

MICS-EMIS Analysis

Data Analysis Manual

Connecting household and school data for better education insights



© UNICEF/UN7779162

October 2026

About this manual

This manual sets out the technical methodology for MICS-EMIS data preparation, linkage, analysis and interpretation. It is the technical companion to the MICS-EMIS Analysis Implementation Manual, which describes the institutional setup, roles, implementation process and policy-question framework.

It is intended for analysts in Ministries of Education, National Statistical Offices, UNICEF and other research partners involved in the work. It focuses on the analytical decisions that need to be applied consistently across country implementations, while allowing for country-specific education structures and EMIS systems.

Core methodological principle

MICS is a household sample survey that is representative of the population. EMIS is an administrative system that covers all schools, students and teachers registered in it. MICS-EMIS Analysis adds school information from EMIS to the MICS survey data, not the other way around. The linked MICS-EMIS dataset therefore remains a child-level MICS dataset: EMIS characteristics are attached to sampled children based on the school they attend. The linked schools are not a representative sample of schools.



© UNICEF/UN7234081

Acknowledgements

This manual was developed by the UNICEF Data and Analytics Education team, drawing on the methodological foundation of MICS-Link and on the pilot implementations of MICS-EMIS Analysis in Eswatini and Fiji in 2021-22.

Sincere appreciation is extended to government counterparts, particularly Ministries of Education and National Statistical Offices, and to UNICEF Country and Regional Offices, for their collaboration in developing and testing this methodology.

This work was supported by the Global Partnership for Education Knowledge and Innovation Exchange (GPE KIX), a joint endeavour with the International Development Research Centre (IDRC), Canada.



© UNICEF/UN7186503

List of abbreviations

Abbreviation	Definition
ECE	Early childhood education
EMIS	Education Management Information System
MICS	Multiple Indicator Cluster Survey
PSU	Primary sampling unit
SFR	Survey Findings Report
TWG	Technical Working Group

Contents

1. MICS-EMIS analytical framework.....	6
1.1 How school linkage is established through MICS-Link	6
1.2 The three core data products	6
1.3 Analytical workflow and units of analysis	7
2. Preparing MICS data	9
2.1 Start from the released MICS data.....	9
2.2 Harmonize the national education structure.....	9
2.3 Construct education variables.....	9
2.4 Prepare MICS-Link school-identification information	10
2.5 Preserve the survey design.....	11
3. Preparing EMIS data.....	12
3.1 Understand the EMIS structure and reference period.....	12
3.2 Establish the school register	12
3.3 Clean school identifiers and school names.....	12
3.4 Build school-level indicators.....	13
3.5 Validate the final school-level dataset.....	13
4. Linking MICS and EMIS.....	14
4.1 Define the linkage population.....	14
4.2 Direct matches	14
4.3 Manual matching of other-specify responses.....	15
4.4 Construct the linked dataset.....	15
4.5 Assess linkage coverage and selectivity.....	16
5. Sampling, analysis and interpretation	17
5.1 MICS, EMIS and linked analysis	17
5.2 What the linked data represent	17
5.3 Applying the MICS survey design	17
5.4 Linkage missingness and analytical denominators	18
5.5 Reporting domains and sample size	18
5.6 Core linked analyses.....	19
5.7 Interpretation	20
6. Outputs, reproducibility and documentation	21
6.1 Final data products.....	21
6.2 Analytical and dissemination products	21
6.3 Reproducibility and technical documentation.....	21
Annex. Practical technical reference.....	22
A. Minimum manual linkage log	22
B. Analysis and interpretation quick reference.....	22
C. Minimum quality-assurance checks.....	23
Related resources	23

1. MICS-EMIS analytical framework

MICS-EMIS Analysis combines nationally representative household survey data from MICS with administrative school data from EMIS. The methodology is designed so that school information can be added to the children already represented in MICS, creating a richer child-level dataset for education analysis. Because the linked data rest on the MICS sample, they keep its representativeness for children, but they do not form a sample of schools and cannot be used to describe or rank individual schools (see Section 5).

1.1 How school linkage is established through MICS-Link

The ability to link MICS children to EMIS schools is established before MICS data collection. MICS-EMIS Analysis builds on MICS-Link, an initiative of the UNICEF MICS team to connect MICS with external data sources. For MICS-EMIS Analysis, school-identification questions are incorporated into the MICS education module before fieldwork, so that the school attended by each eligible child can be identified later in EMIS.

In countries where the MICS survey uses an EMIS school list, respondents can select the school attended by the child, and the survey records a school code that can be mapped to the official EMIS identifier. When the school cannot be selected from the list, the questionnaire can capture additional identifying information such as the school name and location. These other-specify responses provide a second route for resolving the school during the analysis stage.

The data-analysis process therefore does not create school linkage from scratch. It starts from MICS data that already contain the school-identification information collected through MICS-Link, and then validates, resolves and connects those responses to the cleaned EMIS school register.

1.2 The three core data products

The methodology produces three core datasets, corresponding to the three analytical perspectives used throughout MICS-EMIS Analysis: MICS, EMIS and linked MICS-EMIS.

Data product	Unit of observation	Core content	Purpose
Clean MICS dataset	Child / person	Harmonized education variables, child and household characteristics, survey-design variables and MICS-Link school-identification information.	MICS education analysis and the child-level base for linkage.
Clean EMIS dataset	School	Unique school register plus harmonized school characteristics derived from available EMIS components.	EMIS system analysis and the school-level source used for linkage.
Linked MICS-EMIS dataset	MICS child	Relevant MICS variables plus the final matched school identifier, linkage status and selected EMIS school characteristics.	Analysis of the school environments experienced by MICS children.

1.3 Analytical workflow and units of analysis

The workflow proceeds in six stages:

1. Prepare and harmonize the MICS analytical dataset, including the education variables and MICS-Link school-identification fields.
2. Prepare and validate a unique school-level EMIS dataset.
3. Resolve direct and manual school matches and construct the linked MICS-EMIS dataset.
4. Assess linkage coverage and identify whether successful linkage differs across population groups.
5. Carry out MICS, EMIS and linked analyses using the correct unit of analysis, denominator and survey design.
6. Generate documented analytical tables, figures, dashboards and analytical products from the final datasets.

The unit of analysis is not the same across all three datasets. MICS analysis is child- or person-based and follows the MICS survey design. EMIS analysis is normally school-based, since the final dataset has one row per school. Linked MICS-EMIS analysis remains child-based: school characteristics are contextual information attached to the MICS child.

The two sources also differ in what they can represent. MICS interviews a sample of households selected by cluster, so its results are estimates: they are representative at national level and for the domains the survey was designed for, usually regions, but not for smaller areas, where estimates lose statistical stability. EMIS is normally school-based and aims to cover every school in the system, so it provides totals for the population within the system, such as enrolment, sometimes broken down by grade, class or age.

The link runs in one direction only: school information is added to MICS children, and no MICS information is added to EMIS. Many schools have no sampled MICS children at all, and others only one or a few, so the linked data cannot provide reliable information on the background of the children in a particular school.

Below are some examples of how different data sources can be used to answer different questions:

Question	Correct data source and unit	Meaning
What percentage of children are out of school?	MICS: child	A population estimate for the relevant MICS child population.
What percentage of schools have electricity?	EMIS: school	A school-system indicator among in-scope EMIS schools.
What percentage of children attend a school with electricity?	Linked MICS-EMIS: child	A population estimate of children's exposure to a school characteristic.

2. Preparing MICS data

MICS provides the child- and household-level base for MICS-EMIS Analysis. One of the strengths of the approach is that MICS follows standardized questionnaire structures, data files and survey methods across countries. Preparation therefore focuses on harmonizing the data in line with the national education structure, constructing the education variables required for analysis, and preparing the school-identification information collected through MICS-Link.

2.1 Start from the released MICS data

Analysis begins from the released MICS microdata and accompanying survey documentation, such as the Survey Findings Report (SFR), which contains the computed rates for standard MICS indicators. The standard household roster and age-specific MICS datasets contain the information needed for education analysis, with the precise modules depending on the survey. The household roster provides the main person-level basis for school attendance, grade and school-identification questions; other standardized MICS modules add information for relevant populations such as foundational learning or early childhood development.

Preparation should preserve the original MICS identifiers and source variables, so that every derived analytical variable can be traced back to the released survey data.

2.2 Harmonize the national education structure

MICS questions are standardized, but education systems differ in official entry ages, grade structures and the boundaries between education levels. Before deriving indicators, the country education structure should therefore be translated into a documented analytical mapping covering:

- Official ages for pre-primary, primary and secondary education;
- Grade ranges corresponding to each education level;
- Terminal grades used for completion and transition indicators;
- Country-specific level codes used in the MICS questionnaire;
- Any special national structures that affect the interpretation of grade or level.

This mapping should be used consistently across MICS-only analysis and linked analysis. Where international indicators require a specific definition, the derivation should follow the corresponding MICS methodology while reflecting the national grade structure. The SFR for the survey sets out the national education structure as applied in the survey, together with the final indicator calculations, and should be used as the reference for this mapping.

2.3 Construct education variables

The MICS analytical dataset should contain a standard set of documented education variables that can be reused across outputs. The exact set depends on questionnaire availability, but MICS-EMIS preparation commonly includes:

- Current school or early childhood education (ECE) attendance;
- Current education level and grade;
- Age-appropriate or adjusted attendance;
- Out-of-school status;

- Completion of relevant education levels;
- Repetition, dropout and non-transition;
- ECE attendance and early childhood development indicators;
- Foundational reading and numeracy indicators for the eligible child population;
- Household and home learning environment variables relevant to education equity.

Each derived variable should have a clear definition, eligible population, source variables and applicable MICS weight. Analytical outputs should be built from these prepared variables rather than repeatedly recoding the same concept in separate figures or dashboards.

The SFR is also the reference document for the modules available in the survey and for the final calculated indicators, against which derived variables should be checked. Indicators are computed in line with the methodology specified in the [MICS-EAGLE data analysis manual](#). SFRs, questionnaires and datasets for each survey are available on the [MICS website](#).

2.4 Prepare MICS-Link school-identification information

For children currently attending an eligible school or ECE programme, all school-identification information collected through MICS-Link should be retained. Depending on the country implementation, this may include the selected survey school code, reported school name, administrative location, education level and any other-specify school responses.

The original reported information should always be retained. Standardized versions of school names or codes may be created for matching, but the original responses provide the audit trail needed to review linkage decisions.

ED1. Line number	ED2. Name and age.	ED9. At any time during the academic year 2081, did (name) attend school or any Early Childhood Education programme?	ED10. During the academic year 2081, which grade or year is (name) attending?	ED10C1. Check ED10: Attended Grades between 00 to 08 ?	ED10C. Is (name) attending vocational/ technical level of education?	ED10D. In which province is (name) currently attending school or any Early Childhood Education programme?	ED10E. In which district is (name) currently attending school or any Early Childhood Education programme?	ED10F. In which palika is (name) currently attending school or any Early Childhood Education programme?	ED10G. What is the name of the school (name) is currently attending school or any Early Childhood Education programme?	
		1 YES 2 NO $\hookrightarrow$ ED15	CLASS: 00=ECE 01-12=CLASS 1-12 13=BACHELORS $\hookrightarrow$ ED15 14=MASTERS OR ABOVE $\hookrightarrow$ ED15 98=DK $\hookrightarrow$ ED15	1 YES $\hookrightarrow$ ED10D 2 NO	1 Yes 2 No	1 KOSHI 2 MADHESH 3 BAGMATI 4 GANDAKI 5 LUMBINI 6 KARNALI 7 SUDURPASCHIM 8 OTHER COUNTRY $\hookrightarrow$ ED15			Select school name from list. If school name not found, select 96 (Other) and enter school name and address.	
LINE	NAME	AGE	YES NO	GRADE	YES NO	PROVINCE	DISTRICT	PALIKA	SCHOOL NAME	EMIS CODE
01			1 2		1 2	1 2 3 4 5 6 7 8				
02			1 2		1 2	1 2 3 4 5 6 7 8				
03			1 2		1 2	1 2 3 4 5 6 7 8				
04			1 2		1 2	1 2 3 4 5 6 7 8				
05			1 2		1 2	1 2 3 4 5 6 7 8				
06			1 2		1 2	1 2 3 4 5 6 7 8				
07			1 2		1 2	1 2 3 4 5 6 7 8				
08			1 2		1 2	1 2 3 4 5 6 7 8				
09			1 2		1 2	1 2 3 4 5 6 7 8				
10			1 2		1 2	1 2 3 4 5 6 7 8				
11			1 2		1 2	1 2 3 4 5 6 7 8				
12			1 2		1 2	1 2 3 4 5 6 7 8				
13			1 2		1 2	1 2 3 4 5 6 7 8				
14			1 2		1 2	1 2 3 4 5 6 7 8				

Figure 1. Example of MICS school-identification questions capturing current attendance, province, district, palika, and school name. (Nepal MICS 2024-25)

2.5 Preserve the survey design

The prepared MICS data must retain the variables required for design-based analysis. These include the relevant sample weight, sampling strata and primary sampling unit (PSU) or cluster identifier.

Different MICS analytical populations can use different weights. The appropriate weight is determined by the population and module used for the indicator. For example, a foundational-learning outcome uses the weight applicable to the foundational-learning child population, while a roster-based schooling indicator uses the relevant weight for that population.

The table below shows the weight variables for the MICS data files most often used in education analysis. Variable names can differ between survey rounds and should always be confirmed in the survey documentation.

MICS data file	Population and typical education use	Weight variable
Household members (hl)	All household members listed in the household roster: school attendance, grade and school identification.	hhweight
Children under 5 (ch)	Children under 5: ECE attendance and early childhood development.	chweight
Children age 5–17 (fs)	One randomly selected child age 5–17 per household: foundational learning skills.	fsweight

The output of this chapter is a clean, documented MICS analytical dataset, ready for MICS-only analysis and for school linkage.

3. Preparing EMIS data

EMIS provides the school-level administrative information used in MICS-EMIS Analysis. Unlike MICS, EMIS structures vary substantially across countries. The objective of preparation is to turn the available administrative data into a documented, internally consistent school-level dataset that supports both standalone EMIS analysis and linkage to MICS.

3.1 Understand the EMIS structure and reference period

Before cleaning, document the available EMIS components, their units of observation, identifiers, reference periods and coverage. A country may have separate source tables or files for school registry information, student or enrolment records, teachers, facilities, resources, programmes, finance or geospatial information. Where the Ministry of Education provides a codebook explaining the data structure and variable names, as was the case in Ukraine, analysts should refer to it.

The same concept can appear in more than one administrative table. The authoritative source for core variables should therefore be agreed and documented, unless it is already set out in a codebook supplied with the EMIS dataset. The EMIS reference year also needs to be chosen deliberately: the EMIS data used should refer to the same school year as the MICS interviews or, where that year is not available, to the closest available year, with the difference documented.

3.2 Establish the school register

The school register is the anchor of EMIS preparation. It should contain one record for each in-scope school and a unique school identifier that can be used consistently across EMIS components and for MICS linkage. It should include:

- Unique EMIS school identifier;
- Official school name and relevant alternative names;
- Administrative geography and geographic codes;
- School operational status;
- School management, ownership or type;
- Education levels or grades offered;
- School coordinates, where available.

3.3 Clean school identifiers and school names

School identifiers and names should be standardized before any merge or linkage. Common issues include leading zeros, spaces, duplicate identifiers, obsolete codes, missing identifiers, spelling variants, and the same school represented differently across EMIS components.

Identifier problems should be resolved through documented rules. Where a duplicate represents the same school, one authoritative record should be retained after validation. Where an identifier corresponds to genuinely distinct institutions or campuses, the data structure should be corrected before the school file is used for linkage.

School names should be cleaned for analysis and matching while retaining the official version. Normalization may address capitalization, punctuation, spacing and standard abbreviations, but should not erase information needed to distinguish schools.

Most of these issues should have been resolved when MICS-Link was implemented with a clean school list in the questionnaire, but this needs to be validated before the final merge.

3.4 Build school-level indicators

Additional EMIS components should first be cleaned in their original unit and then converted into the school-level measures required for analysis. The available inputs differ by country, but common components include:

- Enrolment or student data, summarized by school and, where relevant, by sex, age, grade or level;
- Teacher records, summarized into school-level teacher counts and characteristics;
- Facilities and infrastructure, harmonized into clear availability indicators (yes or no) or quality indicators (relevant categories);
- Resources, programmes and services, with missing reporting distinguished from genuine absence;
- Geospatial information, checked for completeness, valid coordinates and obvious location errors.

The final school-level file used for linkage should contain one unique row per EMIS school identifier. Richer component-level datasets are intermediate products and may be retained separately for standalone EMIS analysis. Where a school offers several levels, for example primary and lower secondary, and the data allow it, level-specific indicators such as the pupil-teacher ratio should correspond to the level attended by the child (for example, the primary pupil-teacher ratio for a child in primary).

3.5 Validate the final school-level dataset

Validation should cover school coverage, missingness, duplicates, internal consistency and implausible values. At minimum, analysts should review whether core fields are available across the school register, and whether enrolment, teacher and facility information are plausible and internally consistent.

Data-quality issues that cannot be resolved should be documented rather than silently corrected. MICS-EMIS Analysis can itself help identify weaknesses in EMIS coverage and structure, so preserving this information is part of the analytical value of the exercise.

The output of this chapter is a clean, documented school-level EMIS dataset with one unique row per school for linkage.

4. Linking MICS and EMIS

Linkage identifies the EMIS school attended by each eligible MICS child. The linkage is at school level: the child is not matched to an individual administrative student record, and the MICS child remains the observation in the final linked dataset.

4.1 Define the linkage population

The linkage population consists of MICS children reported as currently attending an eligible school or ECE programme for which school-identification information was collected through MICS-Link. The eligibility rules should follow the questionnaire routing used in that survey, that is, the skip pattern applied before the school-identification questions are asked: whether the child is currently attending school, whether the school is in the country, and whether it is at one of the education levels covered by EMIS.

All linkage-eligible children must remain in the denominator when assessing linkage, including those whose school cannot ultimately be identified.

4.2 Direct matches

Where a school was selected during the MICS interview, the recorded MICS school code is mapped to the corresponding official EMIS identifier using the country school crosswalk. School codes and data can change from year to year, so the EMIS year used for the link must be the school year referred to in the MICS interview. Before the final link is made, both sides of the crosswalk should be standardized and checked for duplicate or conflicting mappings.

A direct match is accepted only when the MICS school code maps unambiguously to one valid EMIS school identifier. The school recorded by the interviewer is taken as correct, since an individual response cannot be verified afterwards, and distance is not a reliable check because children can attend schools far from home. Simple validation checks remain useful, in particular comparing the education level, and where available the school type, reported in MICS with those recorded for the school in EMIS.



© UNICEF/UN7244144

4.3 Manual matching of other-specify responses

Manual matching is often a tedious and difficult process, and in some cases it cannot be done reliably. It is always better to be cautious: a school should not be matched unless it is clearly the same school. Whether manual matching is worthwhile should be explored in each country; where the direct linkage rate is already high, it may not be needed.

When the respondent could not select the school from the list, the response in the “other: please specify” category is reviewed against the cleaned EMIS school register. The matching process uses the reported school name together with location and, where useful, education level or programme information.

1. Retain the original school name and create a standardized version for comparison.
2. Restrict candidate EMIS schools using the location of the school recorded in MICS.
3. Where useful, narrow candidates further using education level, programme type or other reported characteristics.
4. Compare the reported school name with candidate EMIS school names, allowing for spelling differences, abbreviations, local naming conventions and transliteration.
5. Assign an EMIS identifier only where the combined information supports one sufficiently clear school.
6. If more than one school remains plausible, classify the case as unresolved rather than forcing a match.
7. Record the decision and the supporting evidence in a manual linkage log.

Automated string comparison or fuzzy matching can help identify possible candidates, but the final manual decision should use all available information and remain auditable.

4.4 Construct the linked dataset

The final link is many-to-one: many MICS children may attend the same EMIS school, while the cleaned EMIS linkage file must contain only one record per school identifier.

All linkage-eligible MICS children should be retained in the linked dataset. Confirmed matches receive the selected EMIS school characteristics; children without a confirmed match retain missing EMIS characteristics and an explicit linkage status. A child can also be matched to a school that exists in the register but has no EMIS data for the reference year, which can happen with private schools, for example. These cases should be flagged separately, so that missing school data are not confused with failed linkage. Linkage can then be reported in two steps, as was done in Eswatini: the share of children matched to a school identifier, and the share matched to a school with EMIS data.

Linkage status	Definition
Direct match	School identified through the MICS-Link school code and the validated crosswalk.
Manual confirmed match	School identified from other-specify information after structured review, with sufficient evidence.
Unresolved / ambiguous	Potential candidates exist, but the available information does not identify one school sufficiently clearly.
No match	No plausible corresponding EMIS school can be identified.

The final merge should be checked to confirm that no MICS child records were inadvertently duplicated or dropped, and that EMIS variables are clearly distinguishable from MICS household and geographic variables.

4.5 Assess linkage coverage and selectivity

Linkage diagnostics are required before substantive linked analysis. The primary operational linkage rate is the unweighted proportion of linkage-eligible MICS child records with a confirmed EMIS match.

A weighted linkage rate should also be calculated using the relevant MICS weight. This supplementary measure shows whether successfully linked and unlinked records differ in their population importance; it does not replace the unweighted operational rate.

Linkage should be examined overall and, where sample size permits, by MICS reporting geography, education level, sex, residence and household wealth. Differences between linked and unlinked children should be documented, because selective non-linkage can affect the interpretation of all subsequent linked estimates. The MICS weights are not adjusted for failed linkage (see Section 5.4), which is why a high linkage rate and this check for selective non-linkage are so important.

5. Sampling, analysis and interpretation

The statistical interpretation of MICS-EMIS Analysis is anchored in the original MICS sample design. Linking school characteristics to MICS children expands the information available for each child, but it does not create a new sample of schools or new geographic representativeness.

5.1 MICS, EMIS and linked analysis

Different kinds of questions can be answered with MICS-EMIS Analysis, but careful consideration has to be taken regarding the type of question and which dataset can answer which question. Some examples illustrate this:

Analysis type	Unit	Example	Denominator
MICS	Child / person	Share of primary-school-age children who are out of school.	Relevant MICS target population.
EMIS	School	Share of schools with a functioning water source.	In-scope EMIS schools with usable information.
Linked MICS-EMIS	MICS child	Share of attending children, by wealth quintile, sex or area of residence, whose school has a functioning water source.	Relevant linked MICS child population with usable school information.

This clearly shows that it is important to always consider the denominator for the interpretation of results. Specifically, MICS-EMIS linked analysis only represents children that attend school and have been linked to an EMIS school.

5.2 What the linked data represent

In linked analysis, school variables become characteristics of the educational environment experienced by the sampled child. If several sampled children attend the same school, the same school information is attached to each of those children, because the estimation concerns children, not schools.

The linked data should therefore not be used to produce school rankings, school-specific learning estimates, or statements about the distribution of child characteristics across all schools. Most schools will contain very few linked MICS children, or none at all, and schools were not selected through a representative school sampling design. The data are, however, representative of children attending school, provided that linkage is not selective (see Section 4.5).

5.3 Applying the MICS survey design

Core rule

Use the original MICS weight appropriate to the analytical population or module. Linking EMIS information to the child does not create a new standard MICS-EMIS weight, and no EMIS or school weight is applied to linked child-level estimates.

The appropriate MICS weight depends on the population being analysed. A roster-based schooling indicator uses the weight specified for that MICS population, while a foundational-learning outcome uses the weight applicable to the foundational-learning child sample.

For standard errors and confidence intervals, the analysis should account for the full MICS complex survey design, including the relevant weight, sampling strata and PSU or cluster. The country MICS survey documentation, in particular the SFR, is the authoritative source for these variables and for the survey's representative reporting domains.

5.4 Linkage missingness and analytical denominators

The standard MICS-EMIS methodology does not create an automatic additional adjustment for failed school linkage, and the MICS sample weights are not recalculated. Reweighting would be a complex process, comparable to the original sampling and weighting of the survey. The methodology relies instead on a high linkage rate and on checking that linkage is not selective (see Section 4.5), and the original MICS weight remains the analytical weight.

When an analysis requires an EMIS school characteristic, the analytical denominator consists of children in the relevant MICS population who have a confirmed school link and non-missing information for that school variable. The unweighted analytical denominator should be retained in the output.

If linkage success differs meaningfully across population groups, the limitation should be stated and considered in interpretation. Linkage diagnostics make this issue visible, rather than masking it through an ad hoc adjustment.



© UNICEF/UN7280256

5.5 Reporting domains and sample size

MICS-EMIS linked estimates should normally be reported only for the geographic domains, usually regions or other subnational administrative units, for which the relevant MICS survey was designed to provide representative estimates. Linkage does not create representative estimates for districts, municipalities or schools, since they are not MICS reporting domains. Estimates for such areas require a separate modelling approach, such as small area estimation.

The unweighted number of observations underlying every linked estimate should be retained. For standard reporting, the following thresholds apply unless the survey documentation requires a stricter rule:

- 50 or more unweighted observations: report normally;
- 25–49 unweighted observations: report with a small-sample caution;
- Fewer than 25 unweighted observations: suppress the estimate.

These thresholds apply to the actual analytical denominator after linkage and variable-availability restrictions.

5.6 Core linked analyses

The core linked analyses follow the five areas of the policy questions framework in Annex 2 of the Implementation Manual. Under each area, the bullets list the questions the analysis should respond to, adapted to the priorities agreed by the Technical Working Group (TWG).

Linkage coverage and representativeness

Assess which children can be linked to EMIS schools and whether linkage differs systematically across MICS population groups. This is a prerequisite for interpreting all subsequent linked analysis.

- What proportion of MICS children can be matched to an EMIS school?
- Does linkage success vary by geography, household wealth, or urban/rural location?
- Which EMIS schools are represented among matched MICS children, and which schools remain unmatched?

Household–school allocation and equity

Examine which school environments are experienced by children from different household backgrounds. Examples include school resources, management type, teacher availability and infrastructure, by household wealth, residence or other MICS equity characteristics.

- Do poorer households send children to systematically different schools than wealthier households?
- Are household deprivations (water, sanitation, electricity, food security) mirrored in school conditions?
- Are education programmes, such as school feeding, health, or inclusion programmes, reaching children and schools with higher levels of vulnerability?

Distance to school and physical access

Where suitable household or cluster location data and school coordinates are available, analyse physical access to school. The air distance as well as walking or driving distances can be calculated by the MICS geospatial experts, whilst preserving the security of the coordinates of the households. For all children, the distance to the nearest suitable school in the EMIS register is calculated; for children attending school, the distance to the school they actually attend is calculated as well.

- How far do children travel to school, and how does this vary across regions and population groups?
- Is distance associated with attendance, dropout, or out-of-school status?

- Which children face the greatest physical access barriers?

Learning outcomes and combined learning environments

For children with valid MICS learning outcomes and linked school information, compare learning across household and school conditions. These analyses describe associations unless a separate analytical design supports causal interpretation.

- How are foundational learning outcomes associated with household characteristics?
- How are foundational learning outcomes associated with school characteristics?
- Which combinations of home and school environments are linked to learning gaps, and do school resources mitigate or amplify household disadvantage in learning outcomes?

Indicator validation and system consistency

Related MICS and EMIS indicators (out-of-school rate, repetition, dropout) can be compared to understand broad consistency, complementary information and possible data-quality issues. Differences should be interpreted in light of different populations, denominators, reference periods and data-generation processes. EMIS-based rates also rely on population estimates or projections for their denominators, which can lead to implausible figures, such as net enrolment rates above 100 per cent. This often occurs at local level, where children attend school in a city or urban area but live outside it, so that more children appear to be in school than the population projection for that area contains.

- How consistent are key indicators (enrolment, attendance, age-for-grade) between MICS and EMIS?
- Can combined data be used to estimate district-level indicators, such as numbers of out-of-school children?
- Where do discrepancies between sources suggest measurement gaps or under-reporting?

5.7 Interpretation

- Linked estimates describe MICS children and the school environments they experience; they do not describe a representative sample of schools.
- Associations between school characteristics and child outcomes should not be presented as causal effects without a separate causal design.
- Differences in MICS and EMIS reference years should be stated whenever they affect interpretation.
- Failed linkage and missing EMIS information reduce the analytical sample and may be selective.
- Differences between MICS and EMIS indicators can arise from definitions, denominators, reference periods and coverage, not only from data error.
- School-level characteristics repeat for children attending the same school; this clustering should be considered if analyses move beyond standard descriptive estimates.

6. Outputs, reproducibility and documentation

MICS-EMIS Analysis should produce reusable data products as well as analytical outputs. The same prepared datasets should underpin every table, figure, dashboard and publication.

6.1 Final data products

Data product	Minimum content
Clean MICS dataset	Harmonized education variables, child and household characteristics, survey-design fields and school-identification information.
Clean EMIS dataset	Unique school register and harmonized school-level indicators from available administrative data.
Linked MICS-EMIS dataset	MICS analytical variables, final school identifier, linkage status and selected EMIS school characteristics.

6.2 Analytical and dissemination products

Depending on national priorities and the questions agreed by the TWG, the final datasets can generate:

- MICS, EMIS and linked MICS-EMIS dashboard outputs;
- Analytical reports and thematic studies;
- Policy briefs and other concise products for decision-making;
- Data-quality and linkage diagnostics supporting future improvements to MICS-Link and EMIS.

Indicators should be calculated in the analytical workflow and passed to dissemination products, rather than being independently recalculated inside dashboards or presentation files.

6.3 Reproducibility and technical documentation

The analytical workflow should allow another analyst to reproduce the final datasets and results from the agreed source data. At minimum:

- Raw MICS and EMIS source data are preserved and not overwritten;
- MICS preparation, EMIS preparation, linkage and output construction are separate documented stages;
- Derived variables, country-specific recodes and methodological decisions are documented;
- Manual school matches are stored in a durable linkage log;
- Intermediate and final datasets are clearly distinguished and versioned;
- Analytical outputs can be regenerated from the final datasets without manual recalculation.

The technical handover should include the final MICS, EMIS and linked datasets and data dictionaries, the school manual-match log, linkage diagnostics, the MICS survey-design information used in the analysis, documentation of key methodological decisions, and all codes used for data cleaning and analysis.

Annex. Practical technical reference

This annex consolidates the reference items that should be applied consistently across country implementations.

A. Minimum manual linkage log

Every manually resolved case should be recorded with the following fields. Linkage status definitions are set out in Section 4.4.

Field	Purpose
MICS internal record key	Allows the linkage decision to be reproduced without relying on the names of children or households.
Original reported school name	Preserves the source response.
Standardized comparison name	Supports systematic matching.
Reported location	Province or region, district and any other available location information.
Education level or programme	Helps distinguish candidate schools where relevant.
Candidate EMIS school(s)	Documents the schools considered during review.
Final EMIS identifier	Recorded only for a confirmed match.
Final linkage status	Direct, manual confirmed, unresolved/ambiguous, or no match.
Decision note	Brief evidence supporting the final decision.

B. Analysis and interpretation quick reference

Question	Unit	Statistical treatment
What share of children are out of school?	MICS child	Use the relevant MICS weight and the full MICS survey design.
What share of schools have a library?	EMIS school	Use the in-scope EMIS school denominator; this is not a MICS-weighted estimate.
What share of attending children, by wealth quintile or other disaggregators, are in schools with a library?	Linked MICS child	Use the relevant MICS weight and full survey design among the applicable linked analytical population.
What share of eligible MICS records were linked?	MICS record	Report the unweighted operational linkage rate; retain a MICS-weighted rate as a supplementary diagnostic.
What is the learning level of School X?	Not applicable	Not a valid MICS-EMIS estimate. Do not use linked MICS children to rank or estimate individual school performance.

C. Minimum quality-assurance checks

- MICS education variables and national education-level mappings are documented and applied consistently.
- The relevant MICS weight, strata and PSU variables are retained for every analytical population.
- The EMIS linkage dataset contains one unique row per school identifier.
- School identifiers and names are standardized, and duplicate mappings are resolved before linkage.
- All eligible MICS children remain in the linked dataset, including unsuccessful matches.
- Ambiguous manual matches are not forced.
- The final merge does not duplicate or drop MICS child records.
- Unweighted and weighted linkage diagnostics are reviewed before substantive linked analysis.
- Linked outputs retain the unweighted analytical denominator and follow the small-sample reporting rules.
- Linked estimates are interpreted as child-level estimates of school exposure, not school-level estimates.
- Reference periods, missing information and country-specific methodological decisions are documented.

Related resources

[MICS-EMIS Analysis webpage](#)

[MICS7 tools and survey-design resources](#)

[MICS website: Survey Findings Reports, questionnaires and datasets](#)

[MICS-EAGLE data analysis manual](#)

Survey-specific MICS questionnaires, sampling reports and datasets should be consulted for the authoritative country-specific survey design, weights, reporting domains and variable availability.