Child Development Project Report: Predictive Analysis of Early Childhood Development Deficits

NGO Applied Project - BREIT

Angel Choquehuanca Luis Torpoco Alan Fraquita Edson

2025-10-10

Table of contents

Executive Summary

This report presents a comprehensive predictive analysis of early childhood development deficits for Asociación Taller de los Niños (TANI), a Peruvian NGO with over 45 years of experience serving vulnerable children and families. Through data science methodologies, we analyze pre- and post-pandemic health records to develop actionable predictive models that identify children at highest risk of developmental delays.

Methodology Overview:

- Multi-phase data consolidation from pre/post-pandemic periods
- Strategic population filtering to address class imbalance
- Feature engineering

Key Findings:

- [To be completed after final analysis]
- Identification of critical risk factors for targeted intervention
- Evidence-based recommendations for resource allocation

Impact: This work provides TANI with a decision-support framework to optimize early intervention strategies, improving outcomes for vulnerable children through data-driven prioritization.

---

1. Introduction and Context

1.1. Background on TANI

Asociación Taller de los Niños (TANI) is a non-governmental organization based in Peru, dedicated to improving the quality of life for children and families in vulnerable situations. With over 45 years of operational experience, TANI focuses particularly on the critical early years of life (0-5 years), implementing a comprehensive approach that integrates:

- **Health and Nutrition:** Growth monitoring, nutritional assessment, and intervention programs
- **Community Strengthening:** Family engagement and capacity building
- **Preventive Care:** Systematic health check-ups and early detection of developmental issues

1.2. Child Development Deficit Context

Fundamental Concepts of Growth and Development:

Child growth and development constitute two interrelated sets of indicators of critical utility for determining health status. Growth refers to the physical increase in weight and height/length measurements while development encompasses the progressive improvement of functional capacity and skills across multiple domains (motor, cognitive, language, and socio-emotional). Both processes are fundamentally dependent on the interaction of genetic, nutritional, and environmental factors. Therefore, systematic monitoring of normal developmental trajectories is essential to identify risk factors that may interfere with these processes, ensuring that each child achieves their full developmental potential.

The Global Burden of Early Childhood Development Deficits:

Early childhood development (ECD) deficits represent a critical public health challenge in low- and middle-income countries. According to WHO estimates, approximately 250 million children under five years in these settings are at risk of not reaching their developmental potential due to poverty and stunting (Black et al., 2017). These deficits manifest across multiple, often overlapping dimensions:

- **Stunting (Chronic Malnutrition):** Linear growth failure resulting from prolonged nutritional deficiencies, affecting approximately 22% of children globally (UNICEF, 2021)
- **Wasting (Acute Malnutrition):** Low weight-for-height indicating recent or current severe nutritional deficit
- **Developmental Delays:** Failure to achieve age-appropriate milestones in gross motor, fine motor, cognitive, language, or social-emotional domains
- **Micronutrient Deficiencies:** Particularly iron-deficiency anemia, affecting cognitive development and immune function

The Peruvian Context:

In Peru, despite significant economic growth over recent decades, substantial disparities persist in child health outcomes:

- **Chronic Malnutrition:** Affects 12.1% of children under 5 nationally, with rates exceeding 25% in rural Andean regions (INEI-ENDES, 2023)
- **Anemia Prevalence:** Reaches 43.1% in children aged 6-35 months, a critical window for brain development (INEI-ENDES, 2023)
- **Developmental Delays:** Studies indicate that children exhibit delays in one or more developmental domains, with higher rates among socioeconomically disadvantaged populations (Diaz et al., 2017)
- **Geographic Inequities:** Urban-rural and coastal-highland disparities create vastly different developmental trajectories for children

The Critical Window: First 1,000 Days:

The period from conception to age 3 (approximately 1,000 days) represents a critical window for brain architecture development. During this time, neural connections form at a rate exceeding 1 million per second, establishing the foundational circuitry for all future learning, behavior, and health (Shonkoff & Phillips, 2000). Nutritional and environmental insults during this window can result in:

- **Irreversible Stunting:** Chronic undernutrition affecting linear growth and cognitive capacity
- **Neurodevelopmental Impairment:** Compromised executive function, memory, and academic readiness
- **Immunological Vulnerability:** Increased susceptibility to infectious diseases
- **Intergenerational Transmission:** Girls who experience stunting are more likely to have low-birth-weight infants, perpetuating cycles of disadvantage

Multidimensional Determinants:

Following Bronfenbrenner's ecological systems framework (1979), child development results from interactions across multiple levels:

- **Individual Factors:** Genetic endowment, birth characteristics (gestational age, birth weight), sex
- **Family/Microsystem:** Maternal nutrition and health, breastfeeding practices, caregiver-child interactions, household food security
- **Community/Exosystem:** Access to healthcare services, availability of nutritious foods, water and sanitation infrastructure
- **Societal/Macrosystem:** Economic policies, health system organization, cultural beliefs about child-rearing

This multidimensional causality necessitates comprehensive assessment approaches that capture not only anthropometric measurements but also functional developmental capacity and contextual risk factors.

1.3. Project Objectives

This project aims to leverage TANI's longitudinal data to develop actionable predictive models. Specific objectives include:

Primary Objective

Develop and validate machine learning models to predict early childhood development deficits, enabling proactive identification of at-risk children for targeted interventions.

Secondary Objectives

1. **Exploratory Analysis:** Conduct comprehensive exploratory data analysis (EDA) to identify patterns, risk factors, and data quality issues
2. **Feature Engineering:** Transform raw clinical measurements into actionable predictive features
3. **Model Development:** Compare multiple predictive algorithms and select the optimal approach for deployment
4. **Actionable Insights:** Provide evidence-based recommendations for TANI's operational and strategic decisions
5. **Data Infrastructure:** Document data quality issues and recommend improvements for future data collection

1.4. Scope and Limitations

Scope

- **Temporal Coverage:** Analysis covers post pandemic observations.
- **Population:** Children aged 0-5 years receiving services at TANI health centers
- **Outcome Variables:** Nutritional status (P/T, T/E, P/E) and developmental domains (gross motor, fine motor, cognitive, language, social)
- **Predictive Modeling:** Focus on classification models for binary outcomes (deficit vs. normal)

Limitations

- **Data Quality:** As with most real-world clinical data, missing values and measurement inconsistencies are present
- **Selection Bias:** Population consists of families actively seeking TANI services, which may not represent the broader vulnerable population
- **Causality:** While we identify predictive associations, establishing causal relationships requires additional study designs
- **External Validity:** Models will be optimized for TANI's specific population and may require recalibration for other contexts

1.5. Key Stakeholders

Primary Beneficiaries

- **Children (0-5 years):** Direct beneficiaries through improved early detection and intervention

- **Families:** Receive targeted support and guidance based on predictive risk assessments
- **TANI Clinical Staff (nurses):** Gain decision support tools for prioritizing cases and planning interventions

Secondary Stakeholders

- **TANI Leadership:** Use insights for strategic planning, resource allocation, and fundraising
- **Public Health Authorities:** Evidence base for policy recommendations and scaling interventions
- **Research Community:** Methodological contributions to applied machine learning in global health

---

2. Data Reception and Consolidation

2.1. Data Sources

The analysis utilizes three primary data sources provided by TANI:

Internal Data

1. **MALNUTRITION Sheet:** Longitudinal records of children's growth monitoring and nutritional assessments
2. **DEVELOPMENT Sheet:** Developmental screening results across five domains

External Reference Standards

- **WHO Growth Standards:** Age and sex-specific percentile calculations for anthropometric indicators
- **Developmental Milestones:** Peru Ministry of Health (MINSA) guidelines for age-appropriate development

2.2. Data Challenges and Quality Issues

Initial data exploration revealed several challenges common to clinical registries:

Structural Issues

- **Inconsistent Formatting:** Age represented in mixed formats (days, months, years: "4d", "6m", "1a8m")
- **Variable Types:** Anthropometric measurements stored as text instead of numeric values
- **Categorical Encoding:** Multiple encoding schemes for the same concept (e.g., "SI"/"Normal" for development)

Data Quality Issues

- **Missing Values:** Systematic patterns of missingness in periods closed to pandemic year.
- **Measurement Errors:** Outliers in weight, height, and head circumference requiring validation
- **Temporal Inconsistencies:** Some records show impossible sequences (e.g., height decreasing over time)
- **Duplicate Records:** Need to establish deduplication rules for repeated visits

Documentation Gaps

- **Sparse Metadata:** Limited information on data collection protocols and quality control procedures
- **Variable Definitions:** Some variables lack clear operational definitions in the data dictionary
- **Missing Codes:** Inconsistent handling of "not applicable" vs. "not measured" vs. "refused"

2.3. Consolidation Process

To address these challenges, we implemented a systematic data consolidation pipeline:

Phase 1: Data Profiling

```
============================================================
DATASET: DESNUTRICION
============================================================

Shape: 257,178 rows × 27 columns
```

```
Data Types Distribution:
object            23
float64            3
datetime64[ns]     1
Name: count, dtype: int64

Top 10 Variables with Missing Values:
                        Missing_Count  Missing_Percentage
Tam_graha                      253123               98.42
Tam_para                       252117               98.03
Recuperado                     247952               96.41
Tam_hb                         243562               94.71
Razón                          238809               92.86
Mantiene_Diag_Fav/Desf         188524               73.30
Lactancia                      186632               72.57
CN-CA                          167056               64.96
ACA                            112621               43.79
P/E                             47695               18.55

============================================================
DATASET: DESARROLLO
============================================================

Shape: 257,178 rows × 18 columns

Data Types Distribution:
object            15
float64            2
datetime64[ns]     1
Name: count, dtype: int64

Top 10 Variables with Missing Values:
            Missing_Count  Missing_Percentage
(M) - FF           130726               50.83
CabPC               42520               16.53
(S) - Soc           12915                5.02
(L) - Len           12902                5.02
(C) - Cog           12905                5.02
(M) - FG            12902                5.02
Talla                  18                0.01
Nº_HC                   9                0.00
Peso                    7                0.00
```

```
Nº_Control              1                0.00
```

Phase 2: Data Cleaning Functions

Functions for parsing variables like Age, Development area, and others were implemented.

Phase 3: Data Integration

```
Data Cleaning Results:
Original records: 257,178
After cleaning and consolidation (pre and post pandemic datasets): 454,901

Age conversion success rate: 100.0%
Weight conversion success rate: 100.0%
Height conversion success rate: 100.0%
```

2.4. Consolidated Dataset Characteristics

```
Consolidated Dataset Summary Statistics:

Anthropometric Measurements:
          edad_meses       Peso_Num      Talla_Num         PC_Num
count  454901.000000  454869.000000  454831.000000  275650.000000
mean       13.996886       9.786695      73.360135      43.174217
std        12.006950       3.365823      12.680113       3.863954
min         0.000000       2.760000      46.700000      32.900000
25%         5.000000       7.570000      64.300000      40.300000
50%        10.000000       9.400000      71.500000      43.200000
75%        20.000000      11.650000      81.900000      46.200000
max        65.000000      24.100000     110.300000      52.500000

Nutritional Status Distribution:
```

3. Exploratory Data Analysis (EDA)

3.1. Dataset Overview

```
================================================================================
SECTION 3: EXPLORATORY DATA ANALYSIS (EDA)
================================================================================

Dataset: 454,901 records × 44 columns
Date range: 2009-01-10 to 2025-08-12
Unique children: 32,011

================================================================================
3.1. GENERAL DATASET DESCRIPTION
================================================================================

Number of records: 454,901
Number of variables: 44
Unique patients: 32,011
Average controls per patient: 14.2

--- Variable Types ---

Numerical variables (continuous):
  Count: 26
  Variables: N_HC, Peso, Talla, CabPC, Tam_hb, edad_meses, flg_cognitivo, flg_lenguaje, flg_m

Categorical variables:
  Count: 16
  Variables: Sexo, Diag_Nacimiento, Ganancia_Peso_Talla, Dx_Nutricional, CN-CA, Mantiene_Diag

Binary flags (0/1):
  Count: 7
  Variables: flg_cognitivo, flg_lenguaje, flg_motora_fina, flg_motora_gruesa, flg_social, flg

--- Temporal Coverage ---
First record: 2009-01-10
Last record: 2025-08-12
Time span: 6058 days

Pandemic periods:
  Pre-Pandemic: 304,516 (66.9%)
```

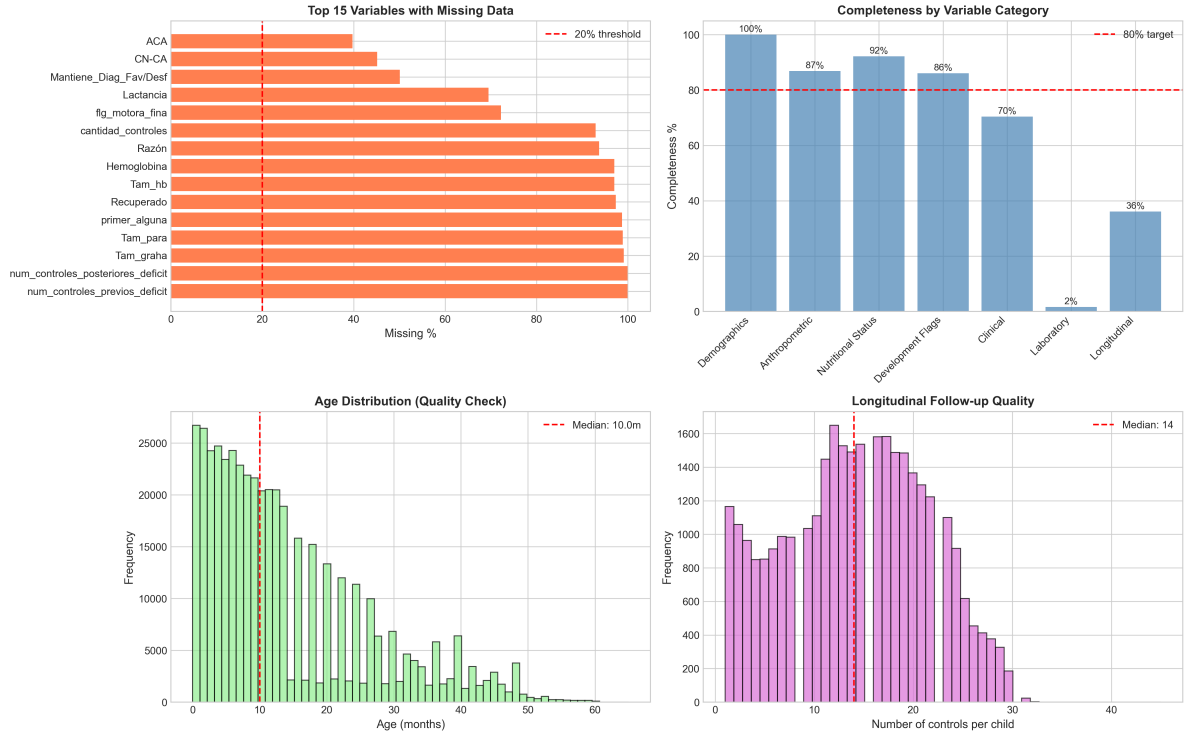
```
  Post-Pandemic: 150,385 (33.1%)

--- Age Distribution ---
count    454901.000
mean         13.997
std          12.007
min           0.000
25%           5.000
50%          10.000
75%          20.000
max          65.000
Name: edad_meses, dtype: float64

Age groups:
  0-6m: 149,746 (32.9%)
  6-12m: 127,788 (28.1%)
  12-24m: 97,104 (21.3%)
  24-36m: 48,427 (10.6%)
  36-60m: 31,814 (7.0%)
```

3.2. Data Quality Assessment

```
================================================================================
3.2. DATA QUALITY ASSESSMENT
================================================================================

--- Completeness Analysis ---

Overall completeness: 68.58%
Total cells: 20,015,644
Filled cells: 13,726,146
Missing cells: 6,289,498

--- Variables with Missing Data (>5%) ---
                         Variable  Missing_Count  Missing_Pct
    num_controles_previos_deficit         454601        99.93
num_controles_posteriores_deficit         454601        99.93
                      Tam_graha         450846        99.11
                       Tam_para         449840        98.89
                  primer_alguna         449064        98.72
                     Recuperado         442898        97.36
```

```
                          Tam_hb         441287        97.01
                     Hemoglobina         441287        97.01
                           Razón         426283        93.71
                cantidad_controles         422890        92.96
                   flg_motora_fina         328466        72.21
                         Lactancia         316182        69.51
           Mantiene_Diag_Fav/Desf         227942        50.11
                         CN-CA          205398        45.15
                           ACA          180885        39.76

--- Completeness by Variable Category ---
  Demographics             : 100.0%
  Anthropometric           :  86.9%
  Nutritional Status       :  92.1%
  Development Flags        :  86.0%
  Clinical                 :  70.4%
  Laboratory               :   1.7%
  Longitudinal             :  36.1%

--- Consistency Validation ---

1. Age range validation:
   Min age: 0.0 months
   Max age: 65.0 months
   Mean age: 14.0 months

2. Sex distribution:
   M: 231,266 (50.8%)
   F: 223,635 (49.2%)

3. Biologically implausible values:
   Weight <1.5kg or >30kg: 0 (0.000%)
   Height <40cm or >130cm: 0 (0.000%)

--- Accuracy Assessment ---

Anthropometric measurements by age group:
              Peso          Talla          CabPC
              mean    std    mean    std    mean    std
Grupo_Edad
0-6m         6.585  1.642  60.447  5.096  40.241  2.538
6-12m        9.283  1.223  70.694  3.346  44.755  1.566
12-24m      11.210  1.577  80.182  4.375  47.095  1.540
```

```
24-36m      13.755  1.879  89.945  4.378  48.672  1.456
36-60m      16.486  2.437  98.768  4.866  49.419  1.675
```



3.3. Descriptive Statistics

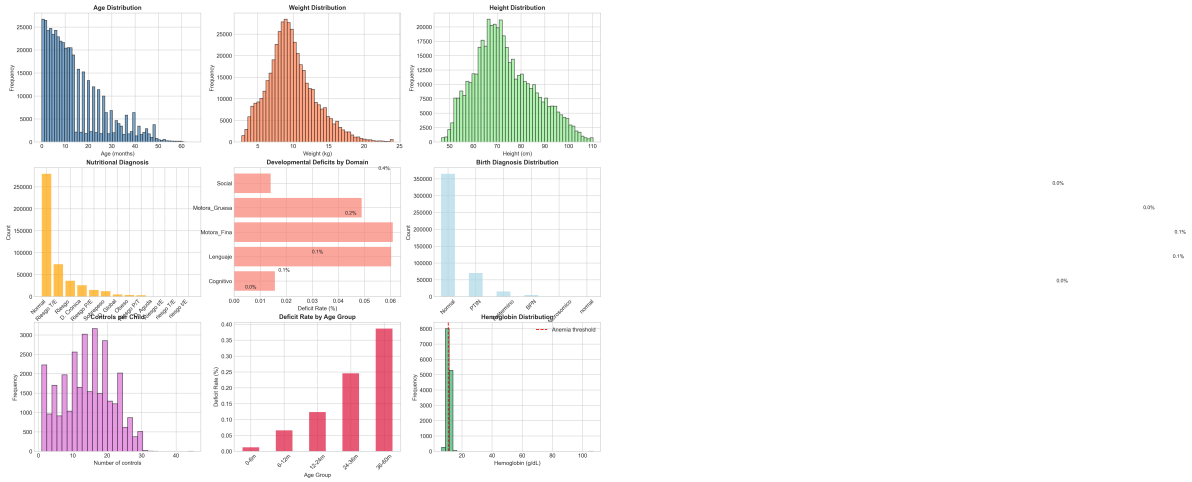
```
================================================================================
3.3. DESCRIPTIVE STATISTICS
================================================================================

--- Age Distribution ---
count    454901.000
mean         13.997
std          12.007
min           0.000
25%           5.000
50%          10.000
75%          20.000
max          65.000
Name: edad_meses, dtype: float64
```

```
Age by group:
               count    mean    std    min   25%   50%   75%   max
Grupo_Edad
0-6m        149746.0   3.403  1.751   0.00   2.0   3.0   5.0   6.0
6-12m       127788.0   9.430  1.717   6.03   8.0   9.0  11.0  12.0
12-24m       97104.0  18.448  3.377  12.03  16.0  18.0  22.0  24.0
24-36m       48427.0  30.412  3.396  24.03  28.0  30.0  33.0  36.0
36-60m       31814.0  43.599  4.724  36.13  40.0  43.0  47.0  60.0

--- Anthropometric Measurements ---

Overall statistics:
              Peso      Talla       CabPC
count  454869.000  454831.00  275650.000
mean        9.787      73.36      43.174
std         3.366      12.68       3.864
min         2.760      46.70      32.900
25%         7.570      64.30      40.300
50%         9.400      71.50      43.200
75%        11.650      81.90      46.200
max        24.100     110.30      52.500

--- Weight by Age Group ---
               count    mean    std   min     25%    50%    75%   max
Grupo_Edad
0-6m        149737.0   6.585  1.642  2.76   5.370   6.70   7.77  24.1
6-12m       127784.0   9.283  1.223  2.76   8.455   9.20  10.00  24.1
12-24m       97095.0  11.210  1.577  2.76  10.140  11.07  12.10  24.1
24-36m       48425.0  13.755  1.879  2.76  12.500  13.54  14.80  24.1
36-60m       31806.0  16.486  2.437  2.76  14.800  16.20  17.80  24.1

--- Height by Age Group ---
               count    mean    std   min   25%   50%    75%    max
Grupo_Edad
0-6m        149727.0  60.447  5.096  46.7  56.5  61.0   64.4  110.3
6-12m       127771.0  70.694  3.346  46.7  68.5  70.6   73.0  110.3
12-24m       97083.0  80.182  4.375  46.7  77.0  80.0   83.2  110.3
24-36m       48422.0  89.945  4.378  46.7  87.0  90.0   92.8  110.3
36-60m       31806.0  98.768  4.866  46.7  95.6  98.6  102.0  110.3

--- Nutritional Status Distribution ---
```

```
T/E_cat:
  N: 374,154 (91.9%)
  R: 26,846 (6.6%)
  DA: 2,693 (0.7%)
  DC: 1,764 (0.4%)
  O: 1,744 (0.4%)
  S: 8 (0.0%)
  DG: 5 (0.0%)
  NN: 1 (0.0%)
  NB: 1 (0.0%)

P/E_cat:
  N: 384,307 (94.4%)
  S: 17,051 (4.2%)
  R: 5,637 (1.4%)
  DG: 182 (0.0%)
  DC: 17 (0.0%)
  O: 13 (0.0%)
  DA: 3 (0.0%)

P/T_cat:
  N: 351,650 (86.4%)
  S: 45,182 (11.1%)
  O: 9,261 (2.3%)
  R: 929 (0.2%)
  DA: 172 (0.0%)
  DG: 9 (0.0%)
  DC: 6 (0.0%)
  SA: 1 (0.0%)

Dx_Nutricional:
  Normal: 279,566 (61.5%)
  Riesgo T/E: 73,633 (16.2%)
  Riesgo: 36,228 (8.0%)
  D. Crónica: 25,969 (5.7%)
  Riesgo P/E: 15,038 (3.3%)
  Sobrepeso: 12,332 (2.7%)
  D. Global: 5,055 (1.1%)
  Obeso: 3,684 (0.8%)
  Riesgo P/T: 3,076 (0.7%)
  D. Aguda: 310 (0.1%)
  Riesgo t/E: 1 (0.0%)
  riesgo T/E: 1 (0.0%)
```

```
  riesgo t/E: 1 (0.0%)

--- Developmental Status Distribution ---

Deficit rates by domain:
  flg_cognitivo            :   0.0% (69/441,617)
  flg_lenguaje             :   0.1% (266/441,620)
  flg_motora_fina          :   0.1% (77/126,435)
  flg_motora_gruesa        :   0.0% (216/441,618)
  flg_social               :   0.0% (62/441,602)
  flg_alguna               :   0.1% (465/454,901)

--- Clinical Characteristics ---

Birth diagnosis:
  Normal: 364,710 (80.2%)
  PTIN: 70,226 (15.4%)
  Prétermino: 15,259 (3.4%)
  BPN: 4,676 (1.0%)
  Macrosomico: 28 (0.0%)
  normal: 2 (0.0%)

Breastfeeding type (0-6 months):
  LME: 112,607 (81.3%)
  LMX: 24,240 (17.5%)
  LA: 1,041 (0.8%)
  LM+AB: 618 (0.4%)
  Lme: 7 (0.0%)
  lmx: 2 (0.0%)
  lmr: 1 (0.0%)
  lme: 1 (0.0%)

--- Laboratory Results ---

Hemoglobin (g/dL):
count    13614.000
mean        11.412
std          1.296
min          6.600
25%         10.800
50%         11.400
75%         12.000
max        107.000
```

```
Name: Tam_hb, dtype: float64
Anemia rate (<11.0 g/dL): 29.4% (4,003)

--- Longitudinal Follow-up Metrics ---

cantidad_controles:
count    32011.000
mean        14.211
std          7.278
min          1.000
25%          9.000
50%         14.000
75%         20.000
max         45.000
Name: cantidad_controles, dtype: float64

primer_alguna:
count    5837.000
mean       18.233
std         5.342
min         3.000
25%        15.000
50%        19.000
75%        23.000
max        26.000
Name: primer_alguna, dtype: float64

ultimo_control:
count    454901.000
mean         19.803
std           5.453
min           1.000
25%          16.000
50%          20.000
75%          25.000
max          26.000
Name: ultimo_control, dtype: float64
```

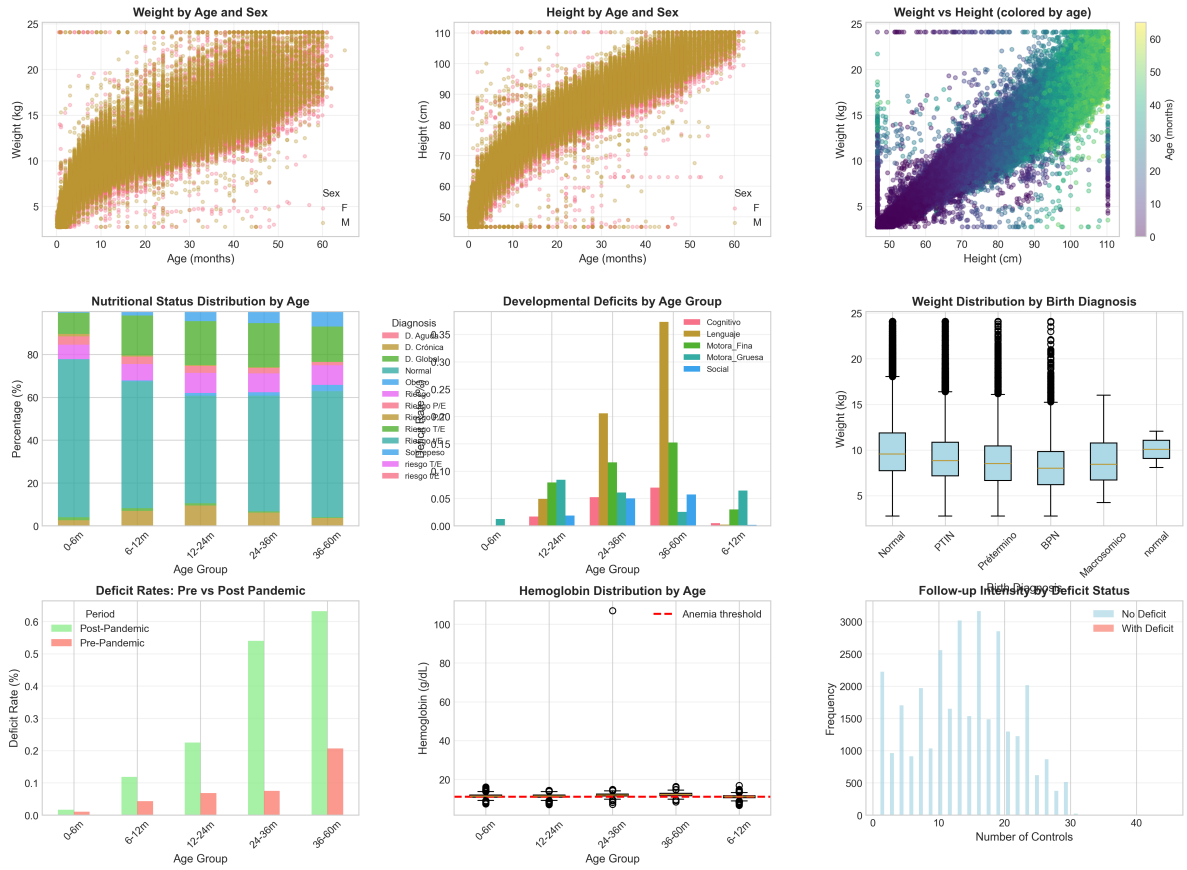


3.4. Initial Visualizations

==========

3.4. INITIAL VISUALIZATIONS

==========



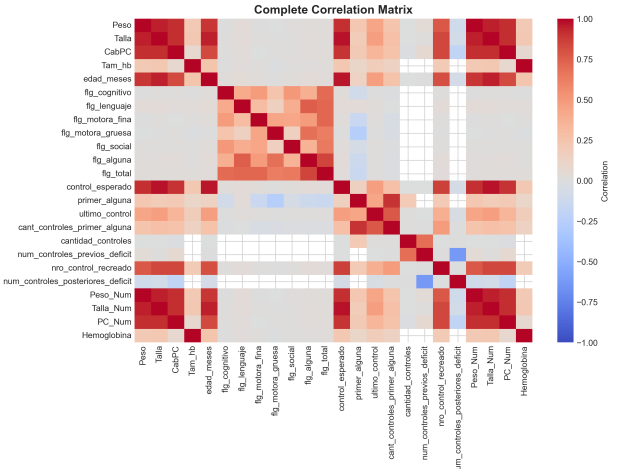

```
--- Statistical Comparisons ---

1. Anthropometric differences by sex (t-tests):
   Peso: t=-52.567, p=0.0000 ***
   Talla: t=-35.055, p=0.0000 ***
   CabPC: t=-65.677, p=0.0000 ***

2. Pandemic impact on deficit rates (Chi-square):
   Chi2=239.096, p=0.0000
   Deficit rate Pre-pandemic: 0.1%
   Deficit rate Post-pandemic: 0.2%

3. Birth diagnosis impact on weight (ANOVA):
   F=1102.696, p=0.0000
```

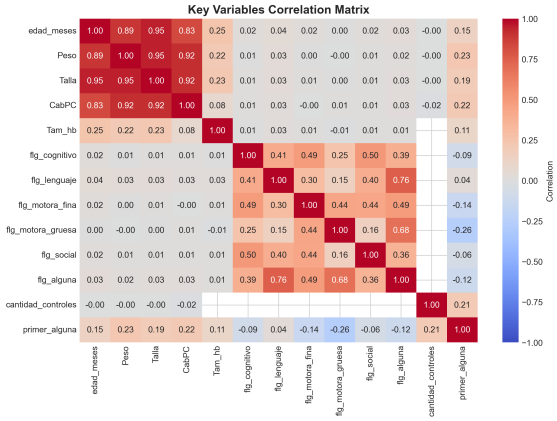
3.5. Outlier Detection and Treatment

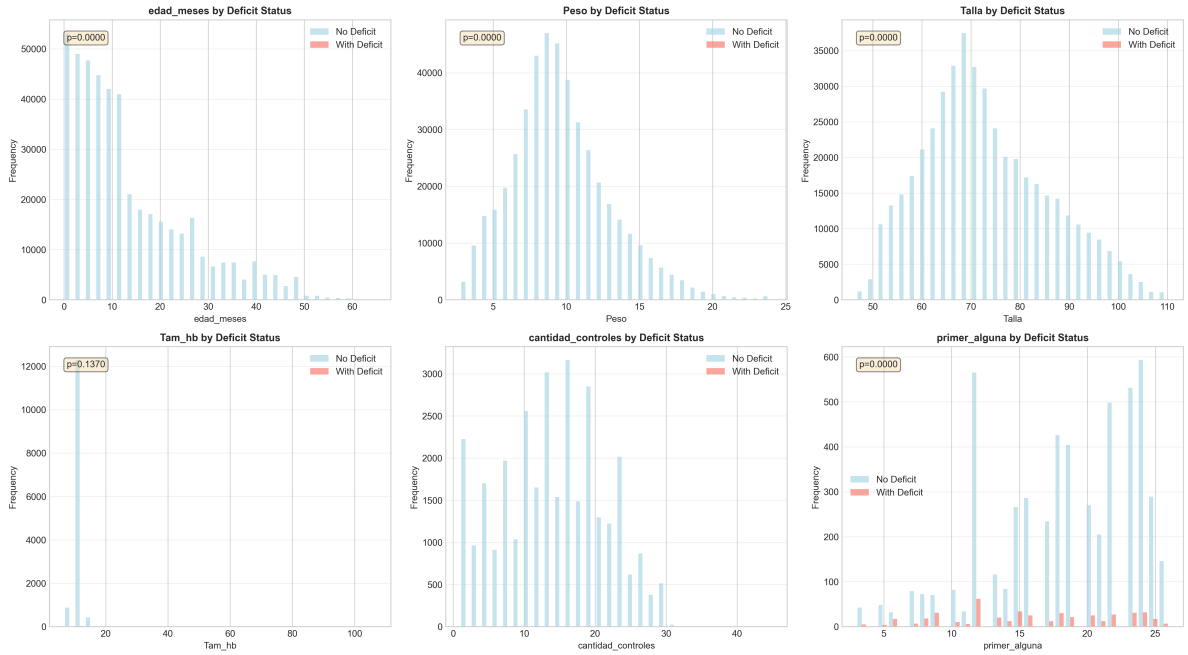
3.6. Correlation Analysis

```
================================================================================
3.6. CORRELATION ANALYSIS
================================================================================

Analyzing correlations among 24 numerical variables

--- Strongest Positive Correlations (r > 0.7) ---
            Var1              Var2  Correlation
          Tam_hb      Hemoglobina        1.000
            Peso         Peso_Num        1.000
           CabPC           PC_Num        1.000
           Talla        Talla_Num        1.000
           Talla control_esperado        0.972
control_esperado        Talla_Num        0.972
      edad_meses control_esperado        0.966
        Peso_Num        Talla_Num        0.949
           Talla         Peso_Num        0.949
            Peso            Talla        0.949

--- Strongest Negative Correlations (r < -0.3) ---
                          Var1                                Var2  Correlation
num_controles_previos_deficit num_controles_posteriores_deficit       -0.626

--- Correlations with Target Variable (flg_alguna) ---

Top 10 positive correlations:
flg_alguna              1.000
flg_total               0.845
flg_lenguaje            0.756
flg_motora_gruesa       0.681
flg_motora_fina         0.488
flg_cognitivo           0.385
flg_social              0.365
edad_meses              0.033
nro_control_recreado    0.032
control_esperado        0.032
Name: flg_alguna, dtype: float64
```

```
Top 10 negative correlations:
Peso_Num                          0.023
Peso                              0.023
ultimo_control                    0.014
Tam_hb                            0.013
Hemoglobina                       0.013
cant_controles_primer_alguna     -0.019
primer_alguna                    -0.124
cantidad_controles                  NaN
num_controles_previos_deficit       NaN
num_controles_posteriores_deficit   NaN
Name: flg_alguna, dtype: float64
```








```
--- Variable Selection Recommendations ---

Variables most correlated with developmental deficits:
flg_alguna              1.000
flg_total               0.845
flg_lenguaje            0.756
flg_motora_gruesa       0.681
flg_motora_fina         0.488
flg_cognitivo           0.385
flg_social              0.365
primer_alguna           0.124
edad_meses              0.033
nro_control_recreado    0.032
control_esperado        0.032
Talla_Num               0.026
Talla                   0.026
CabPC                   0.026
PC_Num                  0.026
Name: flg_alguna, dtype: float64

Recommended predictors for modeling (|r| > 0.1):
  1. flg_total (r=0.845)
```

2. flg_lenguaje (r=0.756)
3. flg_motora_gruesa (r=0.681)
4. flg_motora_fina (r=0.488)
5. flg_cognitivo (r=0.385)
6. flg_social (r=0.365)
7. primer_alguna (r=-0.124)

3.7. Preliminary Insights

---

4. Key Patterns and Insights

[To be developed based on EDA findings]

---

5. Predictive Modeling

[To be developed in next phase]

---

6. Recommendations for TANI

[To be developed based on model results]

---

7. Conclusions and Next Steps

[Final synthesis]

References

- Black, M. M., et al. (2017). Early childhood development coming of age: science through the life course. *The Lancet*, 389(10064), 77-90.
- World Health Organization. (2006). WHO Child Growth Standards.
- Instituto Nacional de Estadística e Informática (INEI). (2023). Encuesta Demográfica y de Salud Familiar.
- Díaz AA, Gallestey JB, Vargas-Machuca R, Velarde RA. Desarrollo infantil en zonas pobres de Perú [Child development in poor areas of Peru]. Rev Panam Salud Publica. 2017 Jun 8;41:e71. Spanish. doi: 10.26633/RPSP.2017.71. PMID: 28614480; PMCID: PMC6660845.
- National Research Council (US) and Institute of Medicine (US) Committee on Integrating the Science of Early Childhood Development; Shonkoff JP, Phillips DA, editors. From Neurons to Neighborhoods: The Science of Early Childhood Development. Washington (DC): National Academies Press (US); 2000. 8, The Developing Brain. Available from: https://www.ncbi.nlm.nih.gov/books/NBK225562/
- Bronfenbrenner, U. (1979). The ecology of human development. Harvard University Press.

Appendices

Appendix A: Data Dictionary

[Include complete variable definitions]

Appendix B: Technical Details

[Code repository, reproducibility instructions]