

# Guide to Dataset: Colombia ECD Project Follow-Up II

---

## Contents

|  |   |
|--|---|
| Project context.....                             | 2 |
| The interventions.....                           | 2 |
| The dataset .....                                | 3 |
| Note on identifiers.....                         | 3 |
| Sources of data .....                            | 3 |
| Household Survey .....                           | 3 |
| Assessment of outcomes of the target child ..... | 7 |
| Assessments for mother .....                     | 7 |
| Assessments for younger siblings .....           | 7 |
| Community Survey.....                            | 7 |
| Additional sources of information .....          | 7 |

## **Project context**

This project covered 96 small towns in central Colombia, representing a large number of small communities across a relatively big geographical area. It exploited structures in place from the government's Conditional Cash Transfer Programme, Familias en Accion (FeA), which targets the poorest 20% of households in the country.

There are currently three waves of data, a baseline, pre-intervention wave collected between February and June 2010, and a follow-up wave 18 months later between September and December 2011, at the end of the intervention period. The second wave of follow-up data collection occurred 2 months after the first follow-up data collection between September and December 2013.

The beneficiaries of FeA periodically elect a female representative, called the Madre Lider (ML). We randomly selected three from each town (municipality), and then from the families represented by the ML we randomly selected 5 children aged 12 to 24 months to be eligible for the intervention. Within each municipality, eligible households were randomly allocated (at the municipality level) to each of the following treatment arms:

- 1. Control**
- 2. Stimulation + Supplementation**
- 3. Stimulation**
- 4. Supplementation**

These MLs were then asked to take on the role of home visitor to implement the intervention of the different treatment arms. However, a significant proportion (nearly 40%) of the MLs either opted out or were not deemed suitable for the role, in which case they were replaced following their recommendations. In order to maintain the randomisation of the design, the families attached to the original ML were maintained, even if in the end she was not the actual home visitor.

## **The interventions**

The stimulation intervention consisted of weekly visits to the homes of the target children, each visit lasting around one hour. The home visitors received a three week training programme in activities designed to stimulate children at different ages. They also received a weekly curriculum as a guide, and a set of locally produced materials (homemade toys from recycling material, picture books, puzzles, etc.).

The supplementation arm consisted of providing daily sachets of multiple micronutrient powder to mothers, via the home visitors, to add to the target child's food. Sachets were designed to provide iron (12.5mg), zinc (5mg), Vitamin A (300 µg retinol equivalent), Vitamin C (30mg) and folic acid (160 µg) for the children targeted.

## The dataset

This dataset forms the second follow-up dataset for the intervention discussed above. Data from baseline and the first follow-up as well as intervention data (including treatment status) are publically available through the UK data service: [10.5255/UKDA-SN-851196](https://ukdataservice.ac.uk/10.5255/UKDA-SN-851196).

### Note on identifiers

Tables 1 and 2 outline the key identifiers used to identify clusters, households and individuals in this second round of follow-up data collection. The main identifier for the cluster id 'dane', 'nim' is for the household. Both 'dane' and 'nim' is identical to the identifiers used in baseline and follow-up I. 'key\_2s' is the new individual IDs assigned to each individual present in the household at the time for follow-up survey. 'key\_2s' is different from the ID used at baseline and follow-up I ('llaveper') however 'key\_1s' which is included in all individual level datasets is equal to 'llaveper' and so should be used for linking individuals across waves. 'key\_1s' is not defined for household members who weren't recorded in previous rounds.

### Sources of data

There are four main sources of data:

- Household survey
- Child Developmental Assessments
- Intervention Data
- Community Survey

### Household Survey

This collected information on the household and on household members in a set of modules (questionnaires) administered by interviewers. The mapping from the questionnaires to the raw datasets is shown in tables 2 and 3. Note that some additional questions were asked in the follow-up II round while other questions were removed, so it is worth checking the wording of questions in the original questionnaires.

Separate questionnaires were asked at household and individual levels. At the individual level there are several standalone modules for the mother, children under six years old, the target child of the intervention, the main care giver, and the home visitor.

The modules contain a rich variety of data: the home environment and composition of the household; education and health; employment and time use etc. There are also a series of modules on child development.

**Table 1: Identifiers and sample sizes**

|                   | Identifier at Baseline | Number at Baseline | Identifier at Follow-up | Number at Follow-up | Identifier at Follow-up II | Number at Follow-up II |
|-------------------|------------------------|--------------------|-------------------------|---------------------|----------------------------|------------------------|
| <b>Cluster</b>    | dane                   | 96                 | dane                    | 96                  | dane                       | 96                     |
| <b>Households</b> | llave                  | 1,429              | nim                     | 1,331               | nim                        | 1312                   |
| <b>Target</b>     | llaveper               | 1,429              | llaveper                | 1,332               | key_1s                     | 1316                   |

|  |            |       |            |       |                        |       |
|--|------------|-------|------------|-------|------------------------|-------|
| <b>children</b>  |            |       |            |       | OR<br>key_2s           |       |
| <b>Biological Mothers</b>                                | llaveper   | 1,395 | llaveper   | 1,259 | key_1s<br>OR<br>key_2s | 1,208 |
| <b>All household members</b>                             | llaveper   | 7,499 | llaveper   | 7,698 | key_2s                 | 6457  |
| <b>Children more than 6 months and less than 6 years</b> | llaveper   | 2,147 | llaveper   | 1,829 | key_2s                 | 1,609 |
| <b>All children less than 6 years</b>                    | llaveper   | 2,210 | llaveper   | 1,889 | key_2s                 | 1742  |
| <b>All persons over age 5</b>                            | llaveper   | 5,288 | llaveper   | 4,665 | key_2s                 | 6069  |
| <b>Primary caregiver</b>                                 | llaveper   | 151   | llaveper   | 122   | key_2s                 | 196   |
| <b>Home Visitor</b>                                      | hv_0_id_bl | 340   | hv_0_id_fu | 326   | n/a                    | n/a   |

**Table 2: Mapping of datasets to questionnaires at Follow-up II**

| Mapping of questionnaires – Follow-up II |               |  |   |  |   |       |
|--|---------------|--|---|--|---|-------|
| Dataset                                  | Questionnaire | Questions  | Questionnaire Name                                  | Sections   | Unique identifier                                       | N     |
| m1_1                                     | modulo 1      | 101 – 325,<br>401 – 403,<br>701 – 710,<br>901 - 909                | Housing and Home                                    | Location and identification, Housing Characteristics, Health, Distances, Expenses, Other revenues and transfers, Incidents in the past 18 months,  | nim   | 1,313 |
| m1_2                                     | modulo1       | 1101-1202  | Housing and Home                                    | Participation in Familias en Accion  | nim   | 1313  |
| m1_3                                     | modulo1       | 1002 - 1003  | Housing and Home                                    | Food consumption of children in the household (for children between 6 months and 6 years)  | nim+orden   | 1,609 |
| m1_4                                     | modulo1       | 300-320  | Housing and Home                                    | Family Structure   | key_1s OR key_2s  | 8293  |
| m1_5                                     | modulo 1      | 325  | Housing and Home                                    | Children born in last two years not living   | nim+orden2  | 4     |
| m2_1                                     | modulo 2      | 101 - 315  | Education, work and health (for persons over age 5) | Education, Health, Work  | key_1s OR key_2s  | 6069  |
| m3_1                                     | modulo 3      | 101-216  | Children less than 6 years                          | Institutional and non-institutional care   | key_1s OR key_2s  | 1742  |
| m3_2                                     | modulo 3      | 216 - 218  | Children less than 6 years                          | Historical care arrangements   | (key_1s OR key_2s)+ ord_216                             | 3,153 |
| m3a_1                                    | modulo 3A     | 101 – 209,<br>311 – 313,<br>401 – 417,<br>501 – 508,<br>701<br>601 | Annex for module 3 – Target Child                   | Interaction with biological parents (if don't live at home), Health, Morbidity, Quality of home environment, Child behavior questionnaire (CBQ), Strengths and difficulties questionnaires (SDQ) | key_1s OR key_2s  | 1316  |
| m3a_2                                    | modulo 3A     | 301  | Annex for module 3 – Target Child                   | Time use   | (key_1s OR key_2s)+ time interval (m3_301a1 – m3_301a4) | 3399  |
| m3a_3                                    | modulo 3A     | 302 – 310  | Annex for module 3 –                                | Time use   | key_2s +orden3  | 1760  |

|      |           |                       | Target Child                            |   |                  |       |
|------|-----------|-----------------------|---|---|------------------|-------|
| m3b1 | modulo 3B | 102<br>401-508        | Annex for module 3 –<br>Younger sibling | Quality of home environment, Health, Mobility,  | key_2s           | 292   |
| m3b2 | modulo 3B | 601                   | Annex for module 3 –<br>Younger sibling | MacArthur-Bates (8-18 months)   | key_2s           | 292   |
| m3b3 | modulo 3B | 601                   | Annex for module 3 –<br>Younger sibling | MacArthur-Bates (19-30 months)  | key_2s           | 292   |
| m3b4 | modulo 3B | 603-610               | Annex for module 3 –<br>Younger sibling | MacArthur-Bates (31-48 months), Use of language   | key_2s           | 292   |
| m4_1 | modulo 4  | 101 – 305,<br>801     | Biological Mother                       | Time use, Health, Empowerment, Contraception and<br>reproductive history, Knowledge of child<br>development | key_1s OR key_2s | 1,208 |
| m4_2 | modulo 4  | 401 – 406,<br>502-504 | Biological Mother                       | Birth and breastfeeding of biological children,<br>Physical development of biological children              | key_1s OR key_2s | 302   |
| m5_1 | modulo 5  | 100 - 601             | Primary Caregiver                       | ID and location, Education, Health, Work, Time use,<br>Knowledge of child development,                      | key_1s OR key_2s | 196   |
| m6_1 | modulo 6  | 101                   | Home Inventory                          | Home Inventory to assess the quality of the family<br>environment   | nim              | 1312  |

### **Assessment of outcomes of the target child**

In addition to the information collected in the questionnaires administered by interviewers at the household level, the target children also attended a centre where her developmental level was assessed by a trained psychologist using a variety of assessments:

- Test de Vocabulario en Imagenes Peabody (TVIP) to assess receptive language
- Daberon (modified) to assess school readiness
- Pencil Tapping Task to assess executive functioning
- Wolke to assess Emotional tone, activity, co-cooperativeness, approach
- Woodcock-Munoz Achievement 14 to assess expressive language
- Woodcock-Munoz Cognitive 5 to assess fluid reasoning/fluid intelligence
- Woodcock-Munoz Cognitive 6 to assess visual perception/processing speed
- Woodcock-Munoz Cognitive 12 to assess fluid reasoning
- Woodcock-Munoz Cognitive 13 to assess visual perception/spatial memory
- Woodcock-Munoz Cognitive 16 to assess speed processing
- Woodcock-Munoz Cognitive 21 to assess long term memory

In the dataset tests\_TC we include the raw scores for each child assessment. These can be linked to other datasets using key\_2s.

More details of the scoring and administration of each test is available in the document “guide to scoring and administration of child assessment data”

### **Assessments for mother**

In the test centre we also performed some assessments with the mother of the target child:

- Raven’s progressive matrices to measure IQ
- CES-D 10 point depression scale
- Big Five to measure personality traits

In the dataset tests\_mother we include the raw scores for each of these assessments. These can be linked to other datasets using key\_2s.

### **Assessments for younger siblings**

We used the ASQ-3 to assess child developmental levels of the younger siblings. This assessment occurred in the household. Item level data is available in the file asq\_sibs.

### **Community Survey**

This collects the information at the municipality level included in the questionnaire “community”. This includes information on shocks, services, wages, prices. The data are included in the datasets com\_1-com\_5 and the key identifying variable for merging to other datasets is “comunidad” which is identical to “dane” in the other datasets.

### **Additional sources of information**

For more detailed information on the intervention and the data, please refer to ongoing work which will be updated on Evaluation of Development Policy department’s website, at the Institute for Fiscal Studies: <http://www.ifs.org.uk/centres/EDePo/about>

If there are any further questions regarding use of the data, please contact the researchers involved in this project at the IFS:

- Emla Fitzsimons: [emla\\_f@ifs.org.uk](mailto:emla_f@ifs.org.uk)
- Marta Rubio-Codina: [marta\\_r@ifs.org.uk](mailto:marta_r@ifs.org.uk)
- Alison Andrew: [alison\\_a@ifs.org.uk](mailto:alison_a@ifs.org.uk)
- Orazio Attanasio: [o.attanasio@ucl.ac.uk](mailto:o.attanasio@ucl.ac.uk)