiles ranges will instead be provided.

The characteristics will be summarised, stratifying by the following features (in turn):

- Study Arm
- Whether subjects were lost to follow-up (immediately after baseline assessment)

When stratifying by Arm, the relative frequency of categories of child sex and child age category (combinations of male/female and 2-5/6-9 years old) will be the same by design.

Templates of the envisioned tables are provided in Appendix B (see Tables B1 to B4).

## Imputation of Missing Data

Although the data analysis utilises a likelihood-based model (discussed below), which therefore implicitly provides a means of handling missing data, not all variables that could be conceived as being important in data imputation are included in the data analysis, termed the substantive, model. Therefore, multiple imputation will first be used to create multiple plausible 'complete' datasets, which will then be analysed using the substantive model. The pooling of final results (estimated effect sizes) obtained from the multiple 'complete' datasets will account for the imputation process (discussed in *Data Analysis – Estimation of Effect Sizes*).

The imputation of missing data will proceed in three steps:

- **Inspecting the missing data**
- **Applying the imputation model** to create multiple 'complete' datasets
- **Assessing convergence** of the imputation model and **checking imputations**

Two **sensitivity analyses** will also be performed.

### *Inspection of data*

The patterns of missing data will be explored and reported: The frequencies of different patterns of which data is missing will be calculated; the empirical distributions of key variables will be plotted, stratified by whether other variables are missing.

### *Imputation model*

The covariates included in the analysis model are Arm, Wave, Sex and Age Group, which will always be known.

However, some outcome data may be missing, and such missing data will be imputed using Multivariate Imputation by Chained Equations (MICE), also known as Fully Conditional Specification (FCS). The imputation model is defined by specifying a set of conditional densities, from which to draw measurements (for observations with missing data) for each variable in turn, using other variables as predictors (which may themselves contain imputed values). The process is iterative, and utilises Markov Chain Monte Carlo (MCMC) techniques.

There are three ways in which a subject may have missing outcome data (not mutually exclusive):

- The subject was lost to follow-up and there are no post-test measurements

- The subject (not necessarily lost to follow-up) did not provide any data on all items used to construct a specific score (at baseline or post-test).
- The subject (not necessarily lost to follow-up) did not provide data for only a subset of items that contribute towards a specific score (at baseline or post-test).

All three cases are handled by the imputation model described below.

To ensure that the underlying imputation model does not change when 12-month follow-up data is also analysed, in the imputation model, only data collected at the same time point or earlier time points is ever used to impute a missing value.

Consider a **particular variable** (or item), such as how often the caregiver plays with the child (the response is captured by a 7-point Likert scale), as measured **at a given time point** (either baseline or post-test). In general, the item (measured at either time point) would contribute to a **specific score** (for that time point), in this case the Supporting Positive Behaviour Frequency score. The specific score is part of a **particular group of scores**, in this case Positive Parenting. When imputing values for the particular variable, the following predictors will be used:

- Child sex (male or female)
- Child age group (2-5 years old or 6-9 years old)
- Arm (Control Arm or Intervention Arm)
- Wave (Wave 1 or Wave 2)
- All other items used to construct the specific score at the same time point (baseline or post-test)
- All items used to construct the same specific score at any earlier time points, if there are any earlier time points (baseline)
- The other scores falling into the corresponding group of scores – at that time point and any earlier time points
- All remaining primary and secondary scores (that is, the scores for all groups of scores) – at that time point and any earlier time points
- Group (for the Intervention Arm)

Reported and observed outcomes will be imputed separately (that is, reported and observed outcomes will not contribute to each other's imputation models).

The large dataset provides motivation for including only those individual items that relate to the same specific score to which that item contributes. All other outcomes enter the (conditional density) only in the form of composite scores.

The functional form of the model used to generate (otherwise missing) measurements for a given variable will depend on the type of data the variable captures:

- For dichotomous variables, logistic regression will be used.
- For the Likert-scales, predictive mean matching will be utilised.

In both cases, a dichotomous predictor will enter into the model as a dichotomous categorical variable, while Likert-scale predictors will enter as numeric (continuous) variables. Since there are only two time points, flat-file imputation will be applied (a particular item/score which is measured at each of the two time points will be included as two separate variables, rather than as two measurements of the same variable within some hierarchical model). The decisions to treat Likert-scales as numeric rather than categorical and use a flat-file approach again aim to restrict the complexity and size of the imputation models. As an exception, the approach for including Group will be explored during the implementation of the imputation model (as a fixed effect categorical variable, or rather as a random effect in a mixed effects regression model, or through data stratification, or neglected if there is little evidence of a Group effect).

If the size of the imputation problem gets too large (for example, too computationally expensive or suffering from instabilities due to multicollinearity), the process will be simplified and number of predictors reduced. Firstly, the imputation will be performed in parts, so as to only work with the subset of data relevant for exploring a particular group of outcomes at a time (see the 8 groups listed in *Construction of Outcome Variables*). This implies that the very last set of predictors in the list above (scores for all other groups of scores) would no longer be included in the imputation model for any given variable. Secondly, the items measured at an earlier time point for the specific score could be replaced by just the overall (measured or imputed) score. Thirdly, the inclusion of only scores as predictors (and no individual items as predictors) could be explored.

A random number generator seed of 30709 will be used to initiate the imputation model.

***Assessment of convergence and checking imputations***

Plots of summary statistics (mean and standard deviation) of the imputed measurements by iteration number, per variable, will be used to identify model misspecifications and assess whether the number of iterations is sufficiently high (sequences for the different imputed datasets should be freely intermingled and variability between sequences should not be larger than within sequences, upon convergence). Plots of the distributions of the variables – observed and imputed – will also be created.

***Sensitivity analyses***

To investigate the sensitivity of results to the assumptions of the imputation model, two alternative (secondary) imputation approaches will be implemented.

The imputation model described above assumes that data is Missing at Random (MAR) – that is, the probability of observations being missing depends only on the observed data (and we can correctly predict values for the missing observations using the patterns seen in the observed observations). The first sensitivity analysis will again utilise a MAR assumption, but will effectively change the underlying imputation model form (including the predictors included). The second sensitivity analysis will extend the multiple imputation model described above to capture a hypothetical mechanism in which data is Missing Not at Random (MNAR) – meaning that the probability of observations being missing depend on the unobserved data too.

More specifically:

- Firstly, the analysis described below (see *Data Analysis – Estimation of Effect Sizes*) will be applied without imputing missing values. Since the substantive model is likelihood-based, it implicitly provides a way of handling missing data – again assuming a MAR process and also that only variables included in the substantive model are important for ‘predicting’ what the missing data would be.
- Secondly, the model described above will be utilised, but with the extension that any imputed value for an item needs be given a marginally more desirable or less desirable value than it would be by the model above. That is, we assume a MNAR process in which subjects who have

relatively more or less desirable responses are the ones for whom data is generally missing (hypothetical scenarios will be selected).

For the main imputation model, and under each of the secondary approaches for handling the missing data, the effect size of interest (with its uncertainty) will be estimated.

## Data Analysis – Estimation of Effect Sizes

The main objective of the analysis is to estimate the **difference between the two arms** in:

**the change in scores from baseline to post-test**.

### *Sample Sizes*

148 dyads were assigned to each study arm. In the study protocol, the sample size calculation suggested that 120 dyads would be required per arm, based on a two-tailed two-sample t-test, and then increasing the sample size by 30% to account for the inclusion of other variables in the analysis. A regression model is presented below, so that the clustering of data (by Dyad and Group) and the inclusion of covariates can be accommodated.

### *Summary statistics*

Before fitting models to the data, the observed changes in measurements from baseline to post-test will be described. For each outcome and each Arm, the average (across subjects) difference between the post-test score and the baseline score will be calculated (with a 95% confidence interval). The template for the table is provided in Appendix B (Table B5).

### *Primary analysis of a given 'complete' dataset and for a chosen outcome*

Each of the 28 outcomes will be analysed in turn. Also, the analysis (the substantive model) will be applied to each of the imputed or 'complete' datasets in turn. Therefore, the analysis is described below in terms of a single outcome and single dataset.

Although each of the items in the dataset is categorical in nature, the composite score (that is, the outcome to be analysed) is constructed as a sum of several individual items. Therefore, it is reasonable to expect that the pseudo-continuous outcome will (approximately) follow a conventional parametric distribution, or more specifically, be approximately normally distributed, and thus parametric regression is likely to be used to analyse the data. The distribution of the outcome will be assessed visually.

Multi-level modelling will be used to capture both that the same child-caregiver dyads are assessed at baseline and post-test (that is, there are repeated measures) and the group-based nature of the intervention (participants in the same Group may respond similarly). More specifically, a linear 'mixed effects' model will be applied. This method of analysis is also appealing as the models can be relatively easily extended to include the additional data to be collected at 12 months post-intervention.

**Model Specification**

For Group $i$, and child-caregiver Dyad $j$ in the group, assessed at Time $k$, the model specifies

$$y_{i,j,k} = \alpha_0 + \alpha_t \cdot \delta_{time,i,j,k} + \alpha_{at} \cdot \delta_{arm,i,j,k} \cdot \delta_{time,i,j,k} + \alpha_w \cdot \delta_{wave,i,j,k} + \alpha_s \cdot \delta_{sex,i,j,k} + \alpha_g \cdot \delta_{age,i,j,k} + \varepsilon_{i,j,k}$$

where

- $i$ = 0,1,2,…,$n_g$ and all participants in the Control Arm are in Group 0 and participants in the Intervention Arm are in one of Group 1 to Group $n_g$;
- $j$ = 1,2,…$n_i$; and
- $k$ equa1s 1 or 2, corresponding to baseline and post-test time points respectively;

and

- $y_{i,j,k}$ is the outcome (score);
- $\delta_{time,i,j,k}$ equals 0 for Time 1 (baseline) and equals 1 for Time 2 (post-test);
- $\delta_{arm,i,j,k}$ equals 0 for the Control Arm and equals 1 for the Intervention Arm2;
- $\delta_{wave,i,j,k}$ equals 0 for Wave 1 and equals 1 for Wave 2;
- $\delta_{sex,i,j,k}$ equals 0 if the Child is Male and equals 1 if the Child is Female;
- $\delta_{age,i,j,k}$ equals 0 if the Child is 2-5 Years Old and equals 1 if the Child is 6-9 Years Old; and
- $\varepsilon_{i,j,k}$ is a draw from a normal distribution with mean 0 and variance $\sigma_e^2$ (identically and independently drawn every time an outcome is measured).

The terms $\alpha_0$, $\alpha_t$, $\alpha_{at}$, $\alpha_w$, $\alpha_s$ and $\alpha_g$ consist of both 'fixed effects' (representing population-level averages) and 'random effects' (representing Group-specific and Dyad-specific deviations from these averages). In particular,

- $\alpha_0 = \beta_0 + b_{0,i,j}$,
- $\alpha_t = \beta_t + b_{t,i,j}$, although the simpler form $\alpha_t = \beta_t$ will first be considered,
- $\alpha_{at} = \beta_{at} + b_{at,i}$,
- $\alpha_w = \beta_w$,
- $\alpha_s = \beta_s$, and
- $\alpha_g = \beta_g$.

The random effects $b_{0,i,j}$, $b_{t,i,j}$, and $b_{at,i}$ follow normal distributions,

- $\begin{bmatrix} b_{0,i,j} \\ b_{t,i,j} \end{bmatrix} \sim N\left(\begin{bmatrix} 0 \\ 0 \end{bmatrix}, \begin{bmatrix} \sigma_1^2 & \sigma_{12} \\ \sigma_{12} & \sigma_2^2 \end{bmatrix}\right)$; and
- $b_{at,i} \sim N(0, \sigma_3^2)$,

while the fixed effects, $\beta_0$, $\beta_t$, $\beta_{at}$, $\beta_w$, $\beta_s$ and $\beta_g$, are model parameters.

The fixed effects ($\beta_0$, $\beta_t$, $\beta_{at}$, $\beta_w$, $\beta_s$, and $\beta_g$) and the remaining model parameters ($\sigma_1^2$, $\sigma_2^2$, $\sigma_{12}$, $\sigma_3^2$ and $\sigma_e^2$) will be estimated using a maximum likelihood approach.

The fixed effects that capture changes over time are

- $\beta_t$, which is the average (additive) change in the outcome from baseline to post-test for the Control Arm; and
- $\beta_t + \beta_{at}$, which is the average (additive) change in the outcome from baseline to post-test for the Intervention Arm.

The **effect size** is therefore represented by $\beta_{at}$, which measures the difference between the two arms.

Pooled parameter estimates will be obtained by averaging over the multiple imputed datasets (discussed below). A template for reporting key effect size estimates is provided in Appendix B (Table B6). In secondary analyses (discussed below), effect sizes by Wave, Sex or Age Group could be obtained, and reported in further tables.

The appropriateness of the model assumptions (and fit of the model) will be assessed by plots of standardised residuals against fitted values (and covariates) and histograms of the residuals. Model fit will be investigated using one of the multiple imputed datasets per outcome.

***Secondary analysis***

The primary analysis above will adopt an **Intention-to-Treat** approach (and therefore use all data).

The analysis will be reproduced under a **Per-Protocol approach**, in which participants must have attended a specified number of sessions.

To investigate subgroup differences, the inclusion of interaction terms could be explored.[1] More specifically, for each of Sex, Age Group and Wave, the variable will be entered into the model, interacting with (1) Time, and (2) the Arm-Time Interaction.

***The pooled effect size***

For each outcome, a given analysis will be performed on multiple imputed or 'complete' datasets, and therefore multiple effect size estimates obtained. Pooled effect size estimates, and confidence intervals (that correctly account for variability arising from imputing the missing data), will be calculated using Rubin's rules.

***Alternative analysis approaches***

If the distributions of the scores are non-normal, simple transformations of the data (to normalise the distributions) will be considered, such as log transformations (complex transformations will not be explored as results will become difficult to interpret). Alternatively, generalised linear mixed models (rather than linear mixed models) may be applied to the data – allowing for the outcome to follow a non-normal, but parametric, distribution (from the exponential family of distributions). For example, meaningful thresholds on the scores may be chosen to create categorical responses, which can then be modelled using a multinomial mixed model. Finally, non-parametric approaches for testing for differences between arms could be applied. However, random effects cannot be included, and so changes in scores (from baseline to post-test) will first be calculated, and these differences will become the new responses of interest.

[1] This will be explored as part of the primary analysis should the model fit prove inadequate.

### *Multiple Testing*

A large number of analyses are to be performed. This leads to multiple testing problems, where the probability of falsely identifying some significant differences is going to be much higher than intended. This limitation will be discussed in the presentation of the analysis results.

## Software

The data will be explored and analysed using both SPSS (IBM SPSS Statistics, Version 22) and R (The R Foundation for Statistical Computing, Version 3.1.3, 64-bit). For multiple imputation (and pooling of estimates), the R package 'MICE' will be used (possibly together with Microsoft Excel to define inputs to the R imputation functions); the mixed models will be fitted using the R package 'nlme'.

# Appendix A: Construction of Outcomes

**Child Behaviour Problems (Primary Outcomes)**

As reported by primary caregivers, using the Eyeberg Child Behaviour Inventory (ECBI) questionnaire. Higher scores indicate less desirable behaviours.

| | **Score** | **Description of score** | **Description of items** |
|---|---|---|---|
| **1** | ECBI Frequency | Sum of 36 items, each one a 7-point Likert scale | 1,2,…,7 for ‘Never’, “Very rarely’, ‘Rarely’, ‘Sometimes’, ‘Often’, ‘Very often’ and ‘Always’ respectively |
| **2** | ECBI Intensity | Sum of 36 items, each one dichotomous | 0 and 1 for ‘no’ and ‘yes’ respectively |

**Positive Parenting (Primary Outcomes)**

As reported by primary caregivers, using the Positive Parenting and Setting Limits Subscales of the Parenting Young Children Scale. Higher scores indicate more desirable behaviours.

| | **Score** | **Description of score** | **Description of items** |
|---|---|---|---|
| **3** | Supporting Positive Behaviour *Frequency* | Sum of 7 (or 8) items, each one a 7-point Likert scale | 0,1,…,6 for ‘Never’, ‘Very rarely’, ‘Rarely’, ‘Sometimes’, ‘Often’, ‘Very often’ and ‘Always’ respectively |
| **4** | Setting Limits *Frequency* | Sum of 7 items, each one a 7-point Likert scale | |
| **5** | Positive Parenting *Frequency* | Sum of Scores 3 and 4 | |
| **6** | Supporting Positive Behaviour *Problem* | Sum of 7 (or 8) items, each one dichotomous | 0 and 1 for ‘no’ and ‘yes’ respectively |
| **7** | Setting Limits *Problem* | Sum of 7 items, each one dichotomous | |
| **8** | Positive Parenting *Problem* | Sum of Scores 6 and 7 | |

**Harsh Parenting (Primary Outcomes)**

As reported by primary caregivers, using the ICAST (IPSCAN Child Abuse Screening Test). Higher scores relate to increasing levels of the behaviour.

| | **Score** | **Description of score** | **Description of items** |
|---|---|---|---|
| **9** | Non-Violent Discipline | Sum of 4 items, each one a 6-point Likert scale | 0,1,…5 for ‘Never’, ‘Has happened but not in past month’, ‘Once or twice’, ‘3-5 times’, ‘6-10 times’, ‘More than 10 times’ respectively |
| **10** | Harsh Discipline | Sum of 13 items, each one a 6-point Likert scale | |
| **11** | Neglect | Sum of 3 items, each one dichotomous | Original data to be recoded such that 0 corresponds to ‘no’ and 1 to ‘yes’, and response relates to behaviour over the last month |

**Observed Parenting and Child Behaviour (Primary Outcomes)**

This scoring system is currently under development. The frequency score assessment has been loosely based on the Dyadic Parent-Child Interaction Scale.

| | **Score** | **Description of score** |
|---|---|---|
| **12** | Child Positive Behaviour | Sum of a number of items, each of which provides a frequency (count) |
| **13** | Child Negative Behaviour | Sum of a number of items, each of which provides a frequency (count) |
| **14** | Parent Positive Behaviour | Sum of a number of items, each of which provides a frequency (count) |
| **15** | Parent Negative Behaviour | Sum of a number of items, each of which provides a frequency (count) |
| **16** | Child Positive Impression | Sum of a number of items, each one a Likert-scale |
| **17** | Parent Positive Impression | Sum of a number of items, each one a Likert-scale |

**Monitoring and Supervision (Secondary Outcomes)**

As reported by parents, using the Poor Monitoring and Supervision Subscale of the Alabama Parenting Questionnaire. A high score corresponds to poor parental monitoring.

| | **Score** | **Description of score** | **Description of items** |
|---|---|---|---|
| **18** | Poor Monitoring and Supervision | Sum of 9 items, each one a 5-point Likert scale | 1,2,…,5 for ‘Never’, ‘Almost never’, ‘Sometimes’, ‘Often’, and ‘Always’ respectively<br>(3 specified items are to be reverse coded before summation) |

**Depression (Secondary Outcomes)**

As reported by the parents, using the Beck Depression Inventory. A higher score corresponds to more severe depression.

| | Score | Description of score | Description of items |
|---|---|---|---|
| **19** | Beck Depression Inventory | Sum of 21 items, each one a 4-point Likert scale | 0,1,2,3 corresponding to increasing levels of experiencing the negative emotion (item-specific categories) |

**Parenting Stress (Secondary Outcomes)**

As reported by parents, using the Parenting Stress Index. A higher score corresponds to less stress.

<table>
<tr><th></th><th>Score</th><th>Description of score</th><th>Description of items</th></tr>
<tr><td>20</td><td>Parental Distress</td><td>Sum of 12 items, each one a 5-point Likert scale</td><td>1,2,…,5 for ‘Strongly agree, ‘Agree’, ‘Not sure’, ‘Disagree’ and ‘Strongly disagree’ respectively</td></tr>
<tr><td>21</td><td>Parent-Child Dysfunctional Interaction</td><td>Sum of 12 items, each one a 5-point Likert scale</td><td rowspan="2">1,2,…,5, typically for ‘Strongly agree, ‘Agree’, ‘Not sure’, ‘Disagree’ and ‘Strongly disagree’ respectively (some items utilise different categories)</td></tr>
<tr><td>22</td><td>Difficult Child</td><td>Sum of 12 items, each one a 5-point Likert scale</td></tr>
<tr><td>23</td><td>Parenting Stress</td><td>Sum of Scores 20, 21 and 22</td><td></td></tr>
</table>

**Social Support (Secondary Outcomes)**

As reported by parents, using the Medical Outcomes Study Social Support Scale. A higher score corresponds to greater social support.

| | Score | Description of score | Description of items |
|---|---|---|---|
| **24** | Social Support | Sum of 8 items, each one a 5-point Likert scale | 1,2,…,5 for ‘None of the time’, ‘A little of the time’, ‘Some of the time’, ‘Most of the time’ and ‘All of the time’ |

# Appendix B: Templates for Tables

**Table B1. Baseline characteristics of child/caregiver/household by Arm**

| | **Control Arm** | **Intervention Arm** |
|---|---|---|
| **Age of caregiver** | | |
| **Age of child** | | |
| **HIV status of caregiver** | | |
| Positive | | |
| Negative | | |
| Unknown | | |
| **HIV status of child** | | |
| Positive | | |
| Negative | | |
| Unknown | | |
| **Family structure** | | |
| Biological mother present in household | | |
| Biological father present in household | | |
| **Child's HIV orphanhood status** | | |
| Single orphan | | |
| Double orphan | | |
| **Gender of caregiver** | | |
| Female | | |
| Male | | |
| **Gender of child** | | |
| Female | | |
| Male | | |
| **Caregiver employment status** | | |
| Unemployed | | |
| Employed | | |
| **Caregiver substance use** | | |
| Alcohol | | |
| Drugs | | |
| **Caregiver experience of intimate partner violence** | | |
| **Caregiver personal experience of maltreatment** | | |
| Physical abuse | | |
| Verbal abuse | | |
| Sexual abuse | | |

**Table B2. Baseline scores by Arm**

| | **Control Arm** | **Intervention Arm** |
|---|---|---|
| **Child Behaviour Problems** | | |
| ECBI Frequency | | |
| ECBI Intensity | | |
| **Positive Parenting** | | |
| Supporting Positive Behaviour Frequency (subscore) | | |
| Setting Limits Frequency (subscore) | | |
| Positive Parenting Frequency | | |
| Supporting Positive Behaviour Problem (subscore) | | |
| Setting Limits Problem (subscore) | | |
| Positive Parenting Problem | | |
| **Harsh Parenting** | | |
| Non-Violent Discipline | | |
| Harsh Discipline | | |
| Neglect | | |
| **Observed Parenting and Child Behaviour** | | |
| Child Positive Behaviour | | |
| Child Negative Behaviour | | |
| Parent Positive Behaviour | | |
| Parent Negative Behaviour | | |
| Child Positive Impression | | |
| Parent Positive Impression | | |
| **Monitoring and Supervision** | | |
| Poor Monitoring and Supervision | | |
| **Depression** | | |
| Beck Depression Inventory | | |
| **Parenting Stress** | | |
| Parental Distress (subscore) | | |
| Parent-Child Dysfunctional Interaction (subscore) | | |
| Difficult Child (subscore) | | |
| Parenting Stress | | |
| **Social Support** | | |
| Social Support | | |

**Table B3. Baseline characteristics of child/caregiver/household by loss to follow-up (after baseline assessment)**

| | Lost to follow-up | Not lost to follow-up |
|---|---|---|
| **Age of caregiver** | | |
| **Age of child** | | |
| **HIV status of caregiver** | | |
| Positive | | |
| Negative | | |
| Unknown | | |
| **HIV status of child** | | |
| Positive | | |
| Negative | | |
| Unknown | | |
| **Family structure** | | |
| Biological mother present in household | | |
| Biological father present in household | | |
| **Child's HIV orphanhood status** | | |
| Single orphan | | |
| Double orphan | | |
| **Gender of caregiver** | | |
| Female | | |
| Male | | |
| **Gender of child** | | |
| Female | | |
| Male | | |
| **Caregiver employment status** | | |
| Unemployed | | |
| Employed | | |
| **Caregiver substance use** | | |
| Alcohol | | |
| Drugs | | |
| **Caregiver experience of intimate partner violence** | | |
| **Caregiver personal experience of maltreatment** | | |
| Physical abuse | | |
| Verbal abuse | | |
| Sexual abuse | | |

**Table B4. Baseline scores by loss to follow-up (after baseline assessment)**

| | Lost to follow-up | Not lost to follow-up |
|---|---|---|
| **Child Behaviour Problems** | | |
| ECBI Frequency | | |
| ECBI Intensity | | |
| **Positive Parenting** | | |
| Supporting Positive Behaviour Frequency (subscore) | | |
| Setting Limits Frequency (subscore) | | |
| Positive Parenting Frequency | | |
| Supporting Positive Behaviour Problem (subscore) | | |
| Setting Limits Problem (subscore) | | |
| Positive Parenting Problem | | |
| **Harsh Parenting** | | |
| Non-Violent Discipline | | |
| Harsh Discipline | | |
| Neglect | | |
| **Observed Parenting and Child Behaviour** | | |
| Child Positive Behaviour | | |
| Child Negative Behaviour | | |
| Parent Positive Behaviour | | |
| Parent Negative Behaviour | | |
| Child Positive Impression | | |
| Parent Positive Impression | | |
| **Monitoring and Supervision** | | |
| Poor Monitoring and Supervision | | |
| **Depression** | | |
| Beck Depression Inventory | | |
| **Parenting Stress** | | |
| Parental Distress (subscore) | | |
| Parent-Child Dysfunctional Interaction (subscore) | | |
| Difficult Child (subscore) | | |
| Parenting Stress | | |
| **Social Support** | | |
| Social Support | | |

**Table B5. Observed changes in scores from baseline to post-test by Arm (mean and 95% confidence interval)**

| | Control Arm | Intervention Arm | Total |
|---|---|---|---|
| **Child Behaviour Problems** | | | |
| ECBI Frequency | | | |
| ECBI Intensity | | | |
| **Positive Parenting** | | | |
| Supporting Positive Behaviour Frequency (subscore) | | | |
| Setting Limits Frequency (subscore) | | | |
| Positive Parenting Frequency | | | |
| Supporting Positive Behaviour Problem (subscore) | | | |
| Setting Limits Problem (subscore) | | | |
| Positive Parenting Problem | | | |
| **Harsh Parenting** | | | |
| Non-Violent Discipline | | | |
| Harsh Discipline | | | |
| Neglect | | | |
| **Observed Parenting and Child Behaviour** | | | |
| Child Positive Behaviour | | | |
| Child Negative Behaviour | | | |
| Parent Positive Behaviour | | | |
| Parent Negative Behaviour | | | |
| Child Positive Impression | | | |
| Parent Positive Impression | | | |
| **Monitoring and Supervision** | | | |
| Poor Monitoring and Supervision | | | |
| **Depression** | | | |
| Beck Depression Inventory | | | |
| **Parenting Stress** | | | |
| Parental Distress (subscore) | | | |
| Parent-Child Dysfunctional Interaction (subscore) | | | |
| Difficult Child (subscore) | | | |
| Parenting Stress | | | |
| **Social Support** | | | |
| Social Support | | | |

**Table B6. Effect size estimates and confidence intervals – differences in changes from baseline to post-test scores for Intervention Arm compared to Control Arm**

| | **Effect size estimate** | **95% confidence interval** |
|---|---|---|
| **Child Behaviour Problems** | | |
| ECBI Frequency | | |
| ECBI Intensity | | |
| **Positive Parenting** | | |
| Supporting Positive Behaviour Frequency (subscore) | | |
| Setting Limits Frequency (subscore) | | |
| Positive Parenting Frequency | | |
| Supporting Positive Behaviour Problem (subscore) | | |
| Setting Limits Problem (subscore) | | |
| Positive Parenting Problem | | |
| **Harsh Parenting** | | |
| Non-Violent Discipline | | |
| Harsh Discipline | | |
| Neglect | | |
| **Observed Parenting and Child Behaviour** | | |
| Child Positive Behaviour | | |
| Child Negative Behaviour | | |
| Parent Positive Behaviour | | |
| Parent Negative Behaviour | | |
| Child Positive Impression | | |
| Parent Positive Impression | | |
| **Monitoring and Supervision** | | |
| Poor Monitoring and Supervision | | |
| **Depression** | | |
| Beck Depression Inventory | | |
| **Parenting Stress** | | |
| Parental Distress (subscore) | | |
| Parent-Child Dysfunctional Interaction (subscore) | | |
| Difficult Child (subscore) | | |
| Parenting Stress | | |
| **Social Support** | | |
| Social Support | | |

# Appendix C

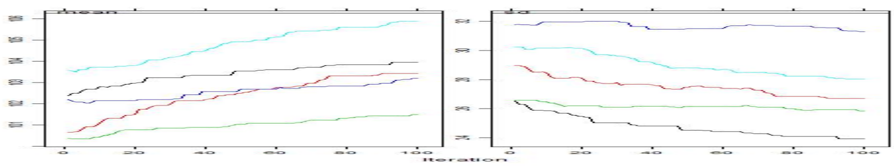
# Checking for convergence

## Imputation using predictive mean matching and logistic regression: convergence checks

Note: All **problem** scores are made up of binary items and therefore logistic regression was used for them, all other scores were imputed using predictive mean matching.

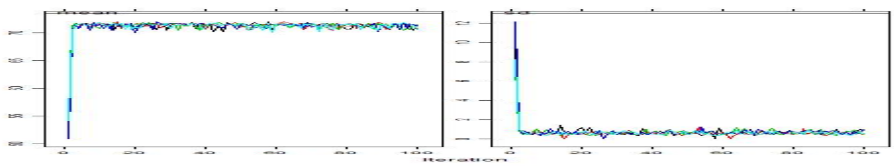
**ECBI Intensity**

(a) **One year follow-up visit**

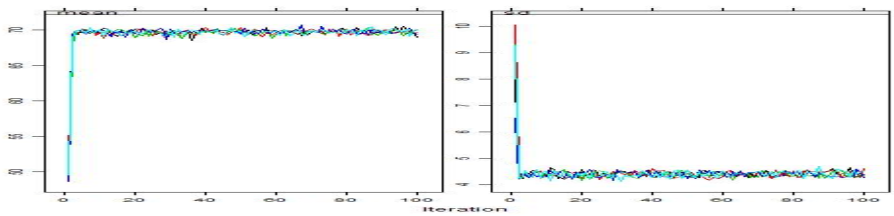


**ECBI Problem**

(a) **Post-test visit**

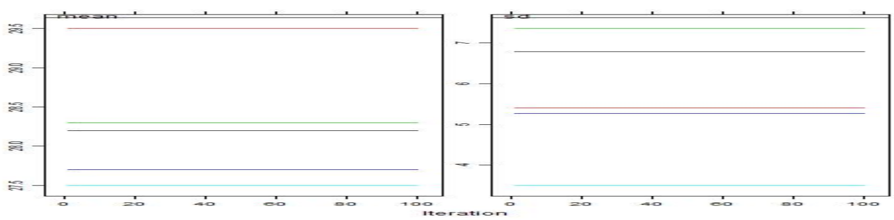


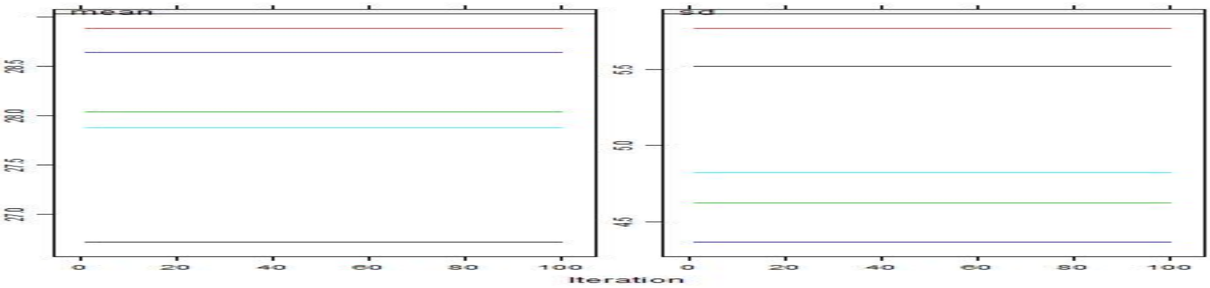
(b) **One year follow-up visit**



**Supporting Positive Behaviour Frequency**

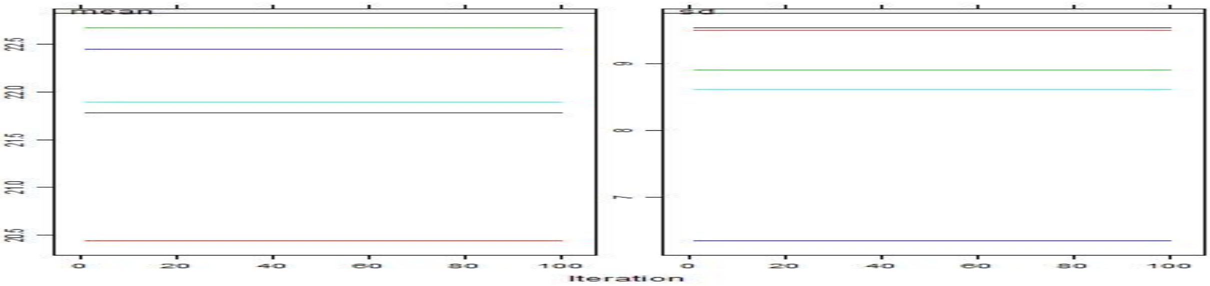
(a) **Post-test visit**



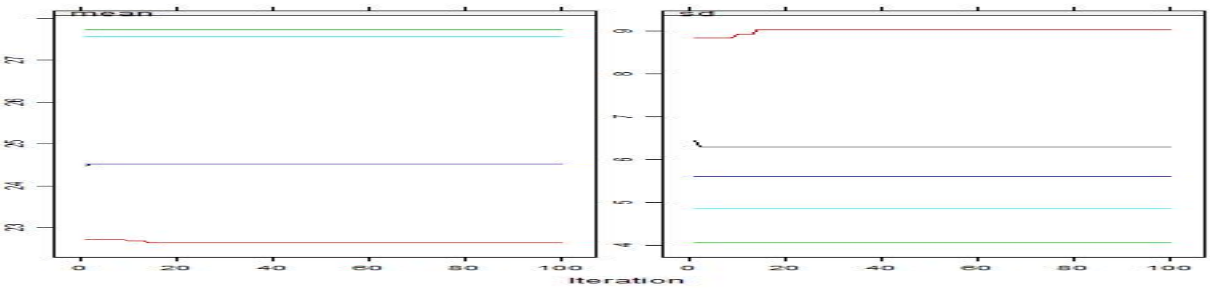



(b) One year follow-up visit

**Setting limits Frequency**

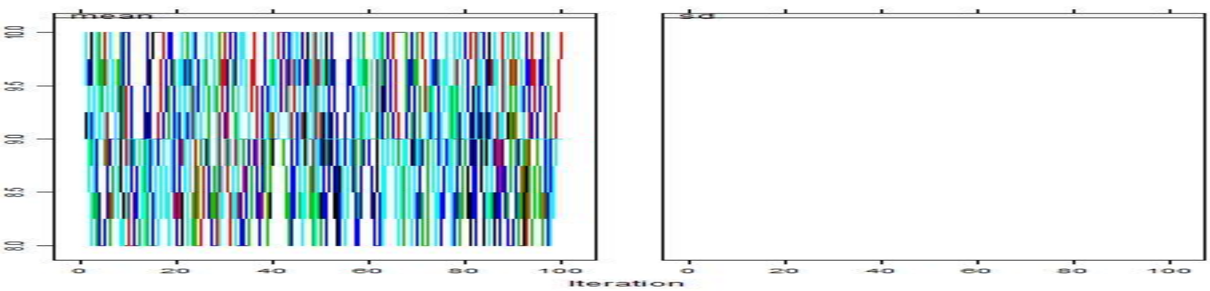



(a) **Post-test visit**




(b) **One year follow-up visit**

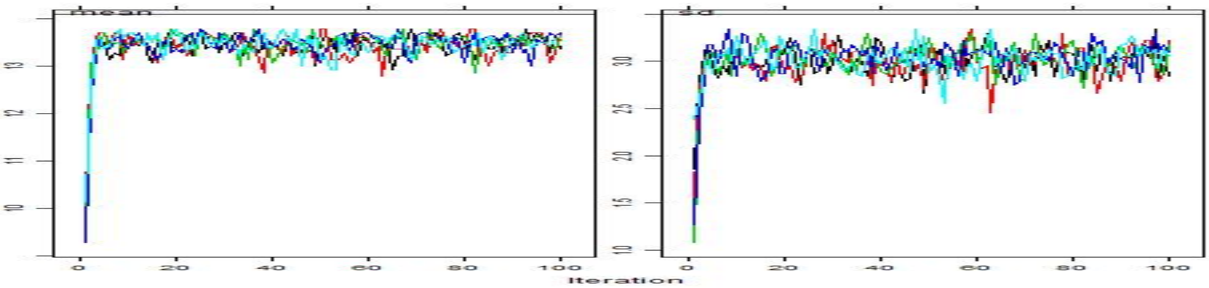
**Supporting Positive Behaviour Problem**



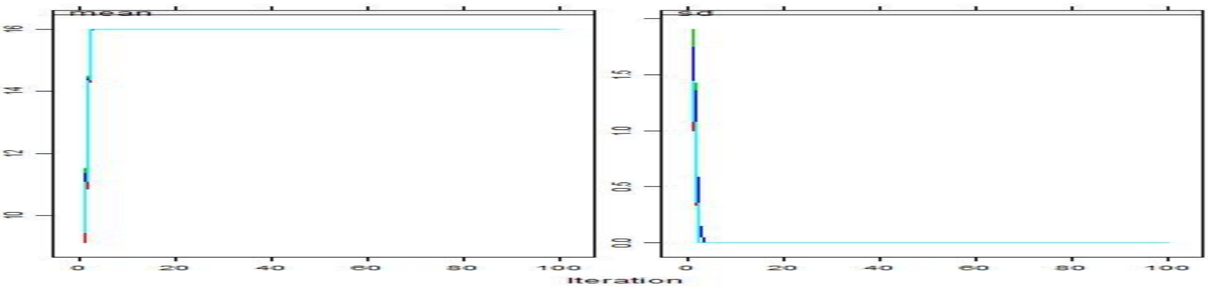

(a) **Baseline visit**

Figure 6: Supporting Positive Behaviour Problem

(a) Post-test visit

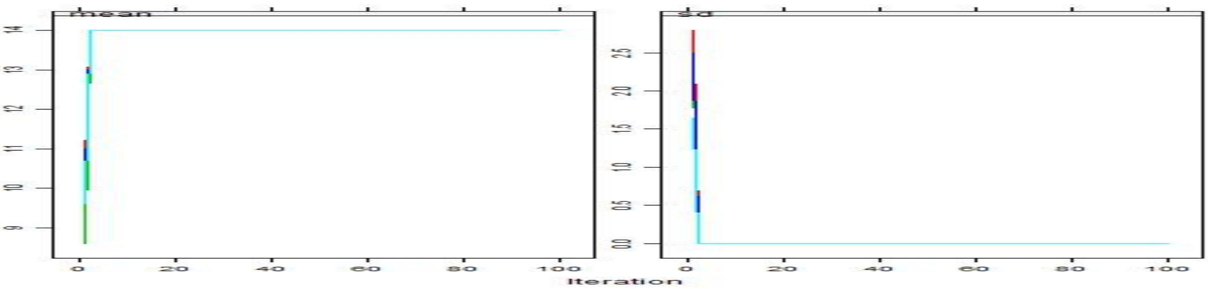



(b) One year follow-up visit

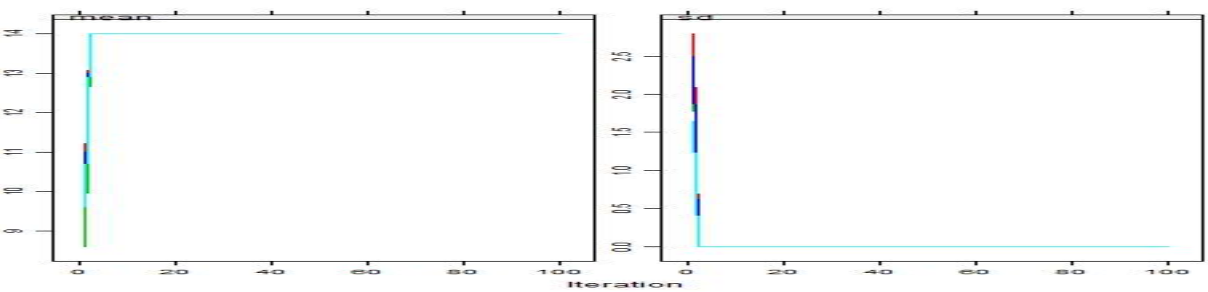


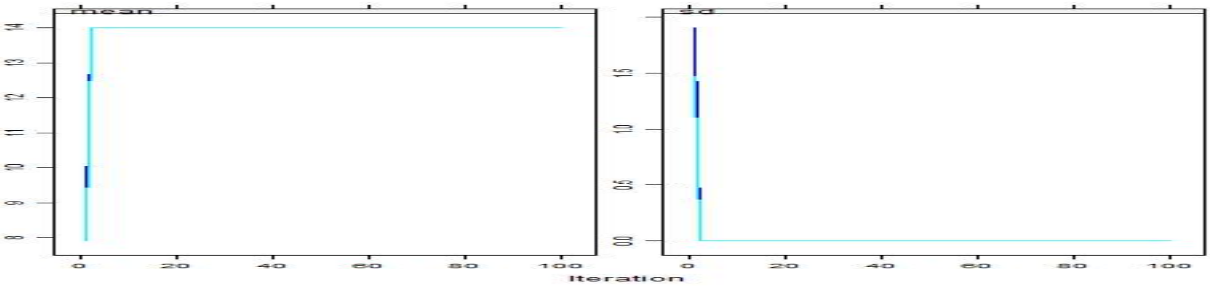

## Setting limits Problem

(a) **Baseline visit**



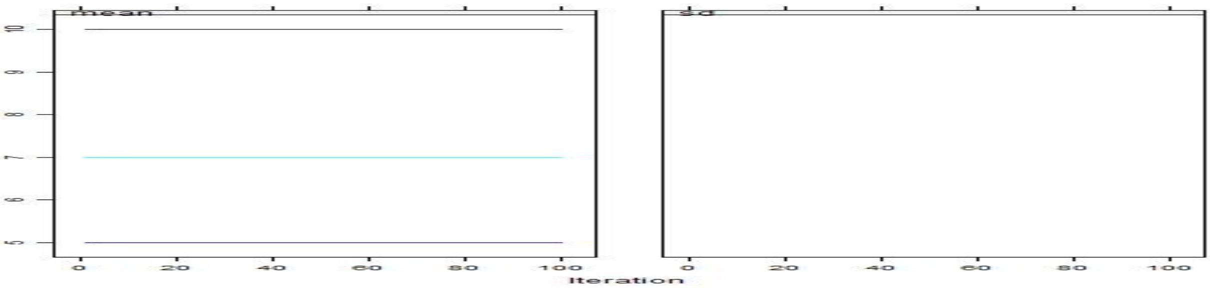

(b) **Post-test visit**



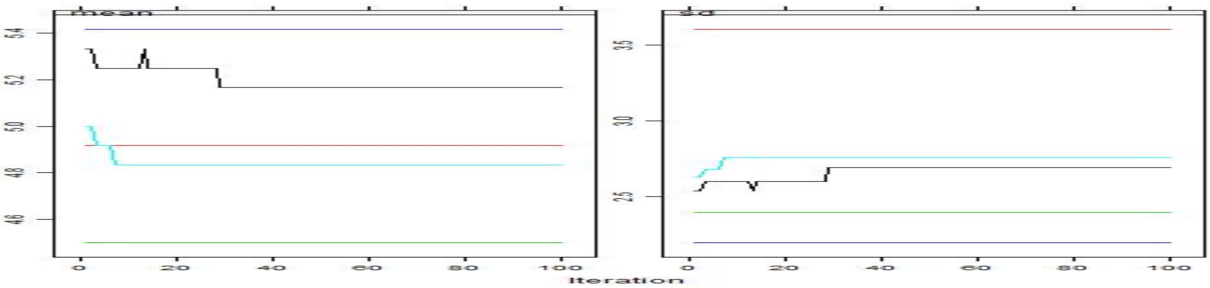



(a) **One year follow-up visit**

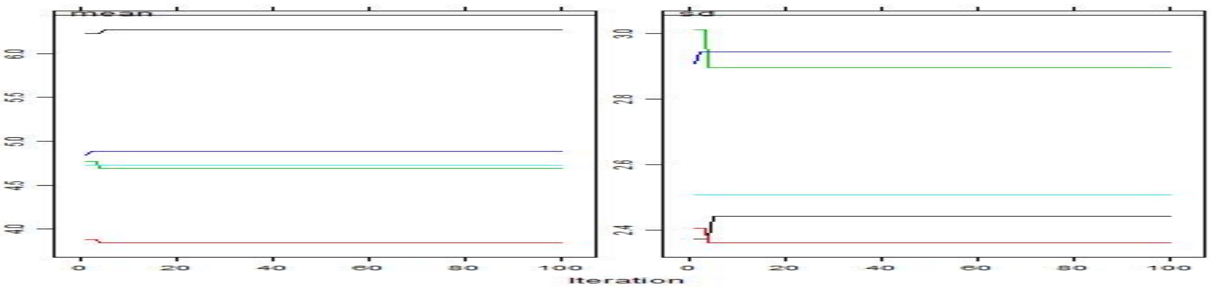
## Non-Violent Discipline




(a) **Baseline visit**



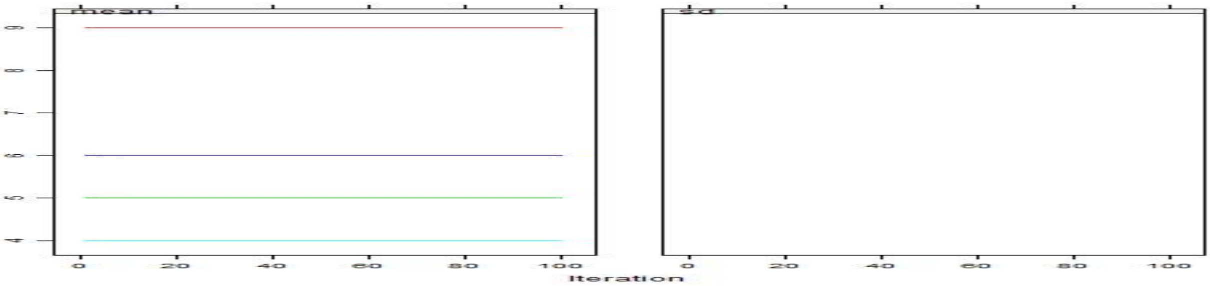

(b) **Post-test visit**



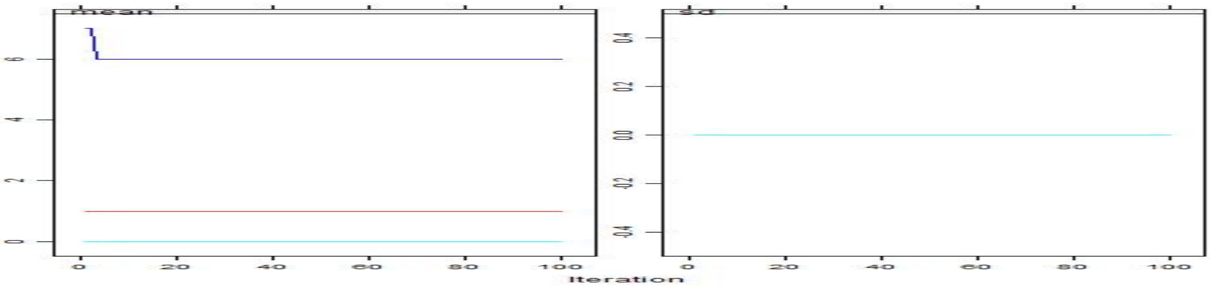

(c) **One year follow-up visit**

## Physical Discipline

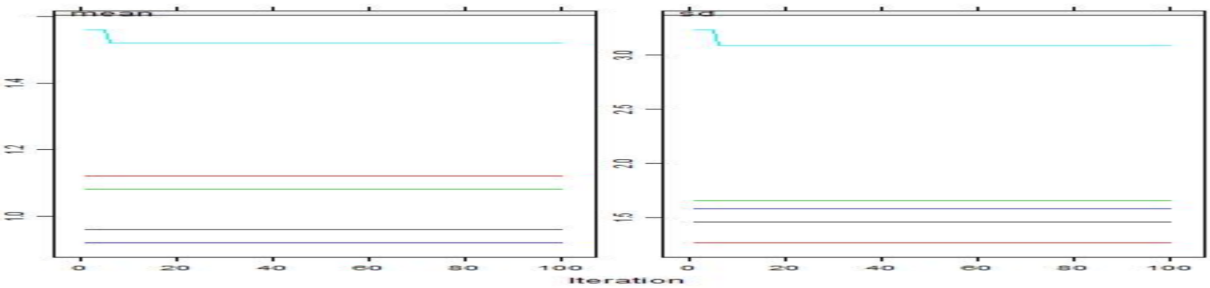
(a) **Baseline visit**



(b) **Post-test visit**

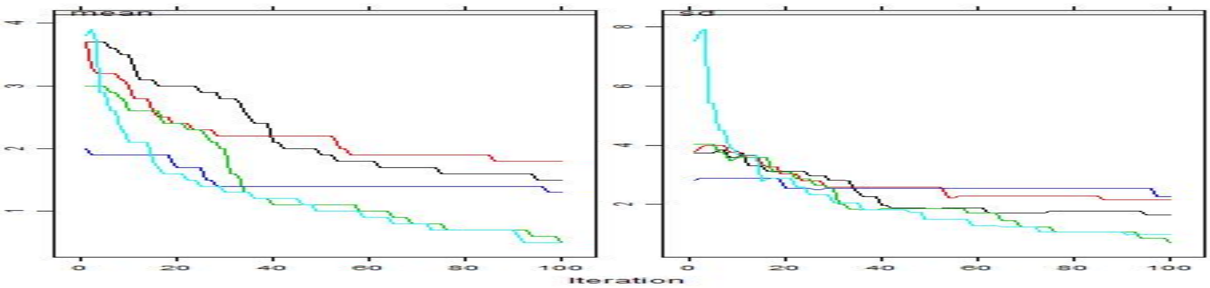


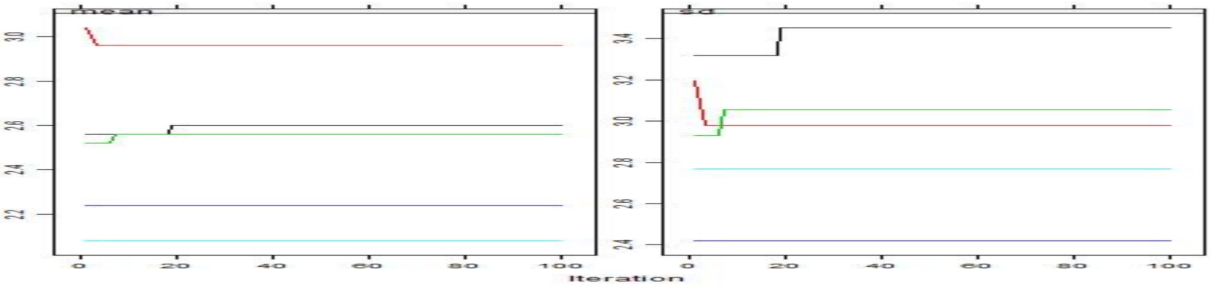
(c) **One year follow-up visit**



## Psychological Discipline

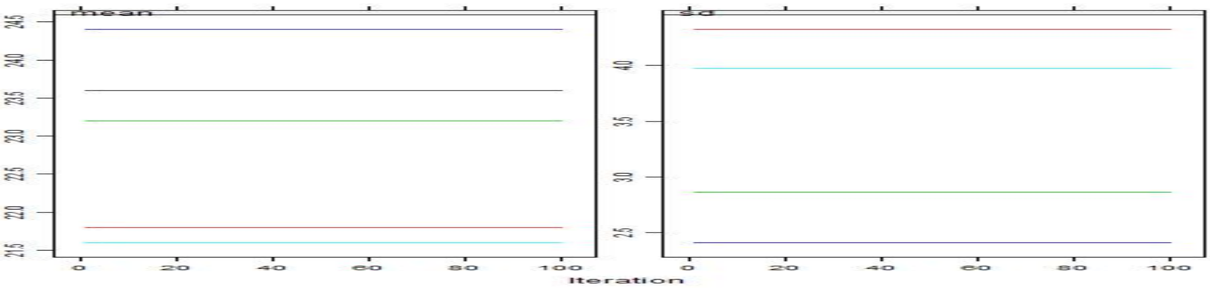
(a) **Post-test visit**



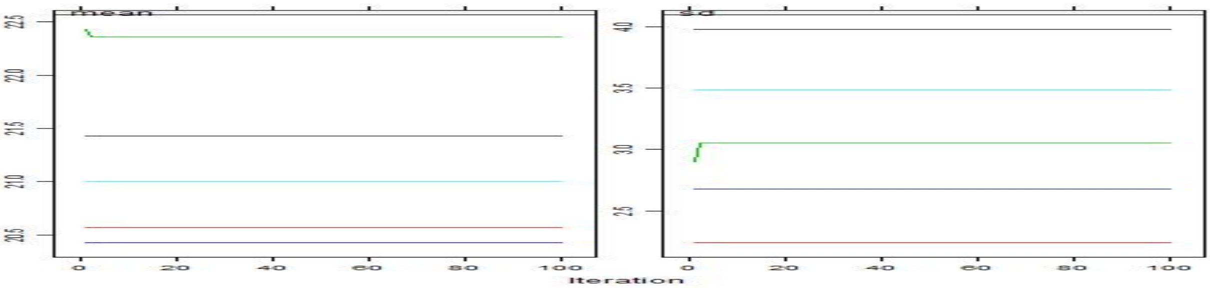



(a) **Post-test visit**

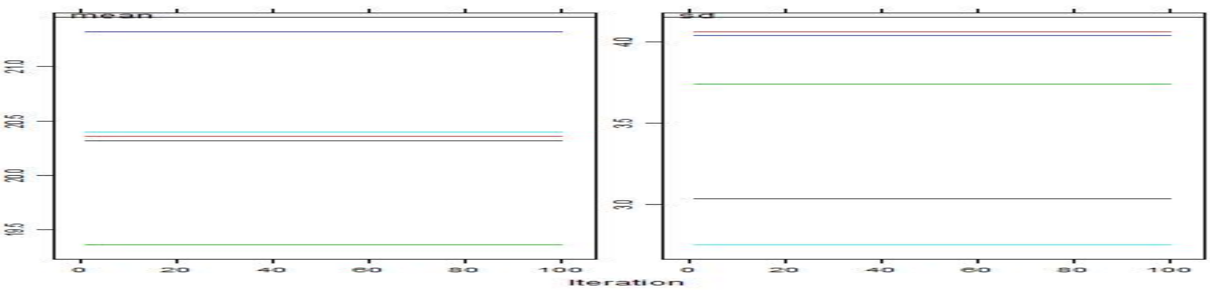
## Poor Monitoring And Supervision




(a) **Baseline visit**



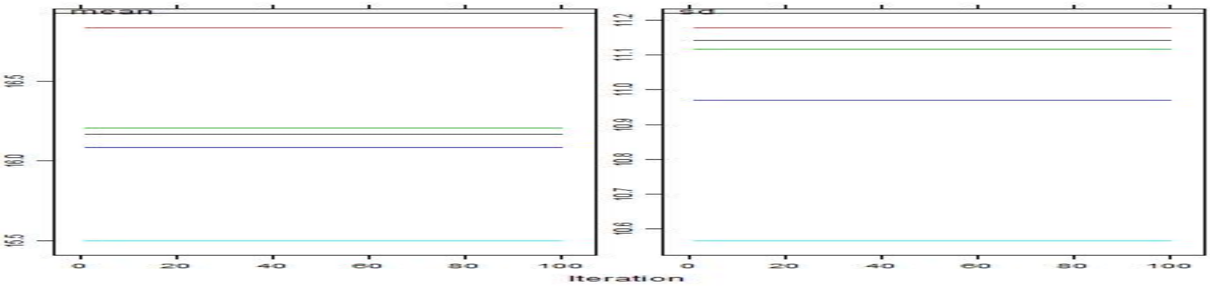

(b) **Post-test visit**



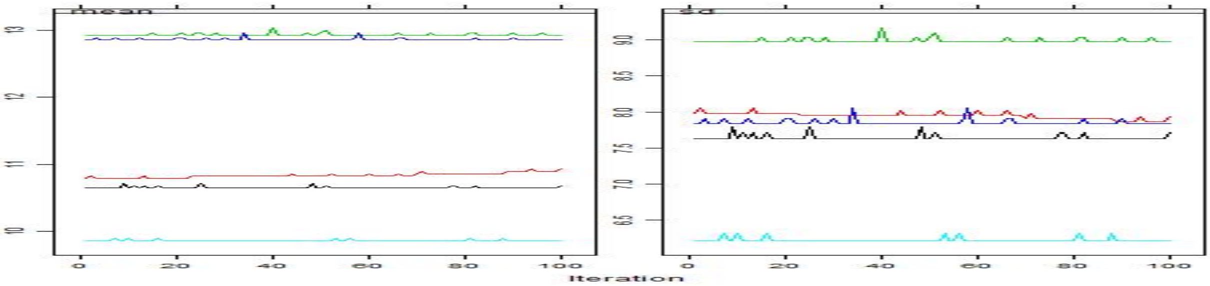

(c) **One-year follow-up visit**

## Beck Depression Inventory

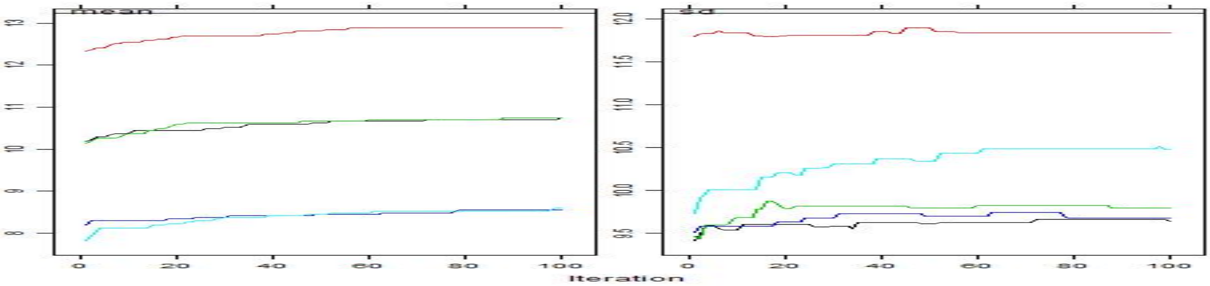
(a) **Baseline visit**



(b) **Post-test visit**

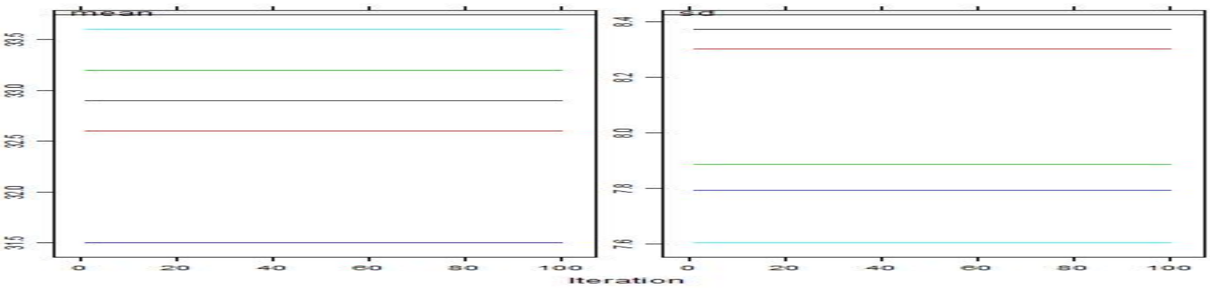


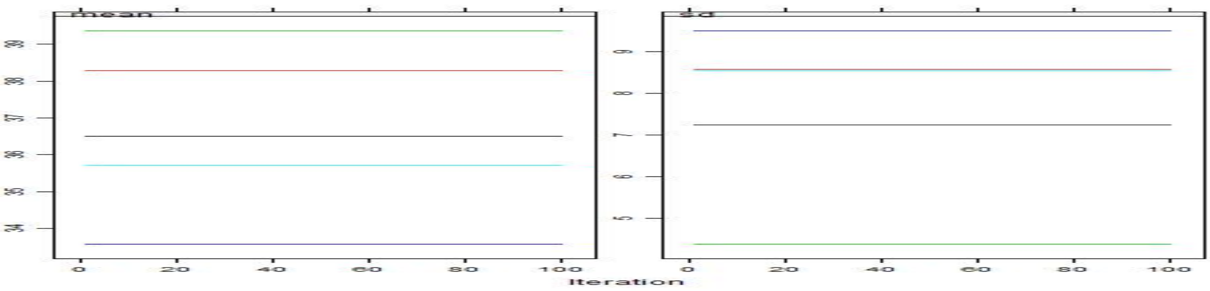
(c) **One-year follow-up visit**



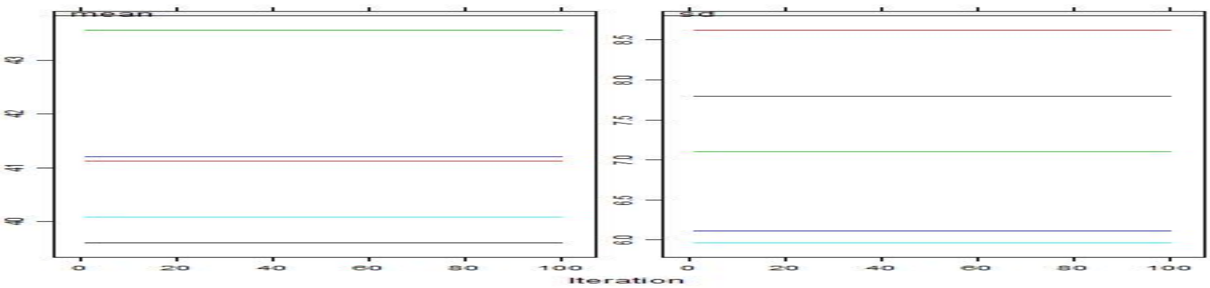
## Parental Distress

(a) **Baseline visit**



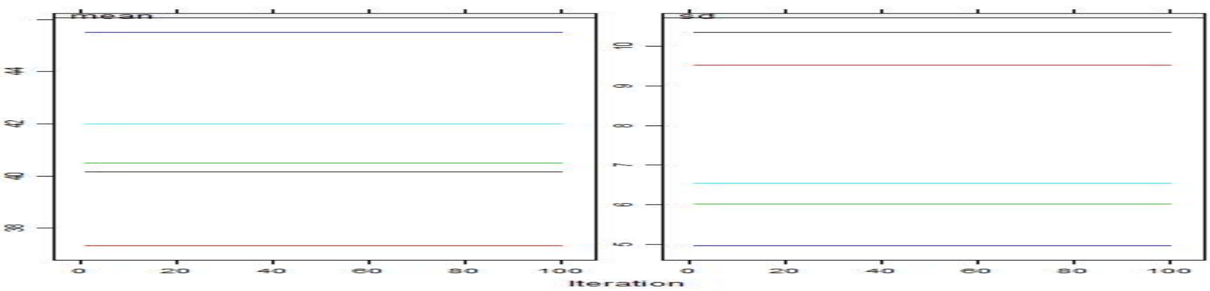


(a) Post-test visit

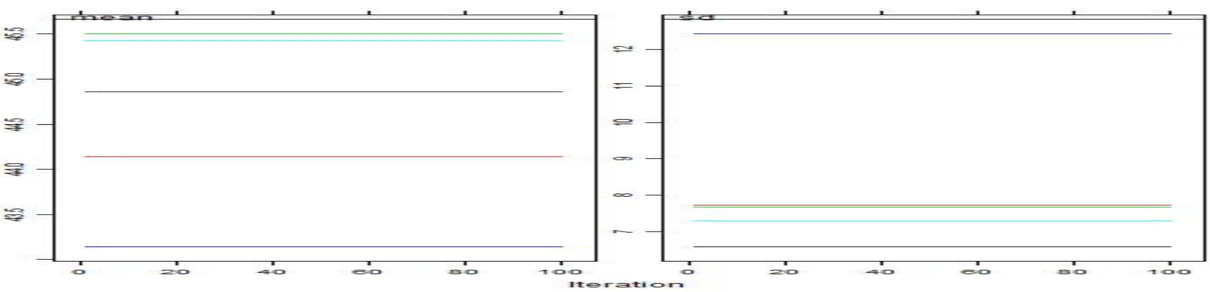


(b) One year follow-up visit

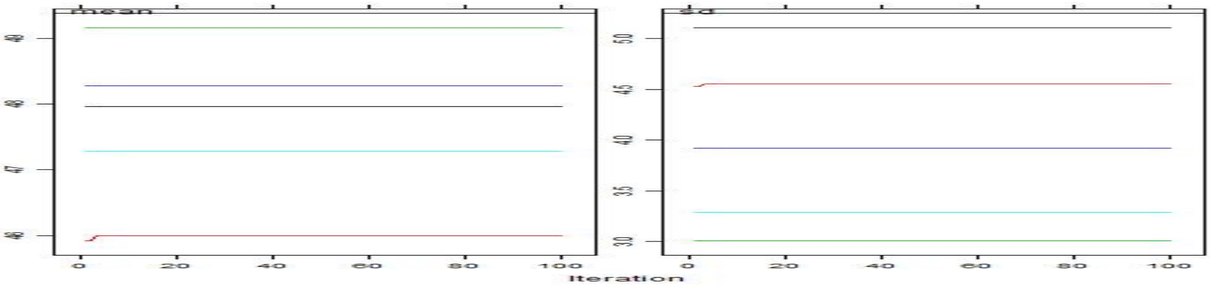
## Parent Child Dysfunctional Interaction



(a) **Baseline visit**

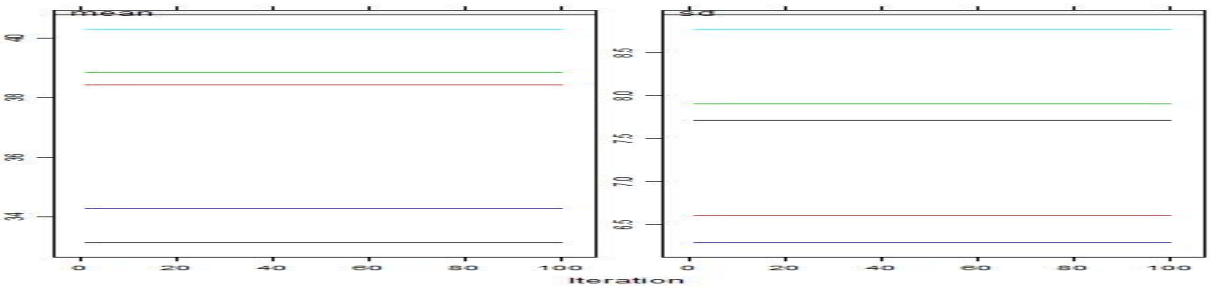


(b) **Post-test visit**

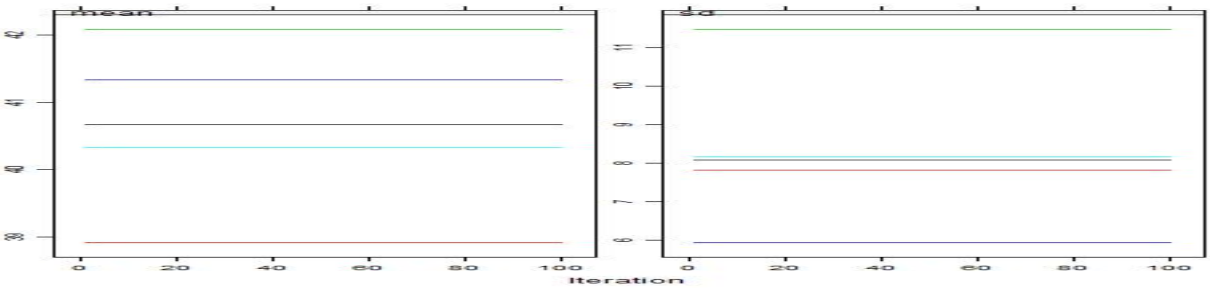


(a) **One-year follow-up visit**

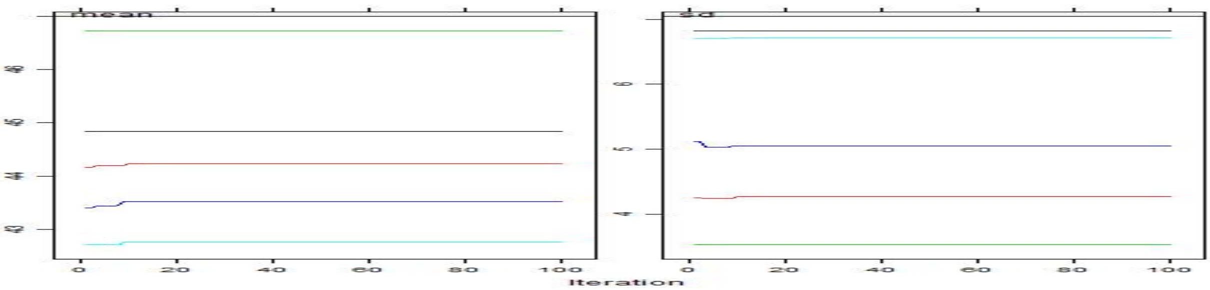
## Difficult Child



(a) **Baseline visit**



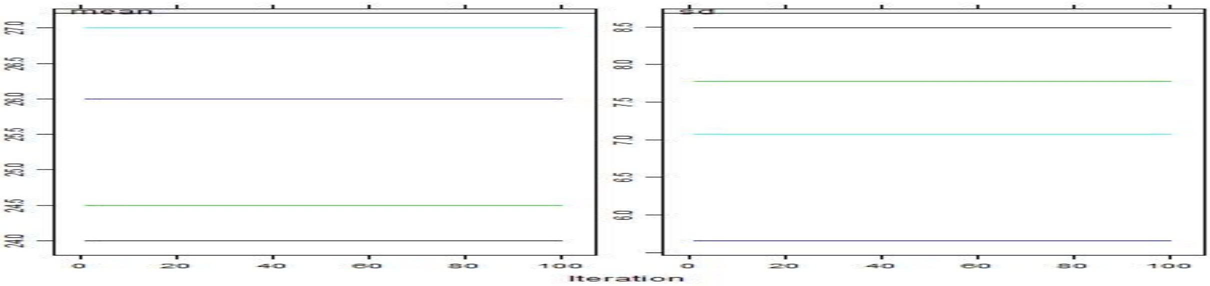
(b) **Post-test visit**



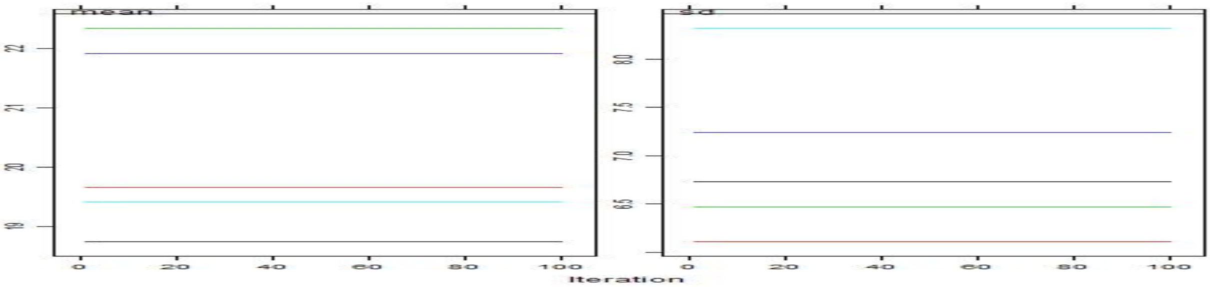
(c) **One-year follow-up visit**

## Social Support

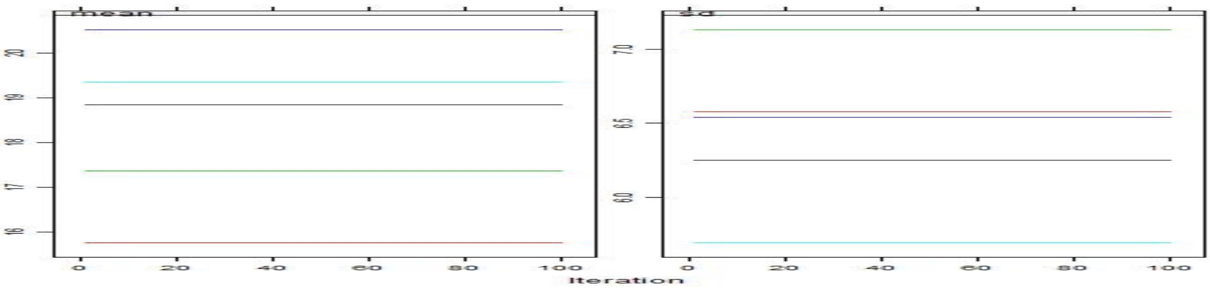
(a) **Baseline visit**




(b) **Post-test visit**

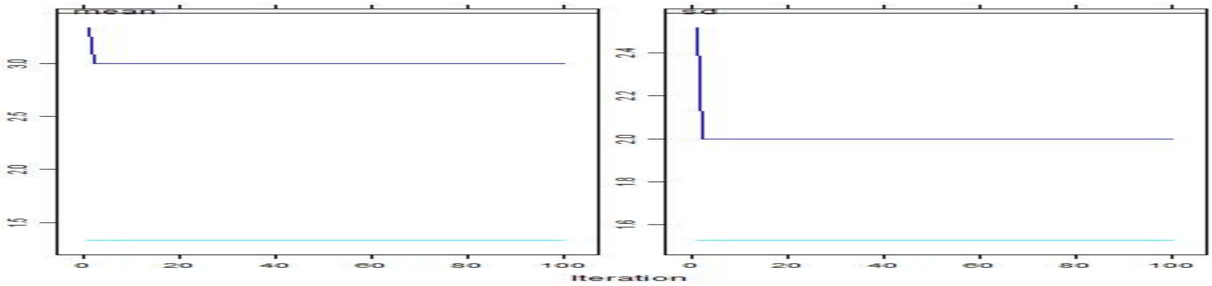



(c) **One-year follow-up visit**

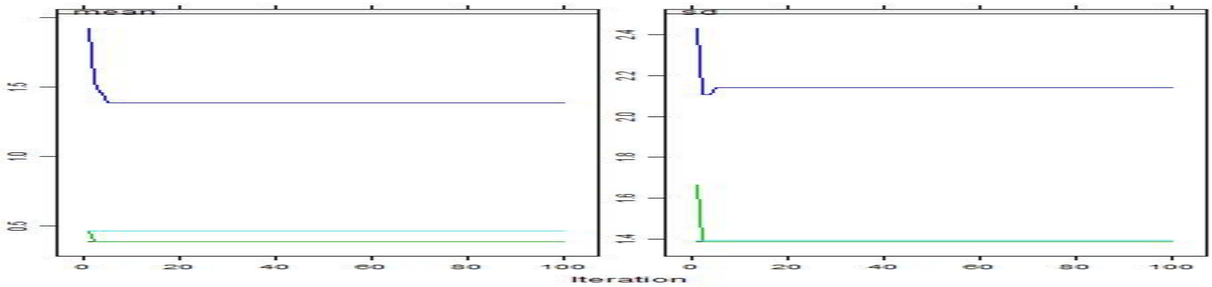


## IPV Chronicity

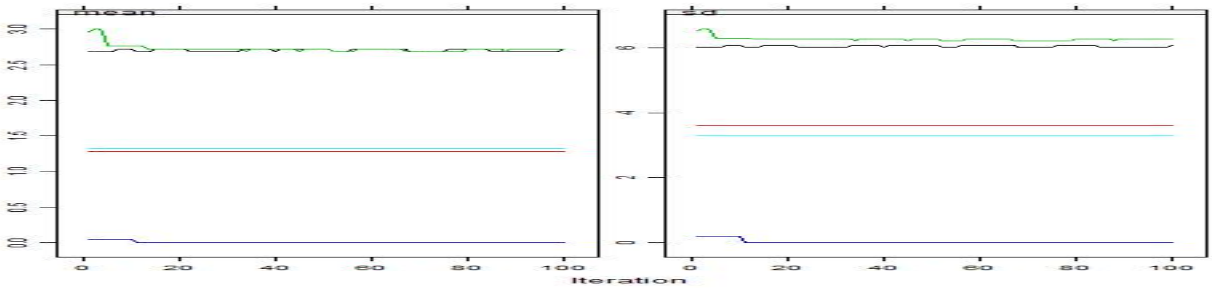
(a) **Baseline visit**



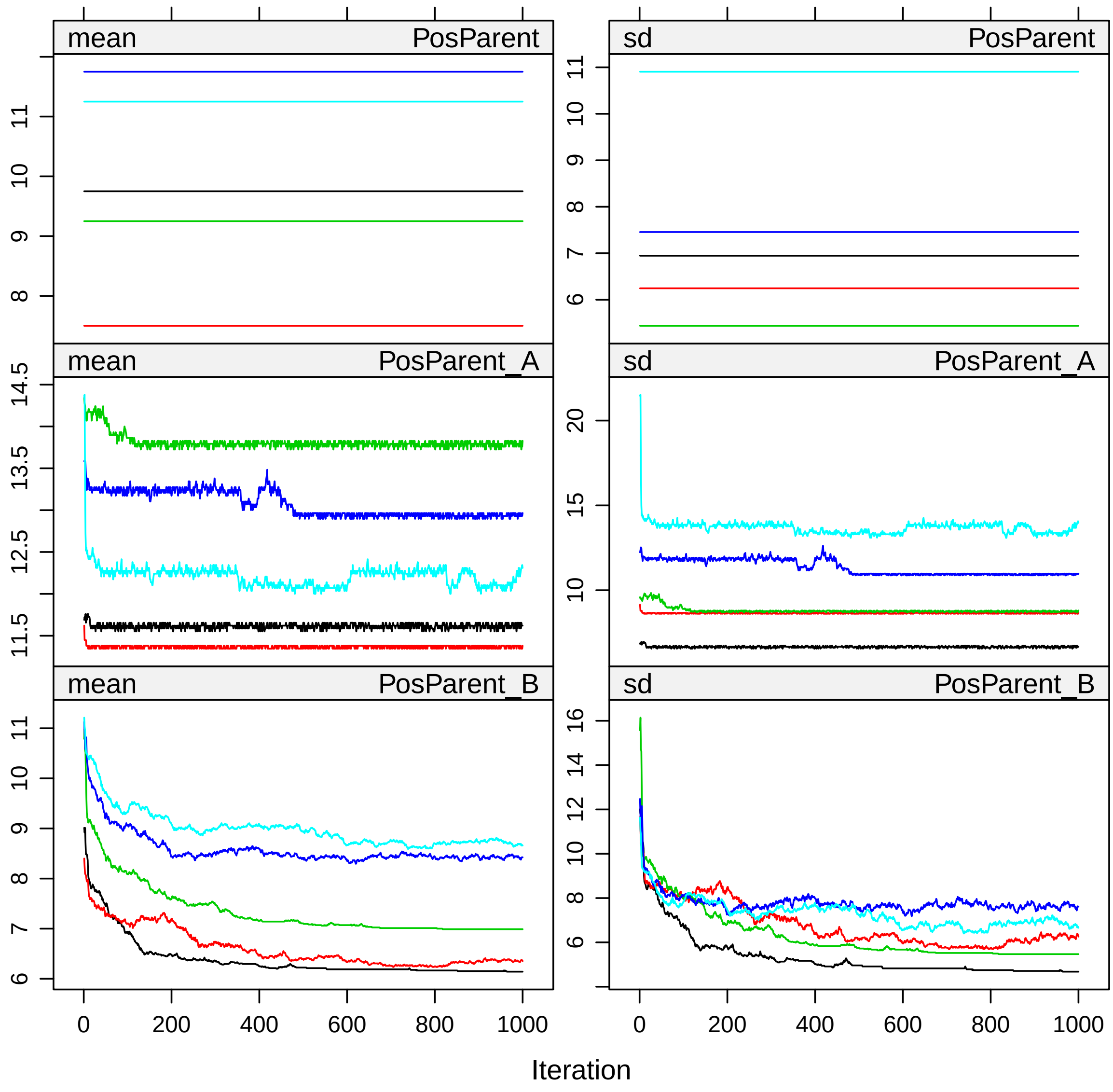

(b) **Post-test visit**



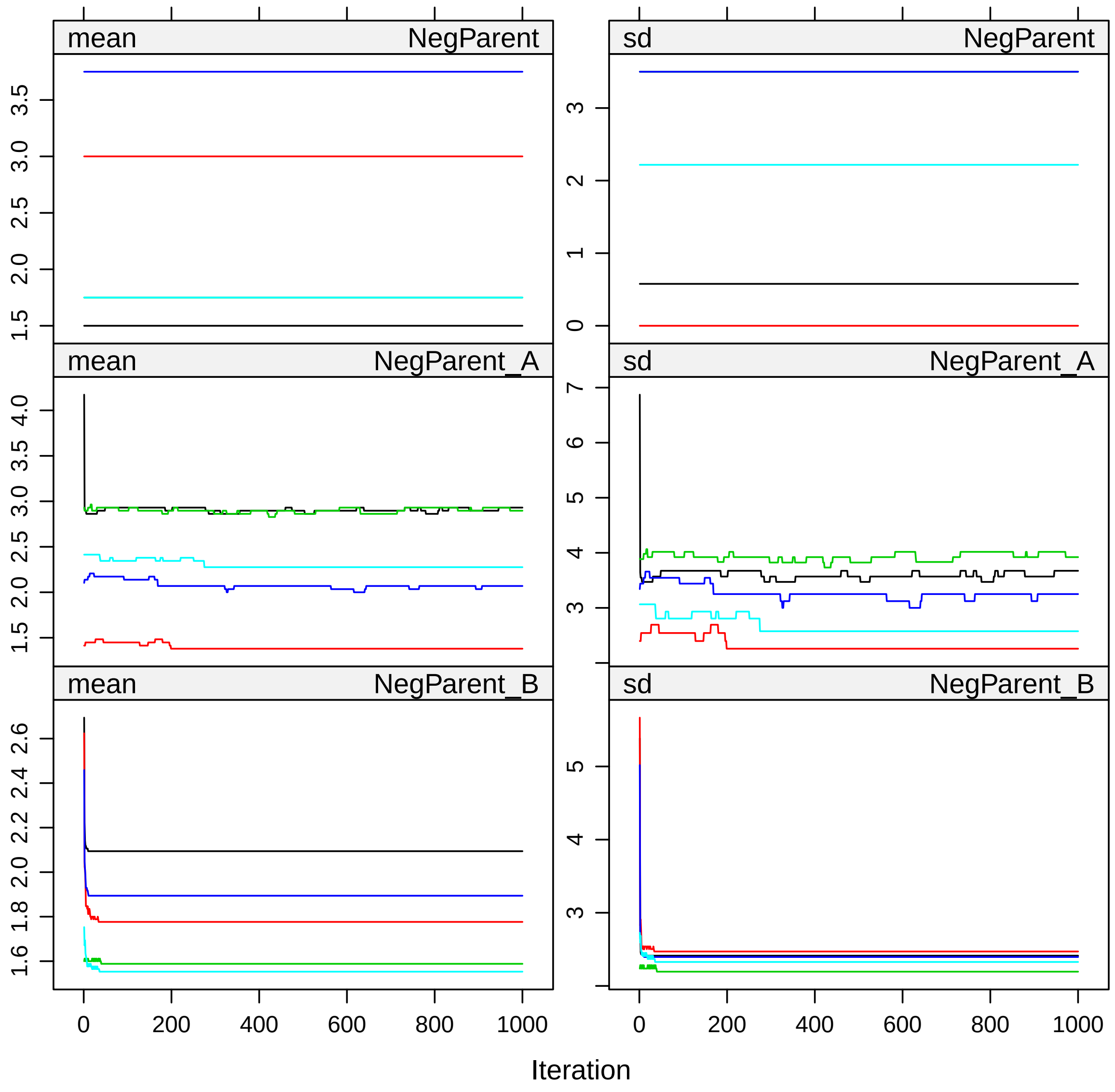

(c) **One-year follow-up visit**



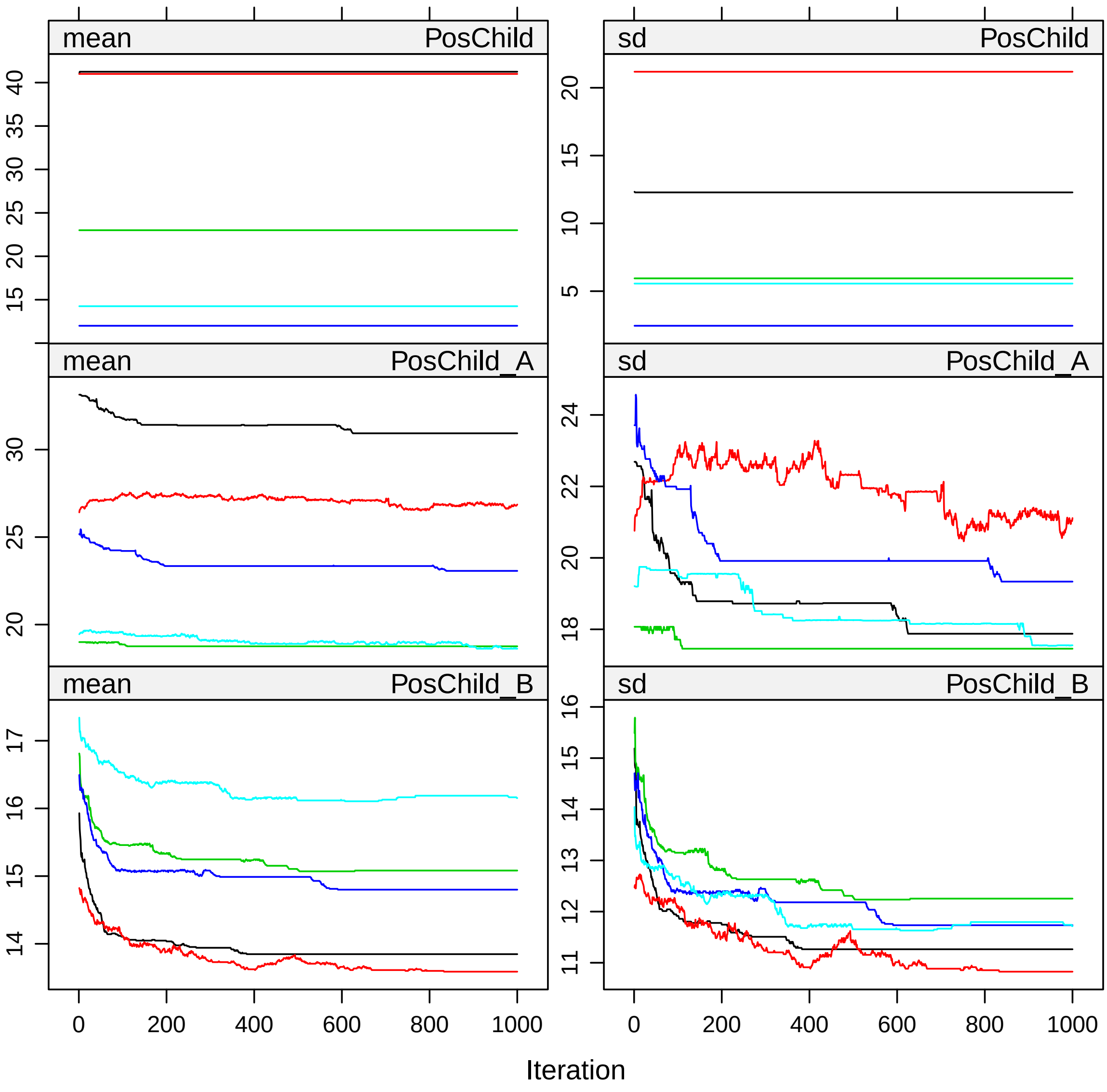
**Parent Positive Behaviour**



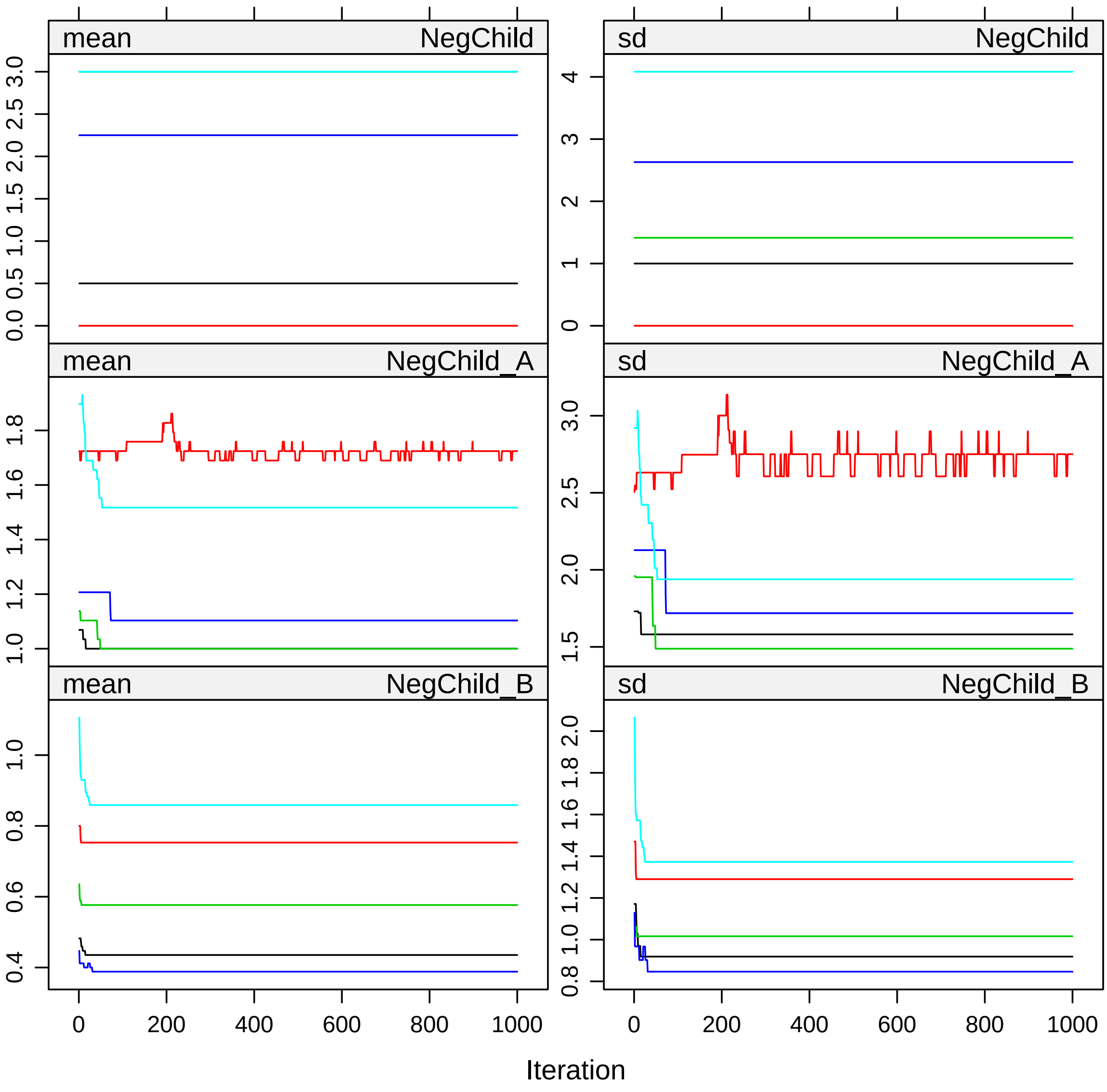
**Parent Negative Behaviour**



**Child Positive Behaviour**



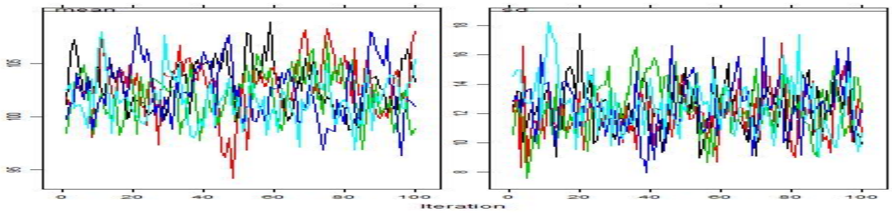
## Child Negative Behaviour



# Imputation using random forests: convergence checks

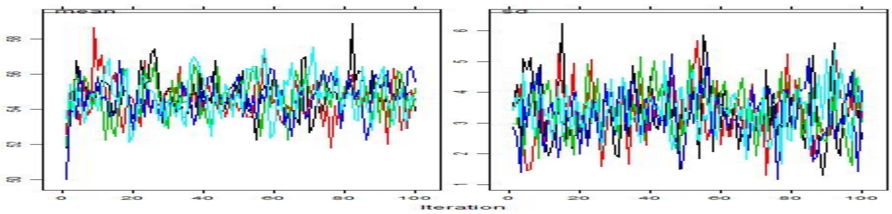
### ECBI Intensity

(a) **One year follow-up visit**

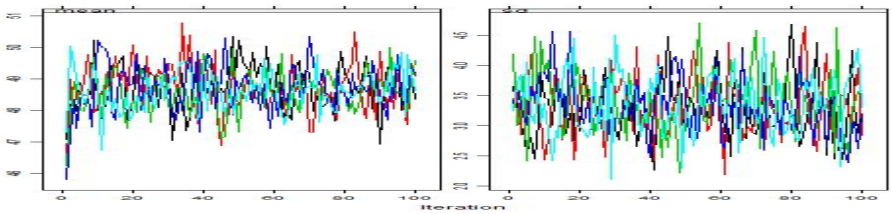


### ECBI Problem

(a) **Post-test visit**

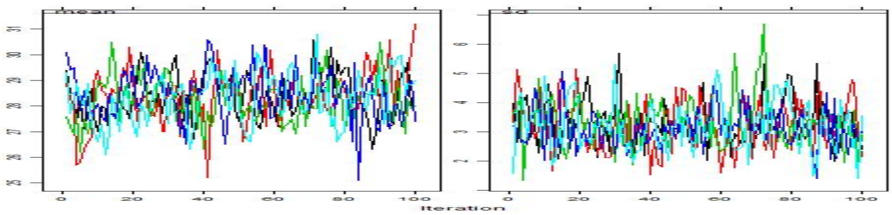


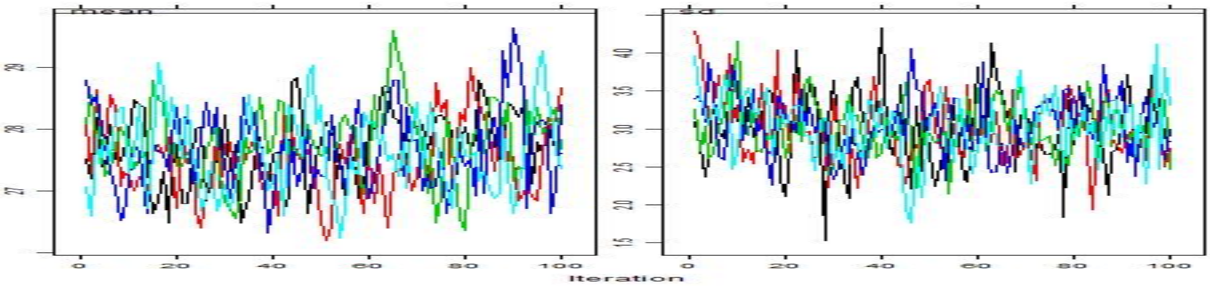
(b) **One year follow-up visit**



### Supporting Positive Behaviour Frequency

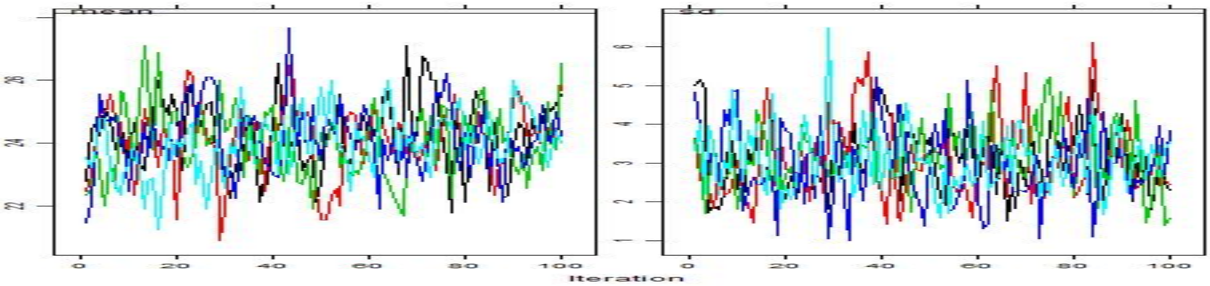
(a) **Post-test visit**



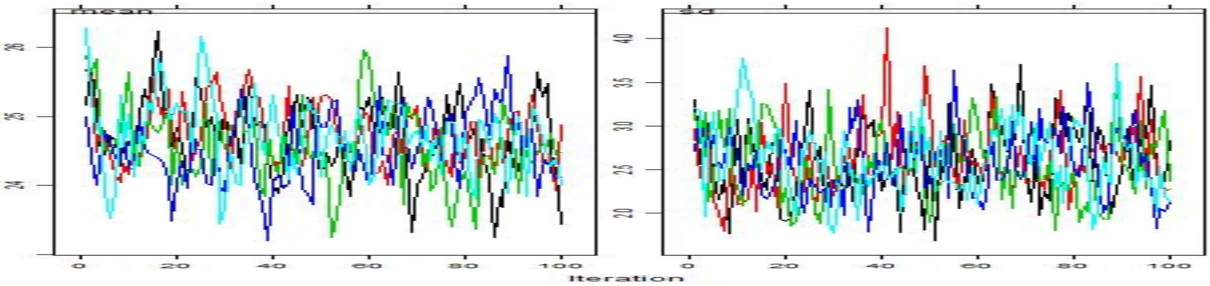


(b) One year follow-up visit

**Setting limits Frequency**

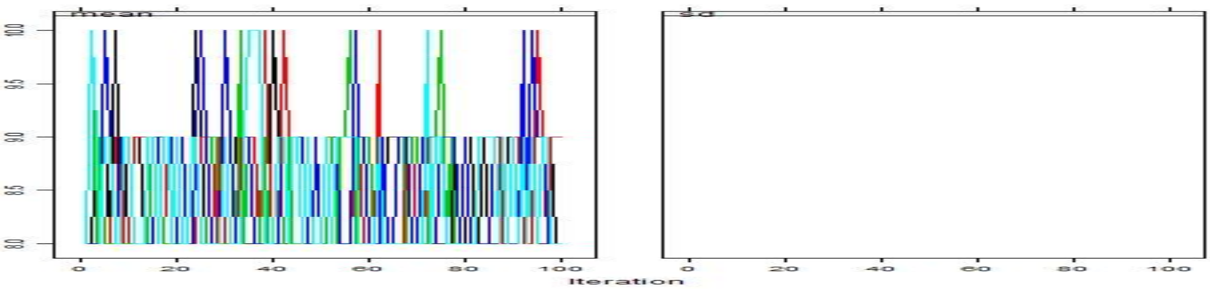


(a) **Post-test visit**



(b) **One year follow-up visit**

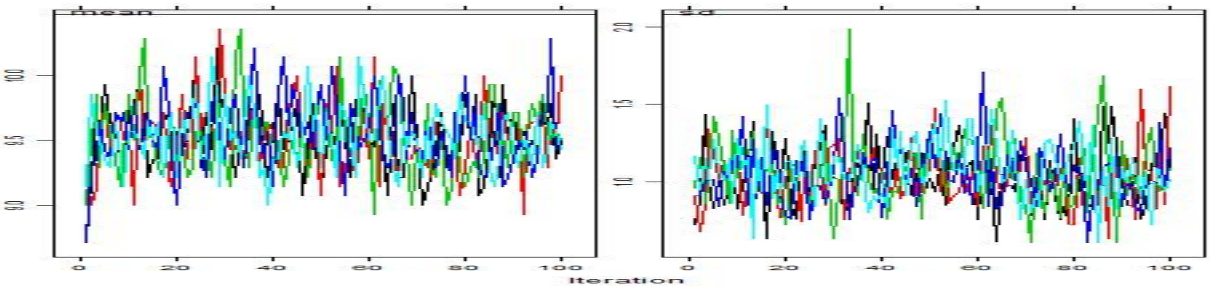
**Supporting Positive Behaviour Problem**



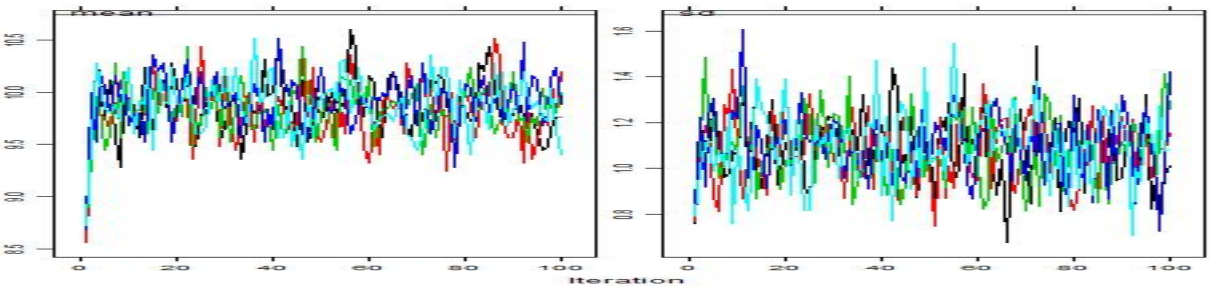
(a) **Baseline visit**

Figure 6: Supporting Positive Behaviour Problem

(a) Post-test visit

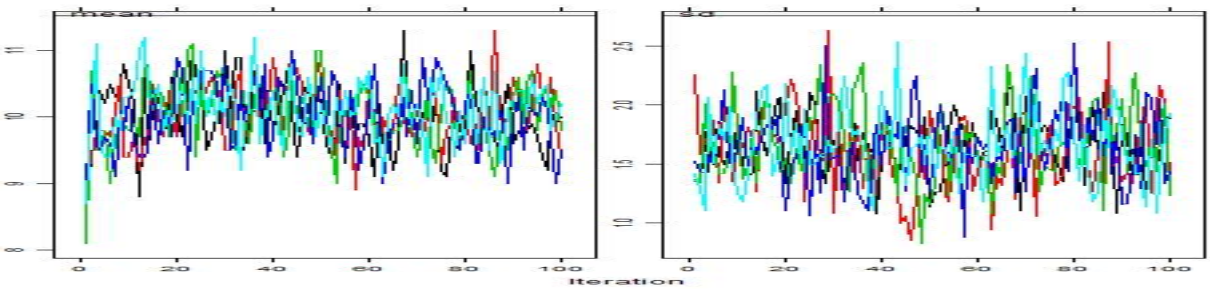


(b) One year follow-up visit

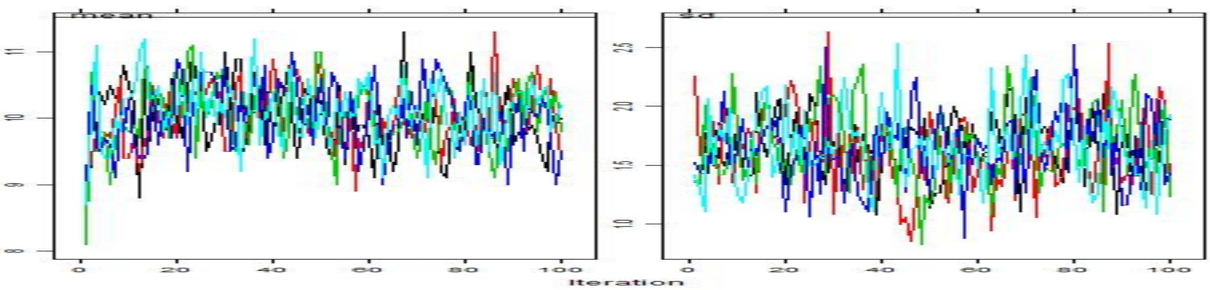


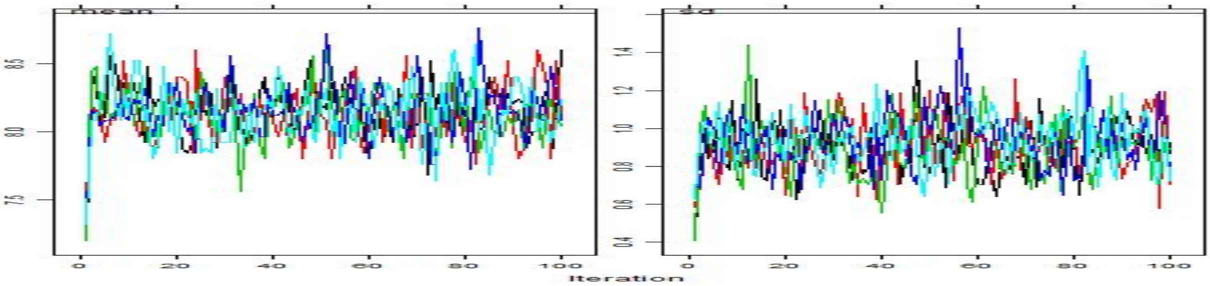
**Setting limits Problem**

(a) **Baseline visit**



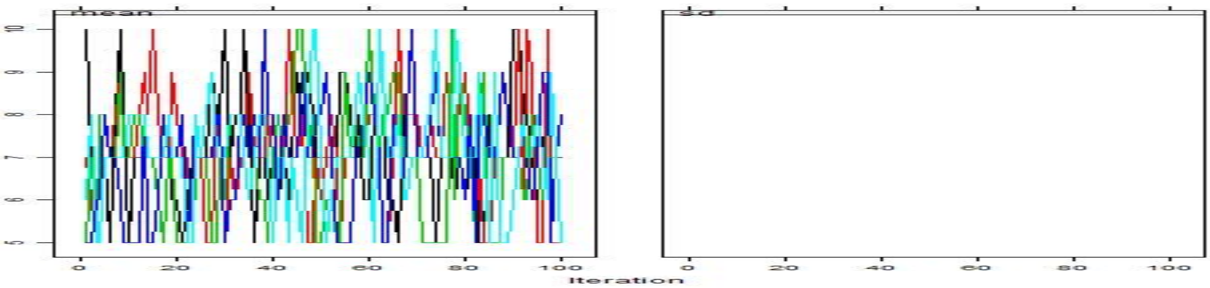
(b) **Post-test visit**



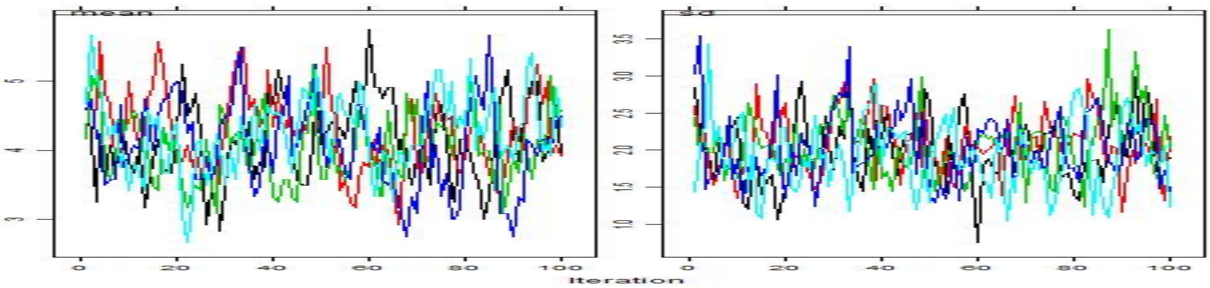


(a) **One year follow-up visit**

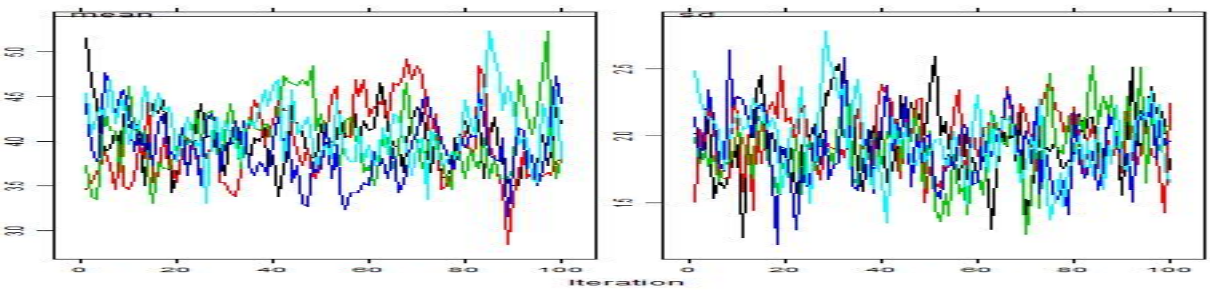
## Non-Violent Discipline



(a) **Baseline visit**



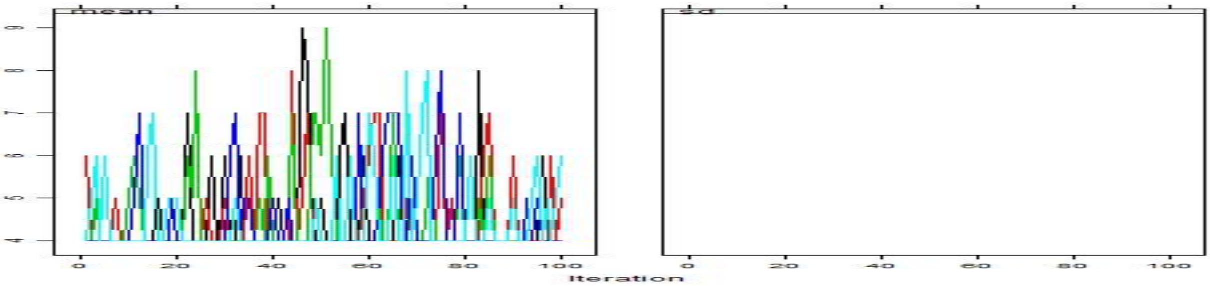
(b) **Post-test visit**



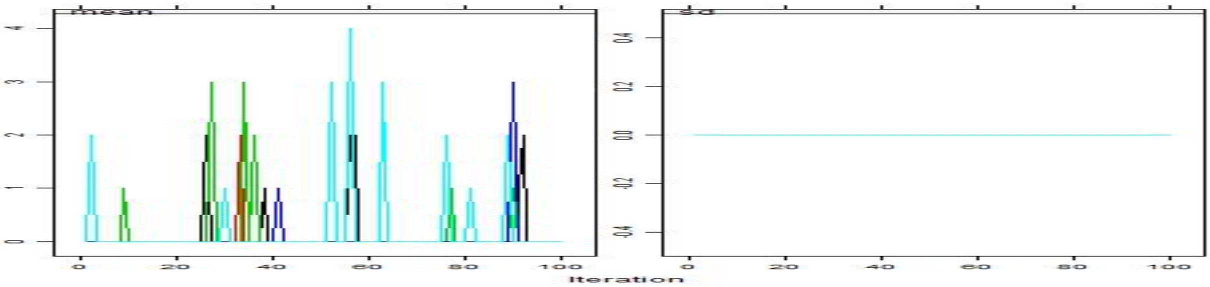
(c) **One year follow-up visit**

## Physical Discipline

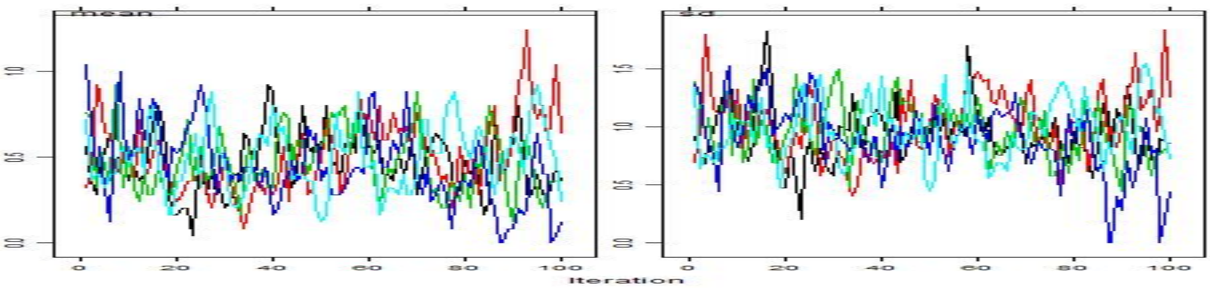
(a) **Baseline visit**



(b) **Post-test visit**

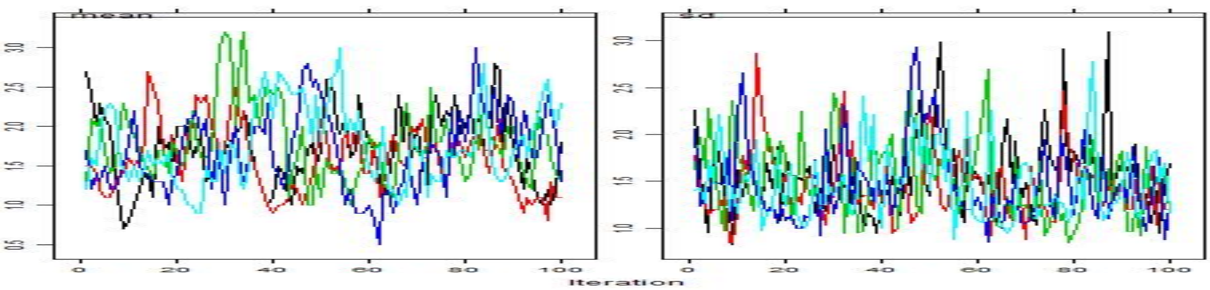


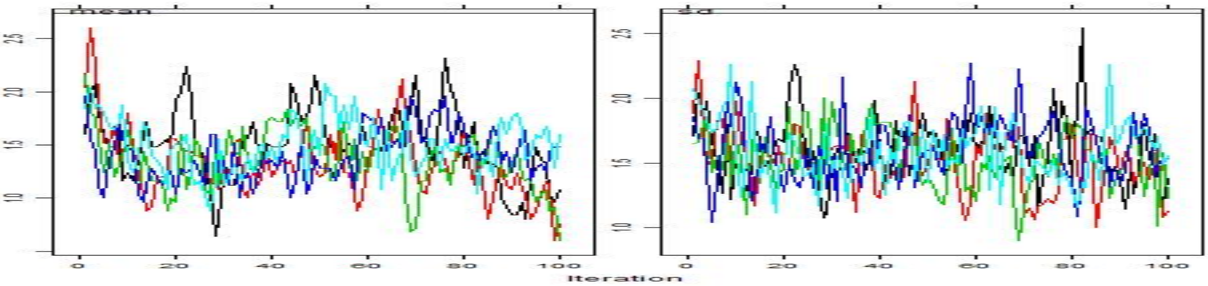
(c) **One year follow-up visit**



## Psychological Discipline

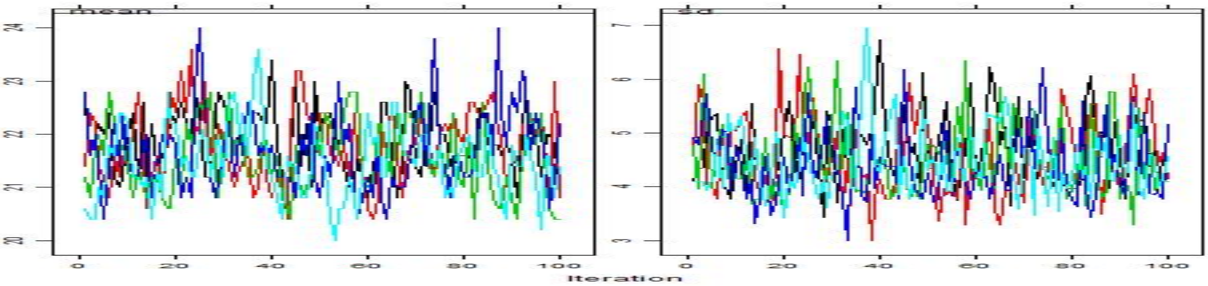
(a) **Post-test visit**



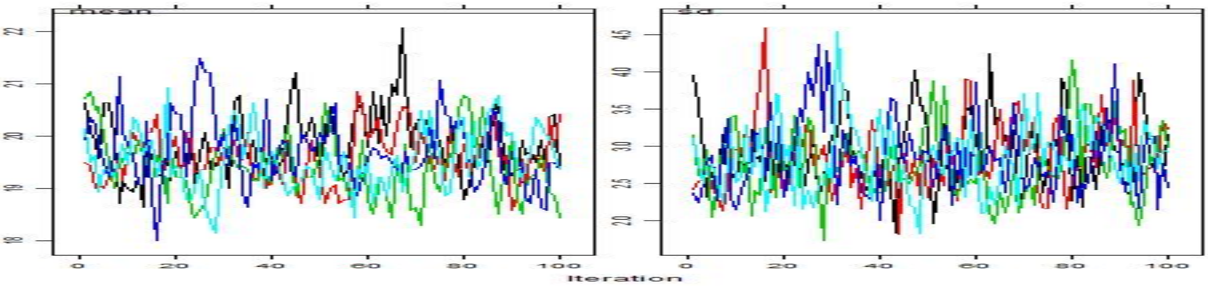


(a) **Post-test visit**

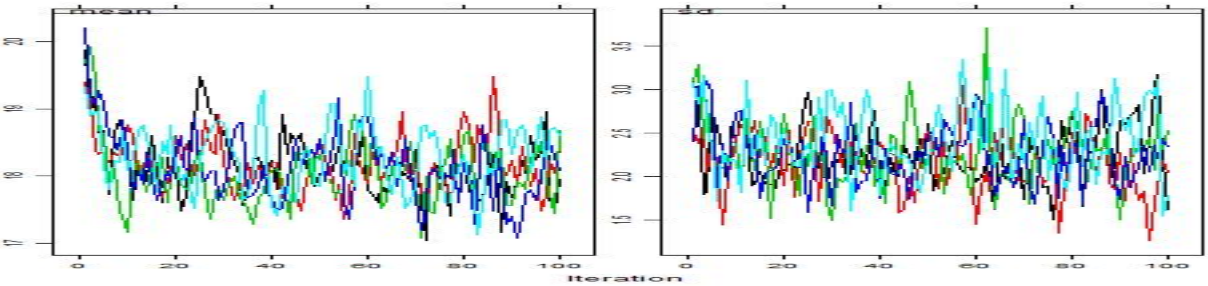
## Poor Monitoring And Supervision



(a) **Baseline visit**



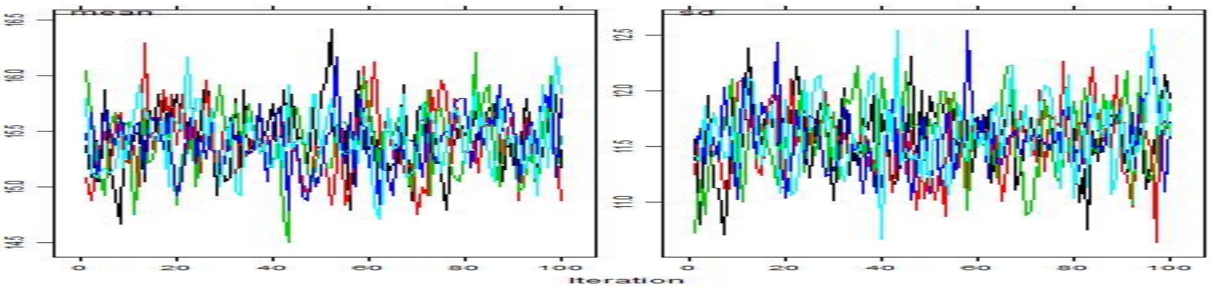
(b) **Post-test visit**



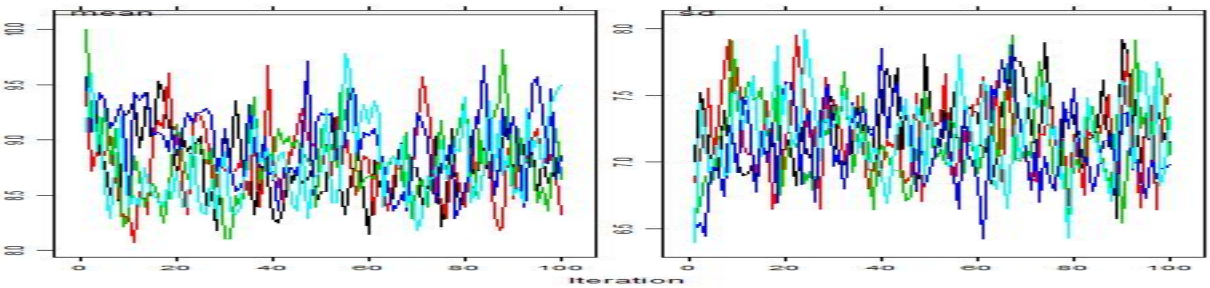
(c) **One-year follow-up visit**

## Beck Depression Inventory

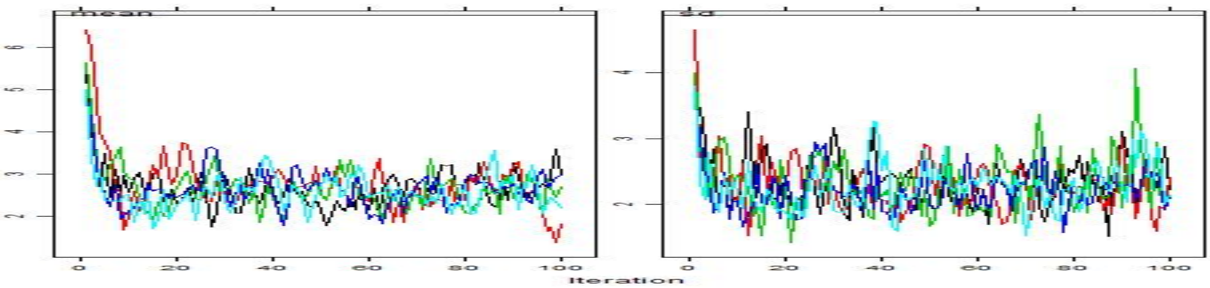
(a) **Baseline visit**




(b) **Post-test visit**

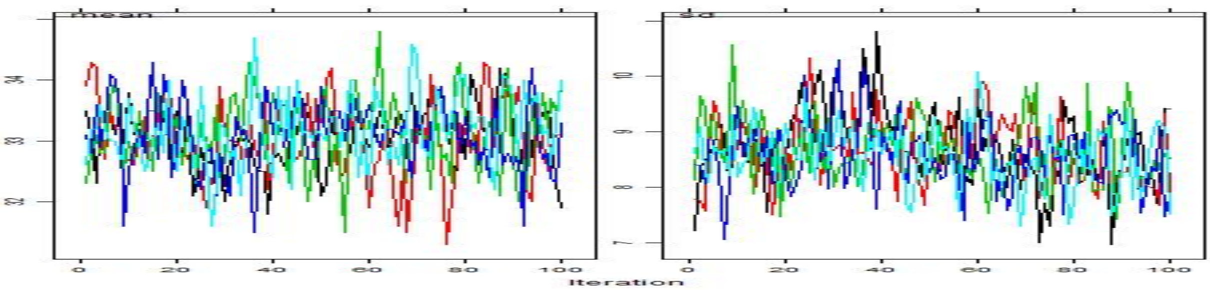


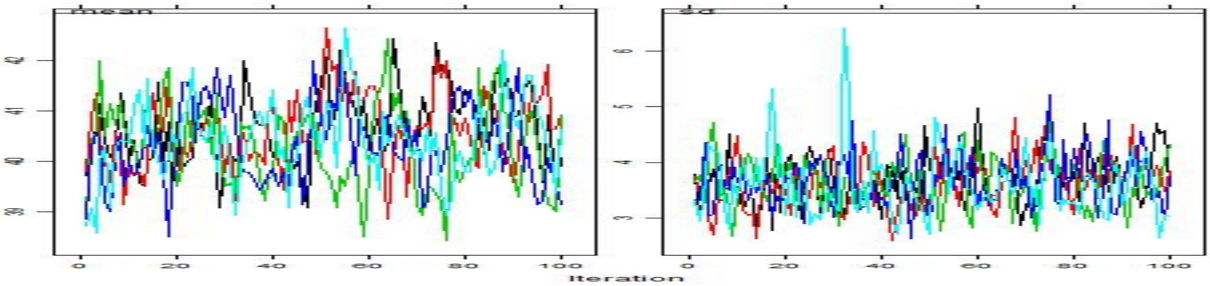

(c) **One-year follow-up visit**



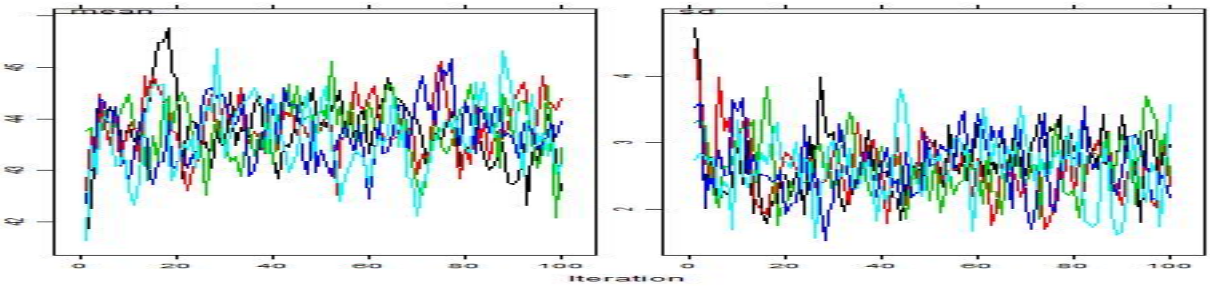

## Parental Distress

(a) **Baseline visit**



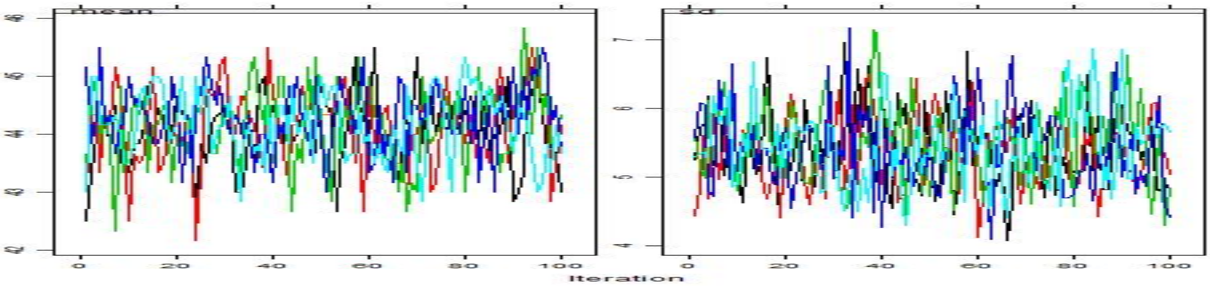



(a) Post-test visit

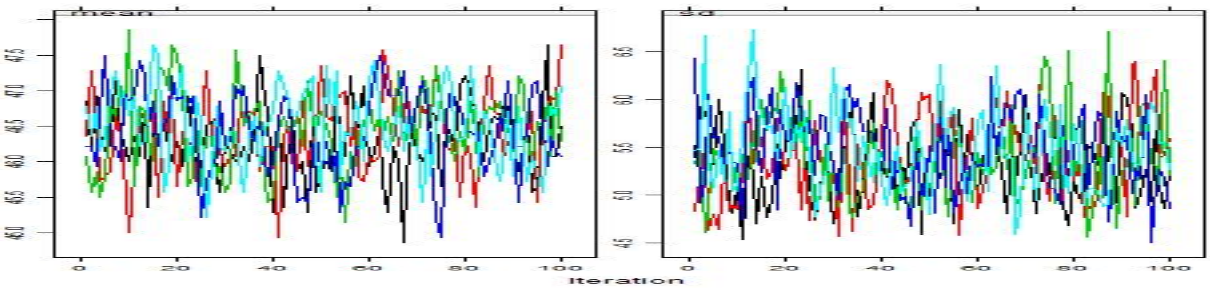



(b) One year follow-up visit

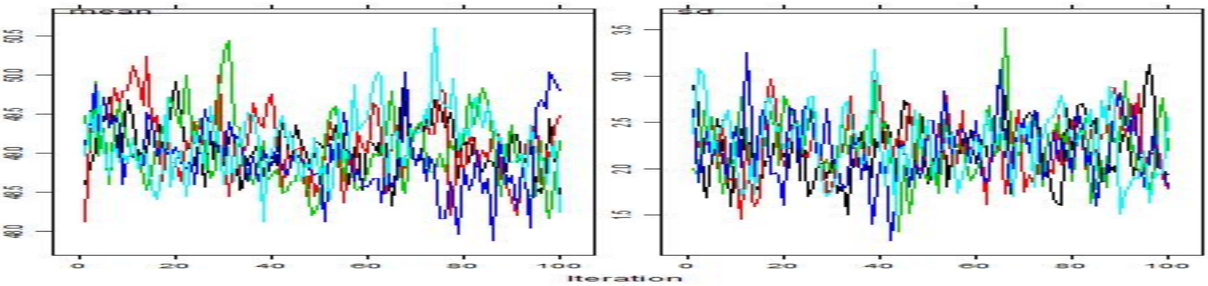
**Parent Child Dysfunctional Interaction**




(a) **Baseline visit**

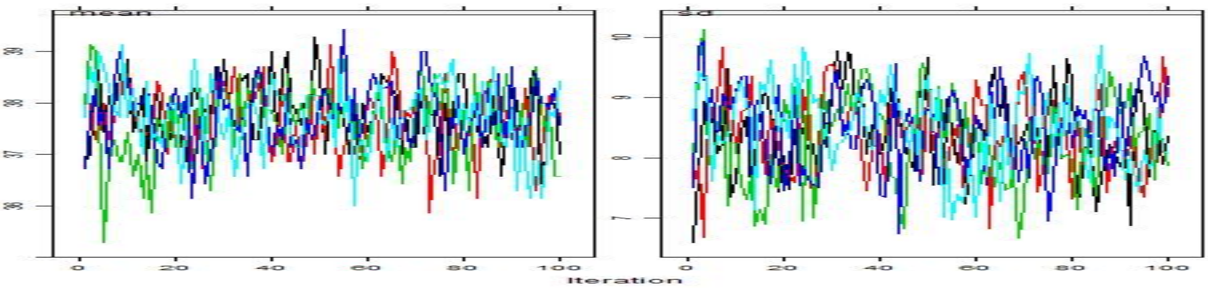



(b) **Post-test visit**

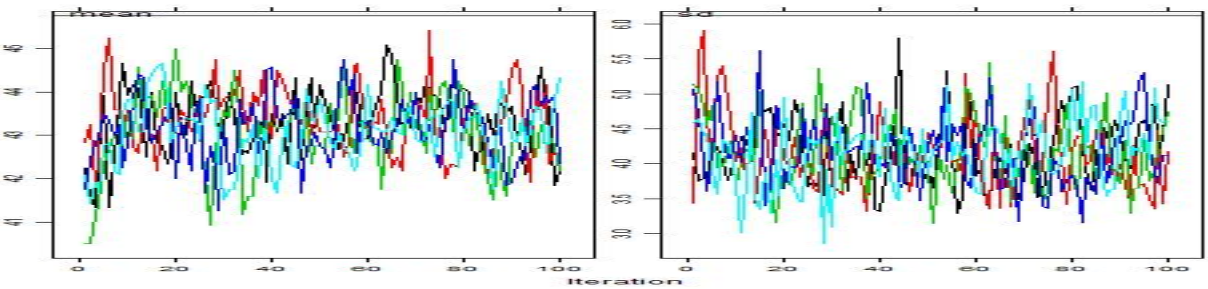


(a) **One-year follow-up visit**

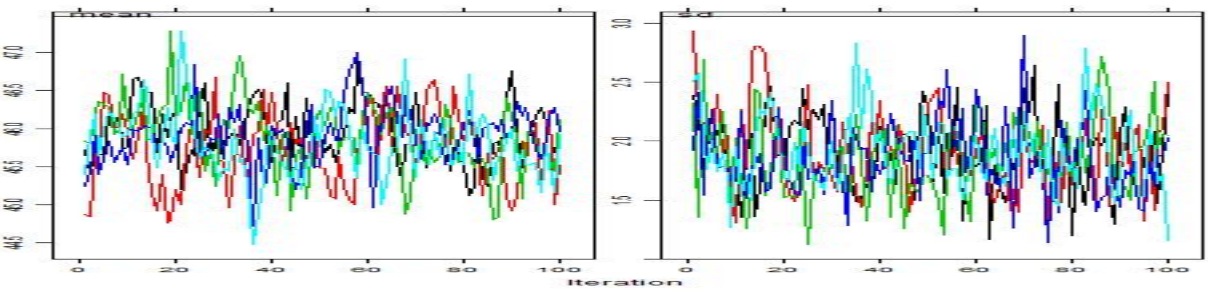
## Difficult Child



(a) **Baseline visit**



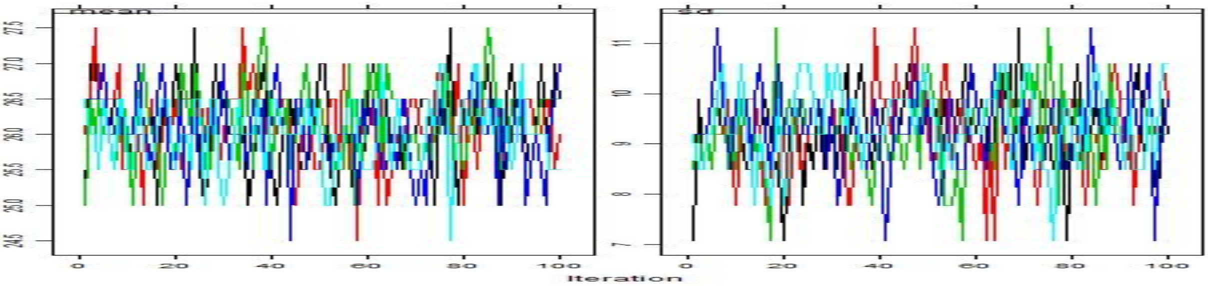
(b) **Post-test visit**



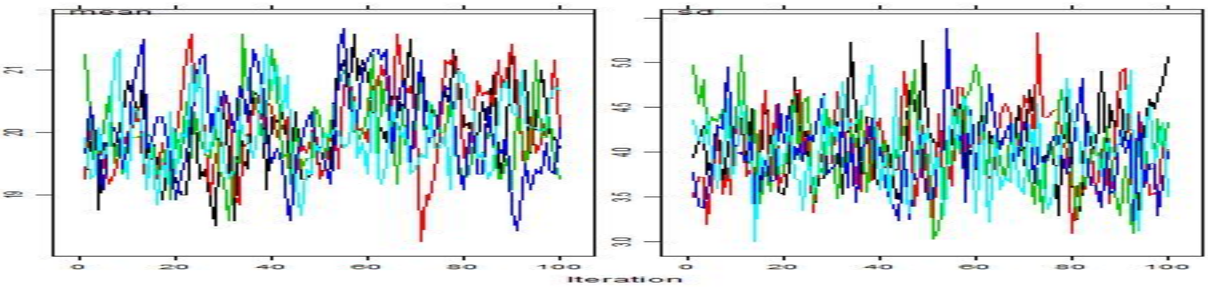
(c) **One-year follow-up visit**

## Social Support

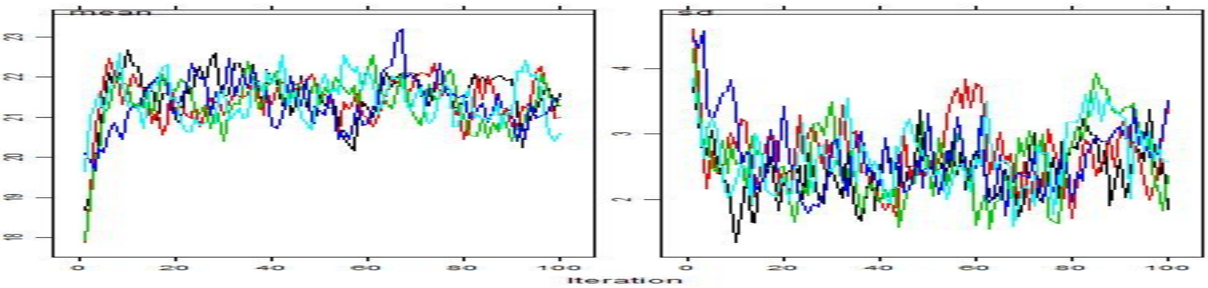
(a) **Baseline visit**




(b) **Post-test visit**

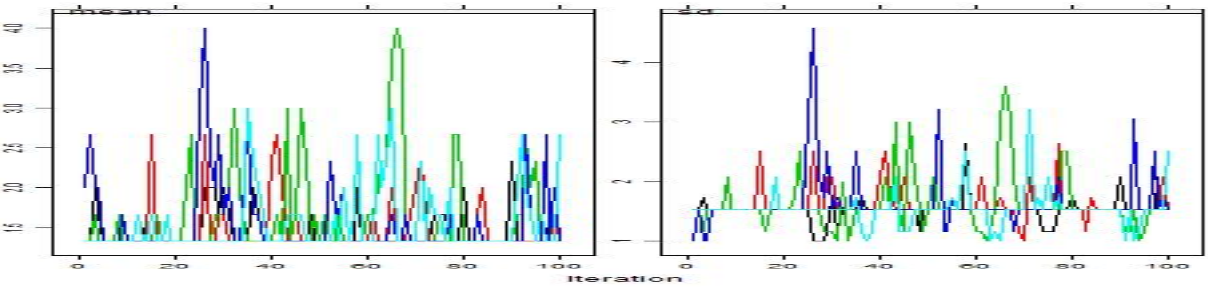



(c) **One-year follow-up visit**

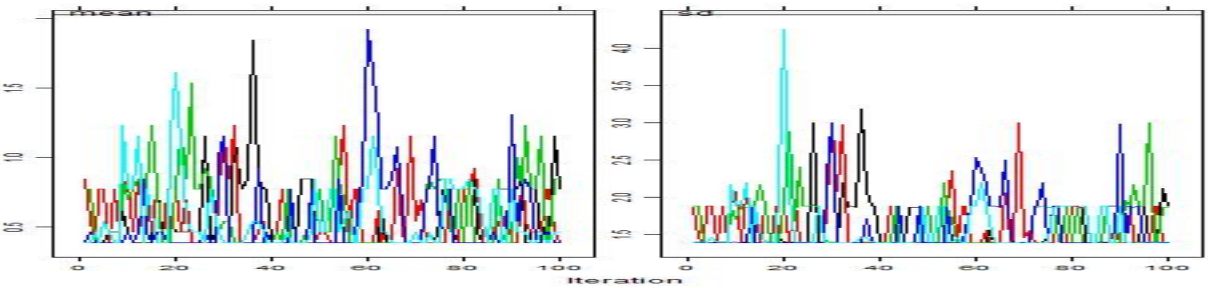


## IPV Chronicity

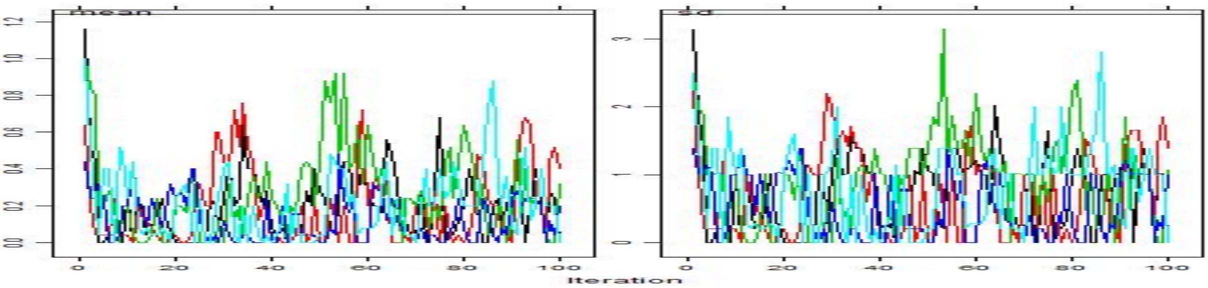
(a) **Baseline visit**



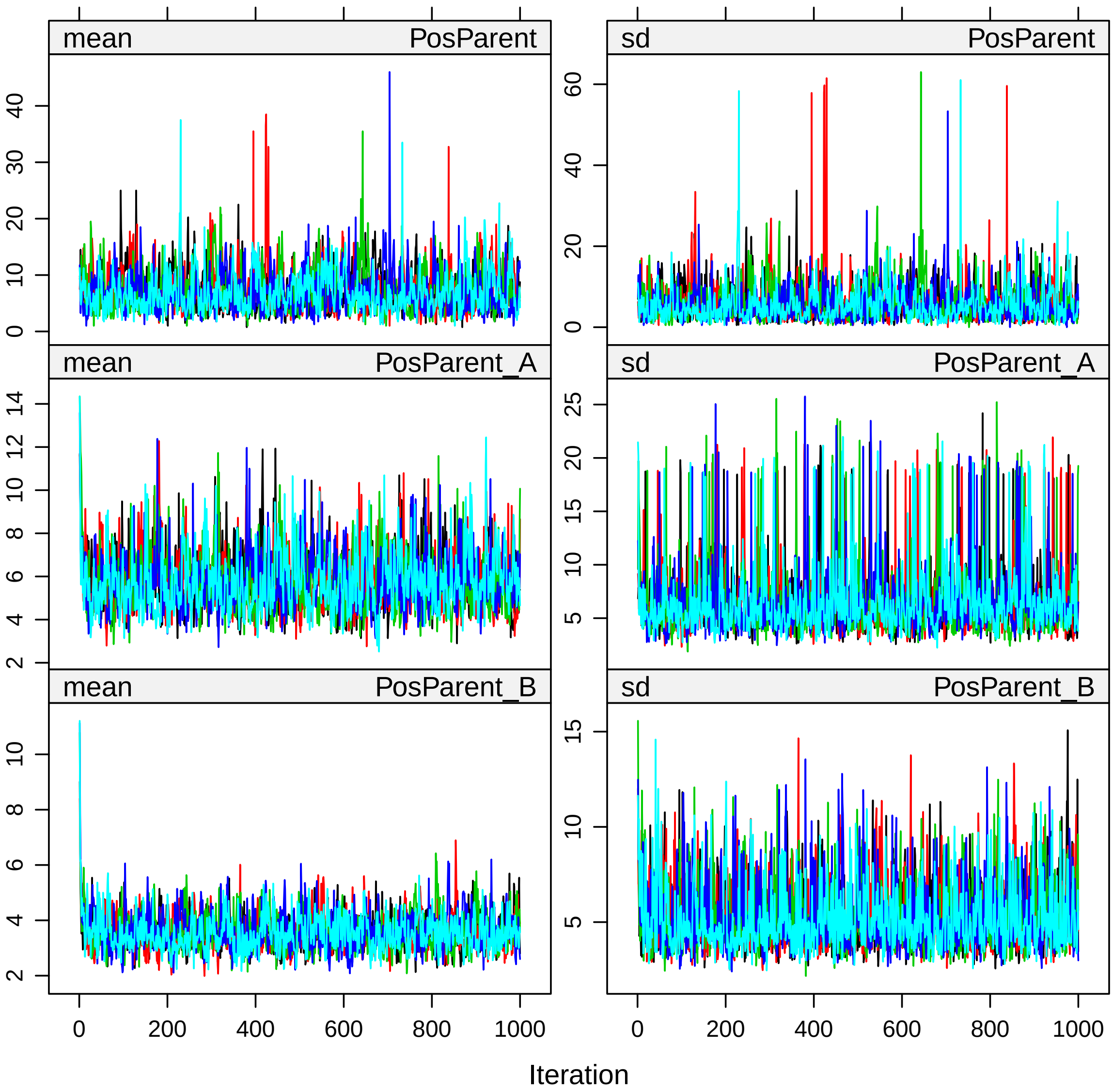
(b) **Post-test visit**



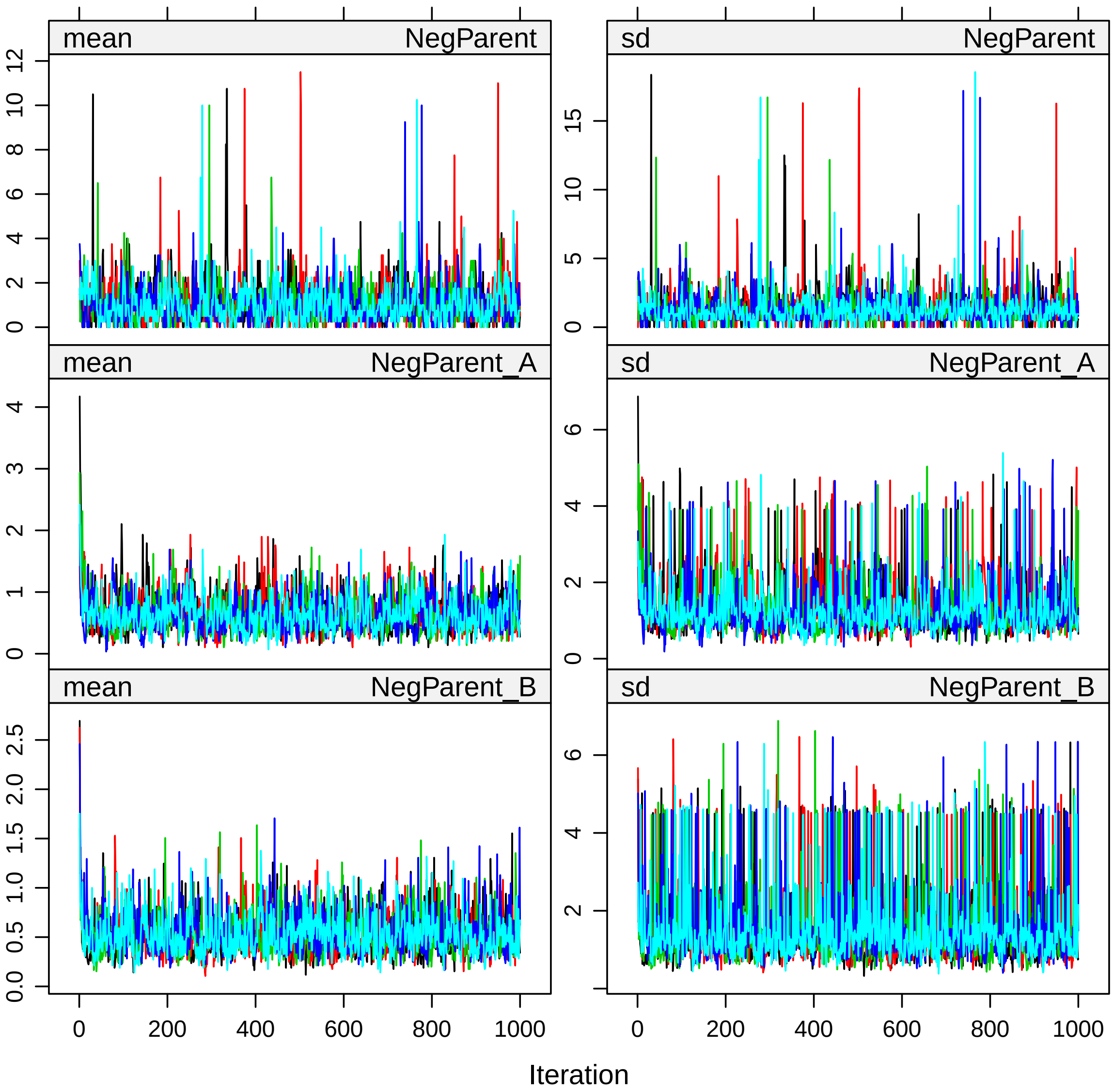
(c) **One-year follow-up visit**



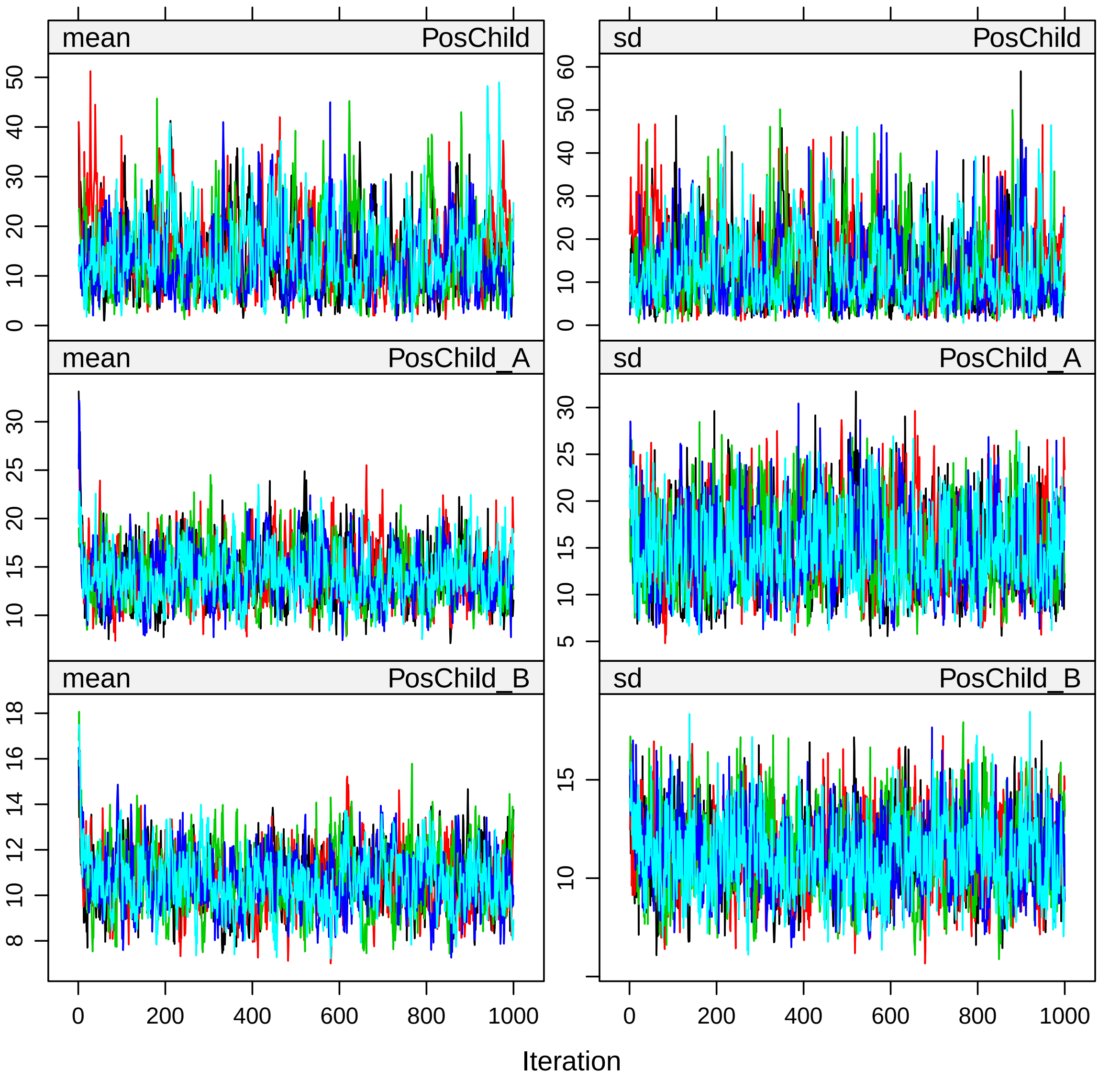
**Parent Positive Behaviour**



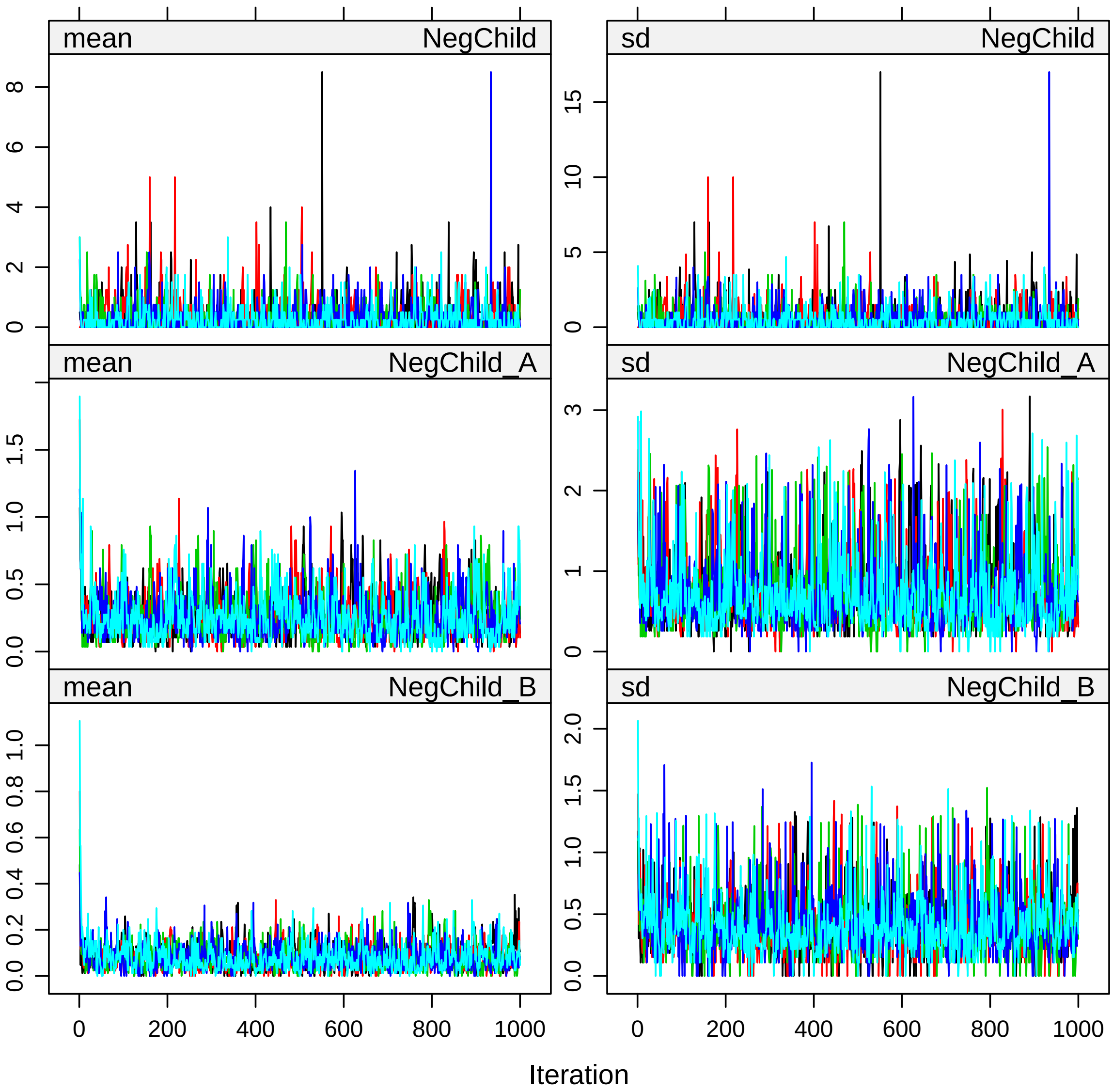
**Parent Negative Behaviour**



**Child Positive Behaviour**



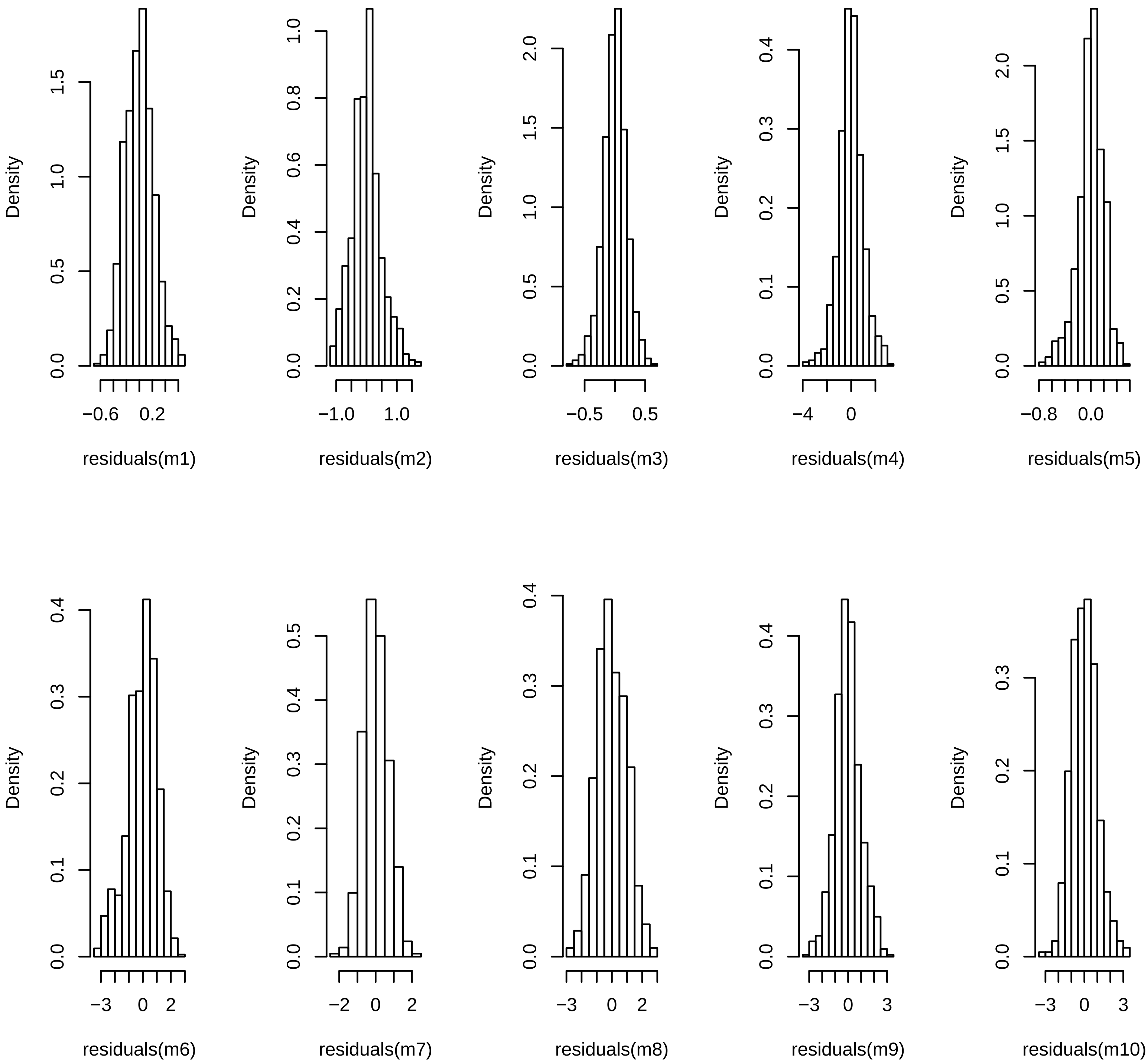
**Child Negative Behaviour**

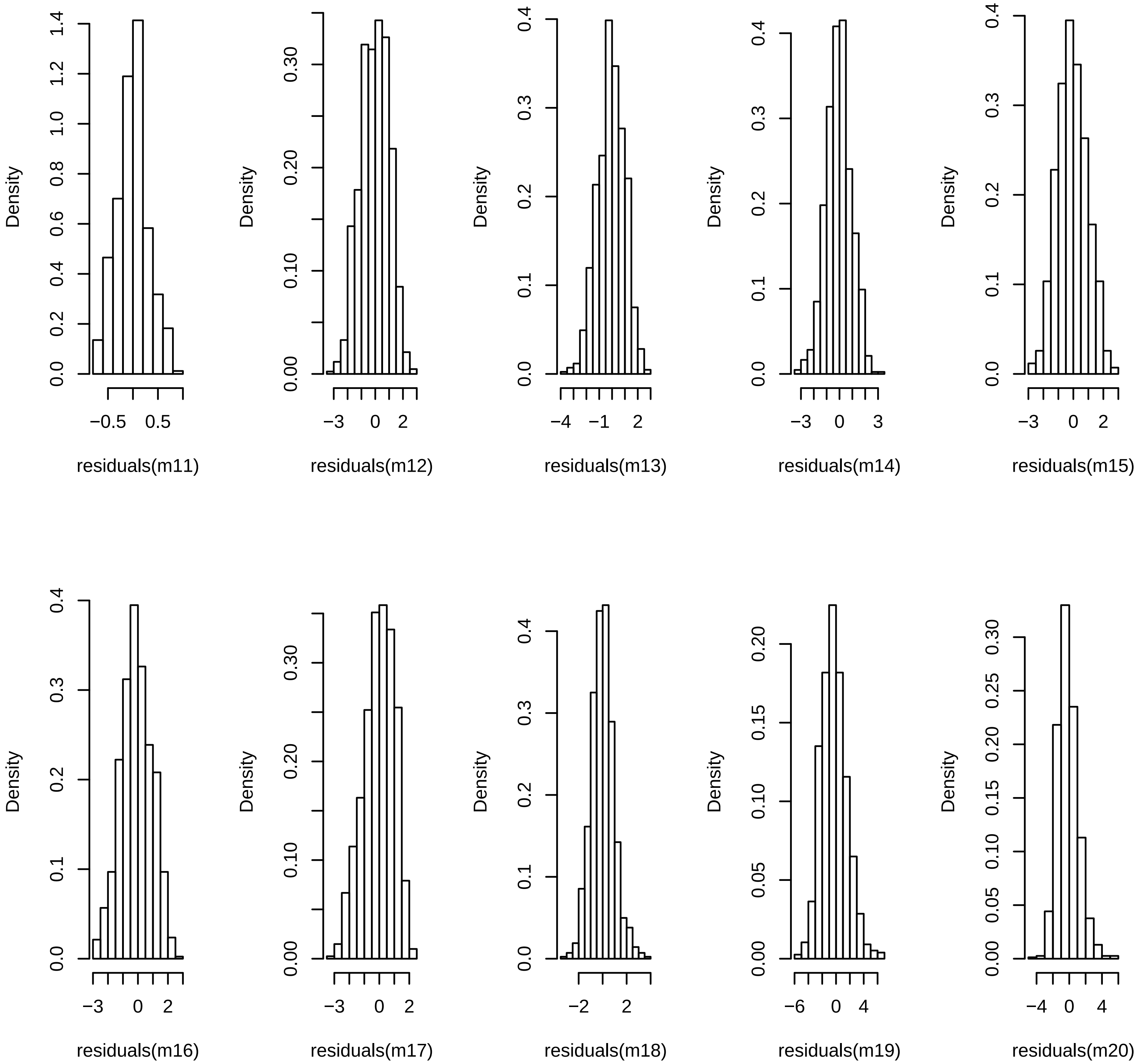


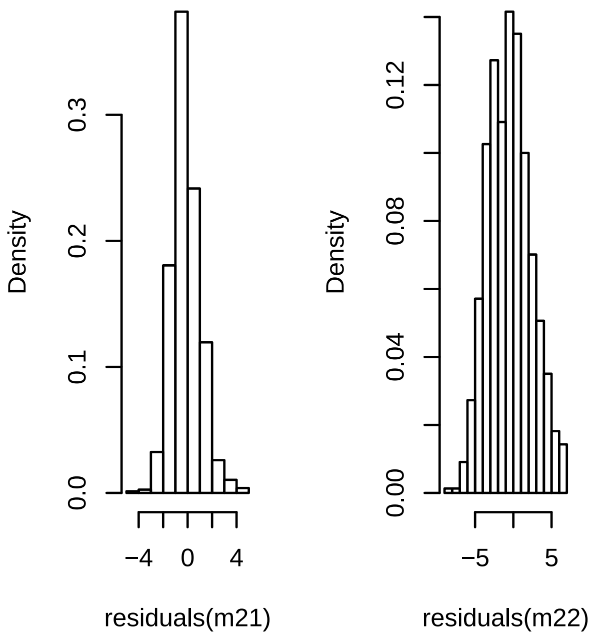
# Appendix D

# Model Diagnostics Plots

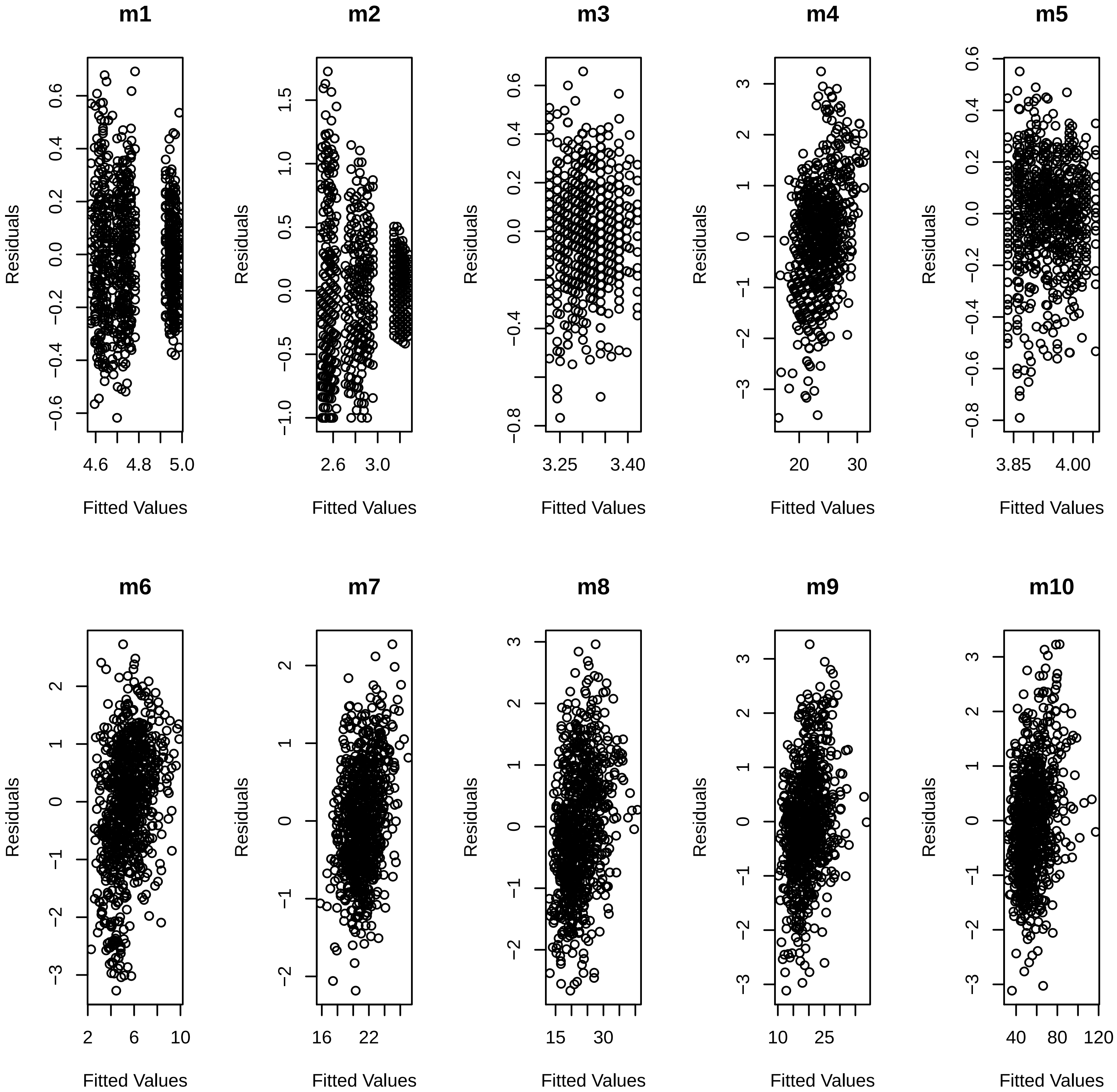
## D.1 Binary Response Models: ITT Models

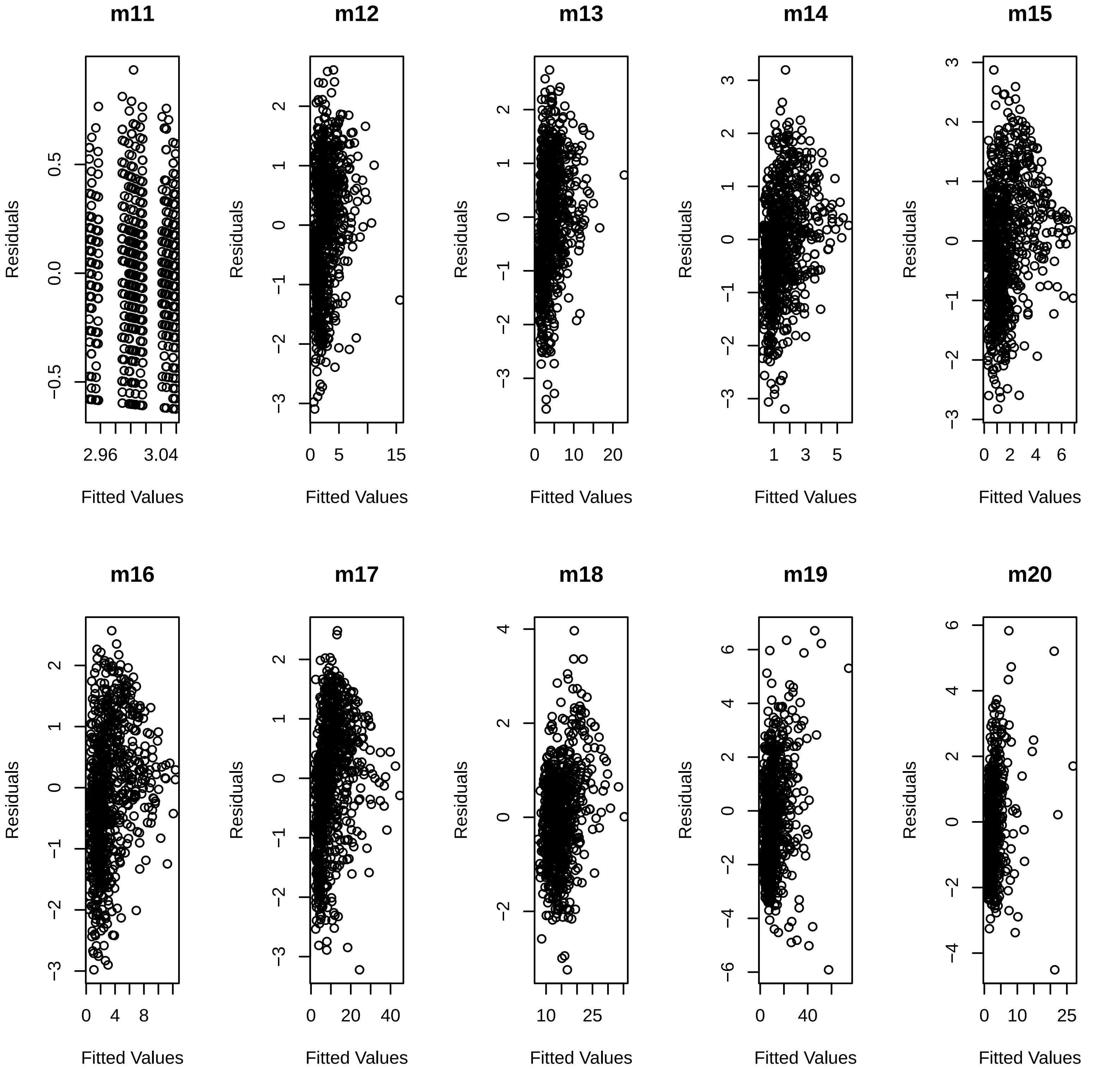

Density
residuals(m1)
residuals(m2)
residuals(m3)
residuals(m4)
residuals(m5)
residuals(m6)
residuals(m7)
residuals(m8)
residuals(m9)
residuals(m10)

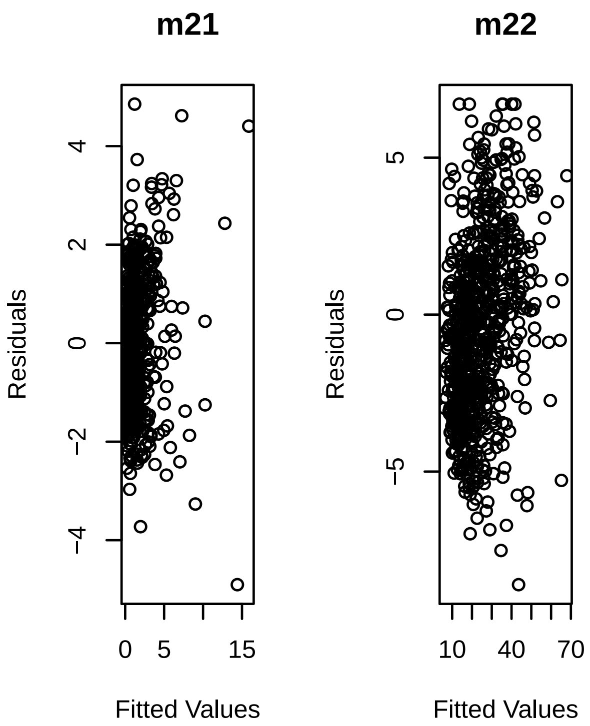

Density
0.0
0.2
0.4
0.6
0.8
1.0
1.2
1.4
−0.5
0.5
residuals(m11)
Density
0.00
0.10
0.20
0.30
−3
0
2
residuals(m12)
Density
0.0
0.1
0.2
0.3
0.4
−4
−1
2
residuals(m13)
Density
0.0
0.1
0.2
0.3
0.4
−3
0
3
residuals(m14)
Density
0.0
0.1
0.2
0.3
0.4
−3
0
2
residuals(m15)
Density
0.0
0.1
0.2
0.3
0.4
−3
0
2
residuals(m16)
Density
0.00
0.10
0.20
0.30
−3
0
2
residuals(m17)
Density
0.0
0.1
0.2
0.3
0.4
−2
2
residuals(m18)
Density
0.00
0.05
0.10
0.15
0.20
−6
0
4
residuals(m19)
Density
0.00
0.05
0.10
0.15
0.20
0.25
0.30
−4
0
4
residuals(m20)



Density
0.0
0.1
0.2
0.3
−4
0
4
residuals(m21)
Density
0.00
0.04
0.08
0.12
−5
5
residuals(m22)

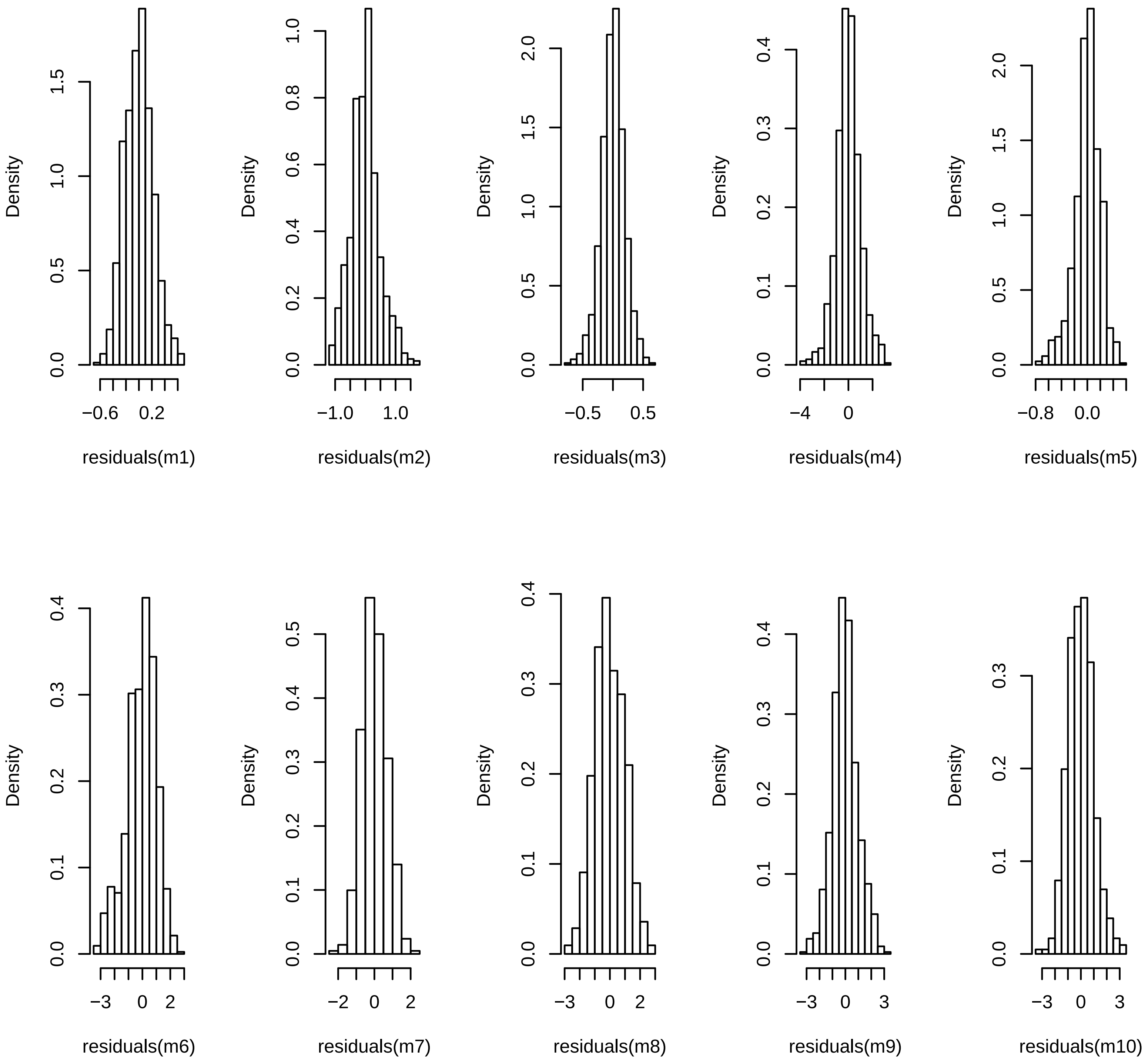

m1
m2
m3
m4
m5
m6
m7
m8
m9
m10
Residuals
Fitted Values


m11
m12
m13
m14
m15
m16
m17
m18
m19
m20
Residuals
Fitted Values

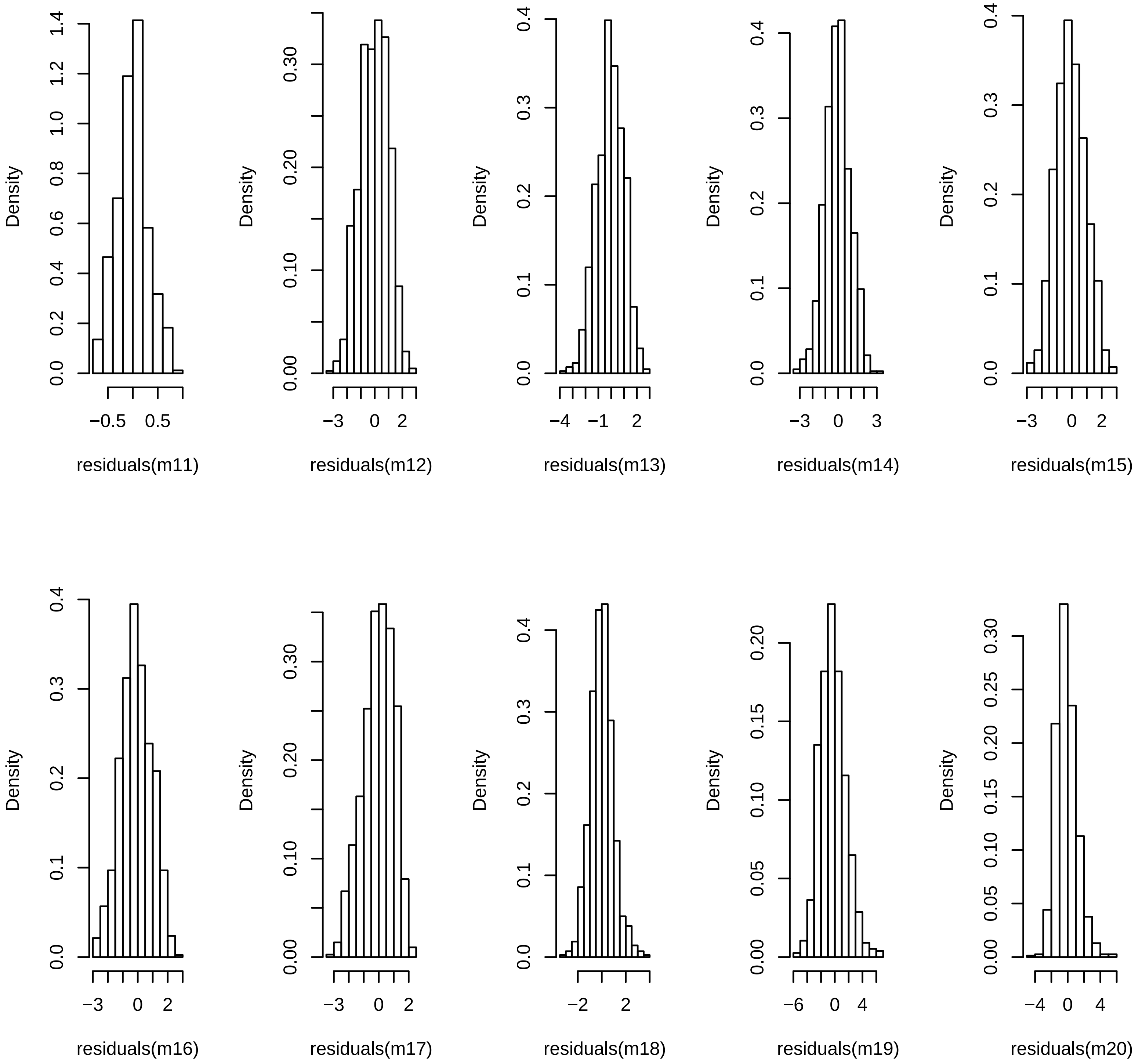


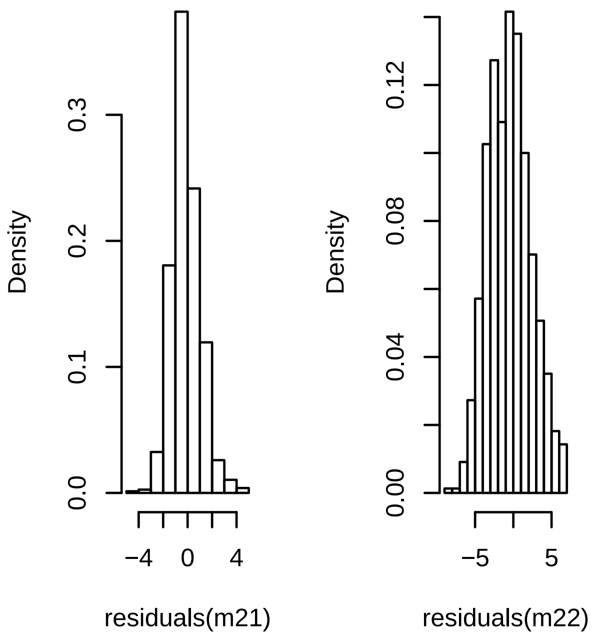
m21
Residuals
Fitted Values
m22
Residuals
Fitted Values

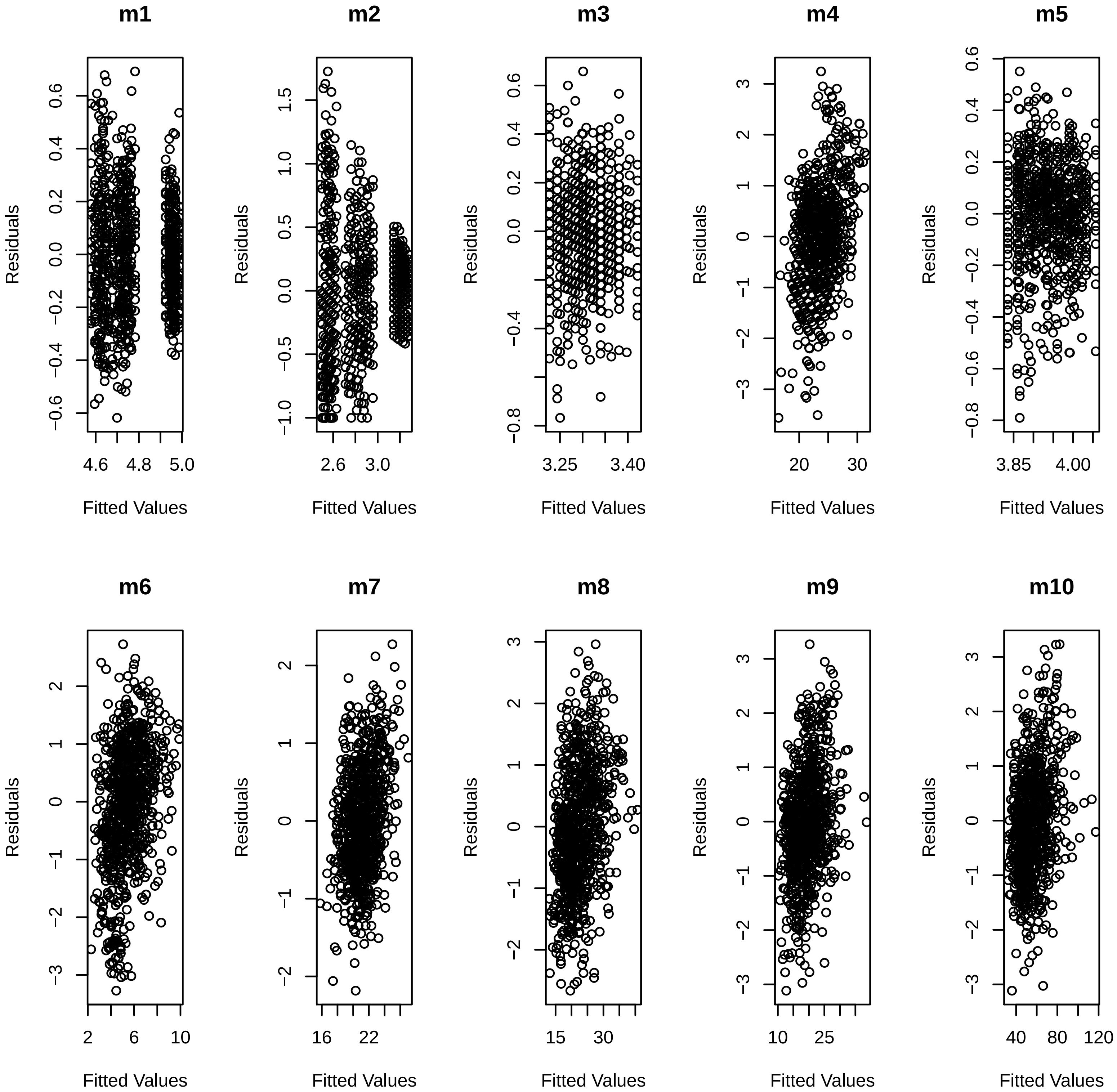
## D.2 Binary Response Models: PP Models

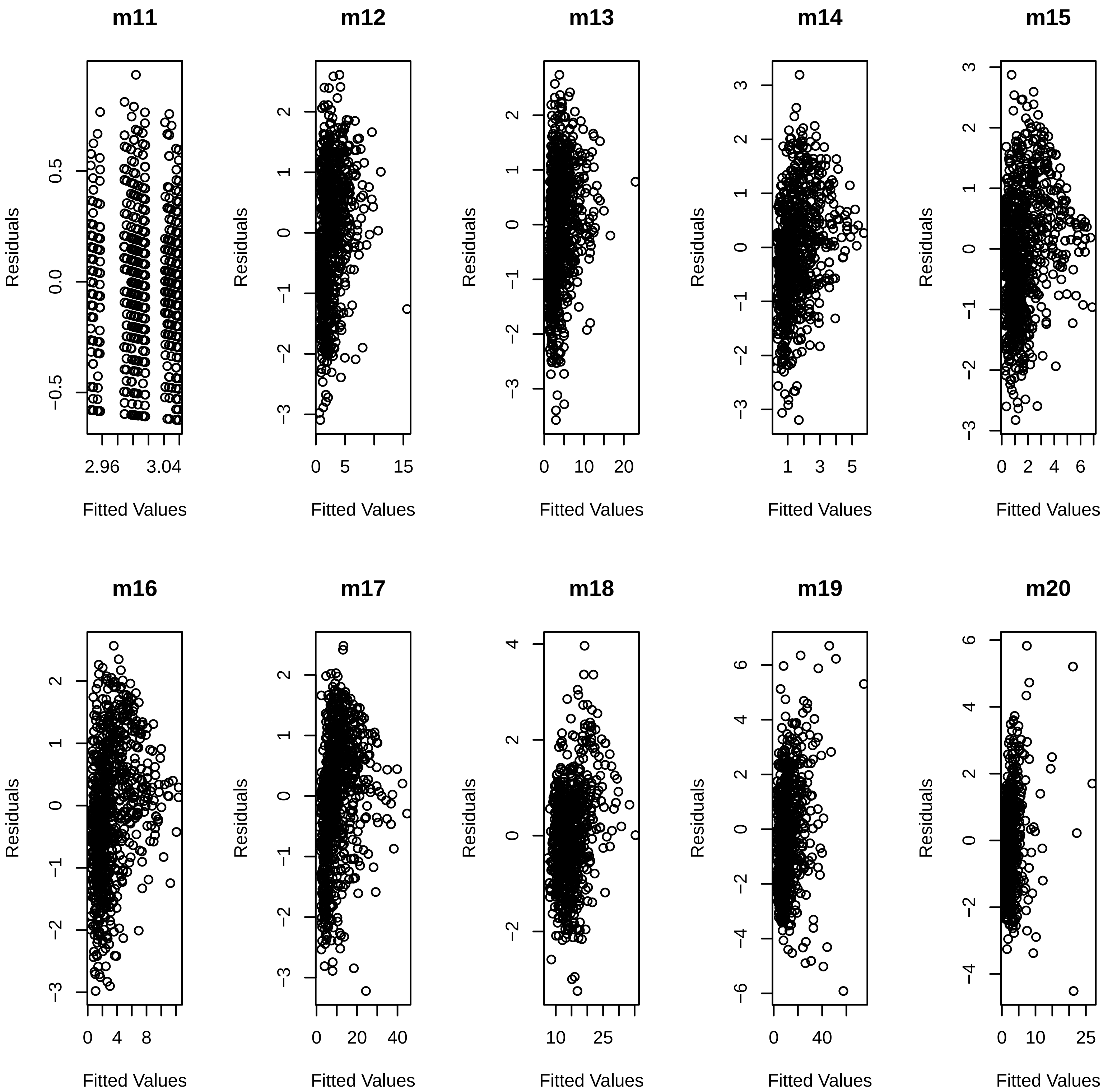


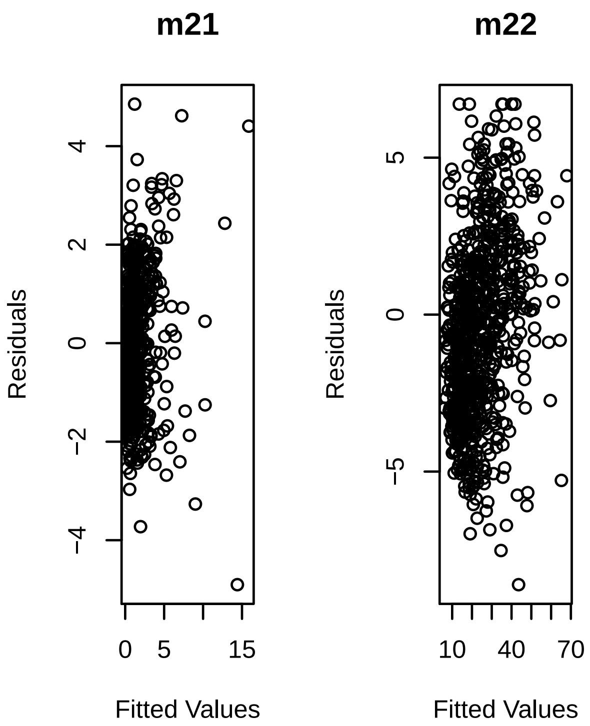
Density
residuals(m1)
residuals(m2)
residuals(m3)
residuals(m4)
residuals(m5)
residuals(m6)
residuals(m7)
residuals(m8)
residuals(m9)
residuals(m10)


Density
residuals(m11)
residuals(m12)
residuals(m13)
residuals(m14)
residuals(m15)
residuals(m16)
residuals(m17)
residuals(m18)
residuals(m19)
residuals(m20)


Density
0.0
0.1
0.2
0.3
−4
0
4
residuals(m21)
Density
0.00
0.04
0.08
0.12
−5
5
residuals(m22)


m1
m2
m3
m4
m5
m6
m7
m8
m9
m10
Residuals
Fitted Values


m11
m12
m13
m14
m15
m16
m17
m18
m19
m20
Residuals
Fitted Values
2.96
3.04
-0.5
0.0
0.5
0
5
15
10
20
1
3
2
4
6
8
25
40
-3
-2
-1
-4
-6


m21
m22
Residuals
Residuals
Fitted Values
Fitted Values
4
2
0
−2
−4
5
0
−5
0
5
15
10
40
70