



RESEARCH METRICS AFRICA

FIELD DATA INTEGRITY

Turn field data into decisions you can defend.

SYNTHETIC DEMONSTRATION · NOT A CLIENT ENGAGEMENT

NGO Data Quality Audit

Juja hypertension screening practice dataset

122	70	15	0
RECORDS REVIEWED	SCREENED OUTCOME	REVIEW-SIGNAL RECORDS	PARTICIPANT PII

Executive decision brief

This sample shows how a field-data export can be screened before it reaches a donor report, management dashboard or inferential analysis. The workflow separates confirmed data defects from patterns that still require field context.

BOTTOM LINE

The dataset was usable for a controlled demonstration, but not ready for unrestricted inferential reporting.

Fifteen records required human review, documented range and logic failures needed correction, and the planned chi-square test was withheld because expected-cell assumptions were not met.

What this audit answers

- Which records need correction or supervisor review before reporting?
- Which collection or coding controls should be strengthened for the next cycle?
- Which statistical analysis can proceed, and which must be withheld?
- What evidence can the M&E lead share with management without overstating the findings?

This report is a marketing demonstration built from synthetic data. It does not describe work completed for an NGO and does not imply endorsement by any organisation.



01 · CONTEXT AND SCOPE

The audit begins with the data-generating process.

Practice study objective: assess factors associated with uptake of hypertension screening among adults in Juja Sub-County. The workbook was designed to demonstrate a realistic NGO-style quality review while containing no participant-identifiable information.

DATA PACKAGE	122 synthetic records; README; Raw Data; Codebook; Expected Issues; 25 planted scenarios.
STRUCTURE	Outcome, predictor/group, identifier and covariate roles mapped before analysis.
QUALITY SCREEN	Missingness, duplicates, range and logic, interview duration and possible straight-lining.
LOCATION SCOPE	Available coordinates can be screened. Historical network/IP evidence cannot be reconstructed from a static upload.
ANALYSIS GATE	Statistical testing proceeds only after data sufficiency and method assumptions are checked.

Evidence boundary

- Automated flags are screening tools, not proof of fraud, intent or enumerator misconduct.
- Records should be checked against the questionnaire, field context, supervisor notes and data-management plan.
- Direct identifiers should not be sent through a general enquiry channel; access and de-identification must be agreed first.
- GPS Layer 2 requires a consented live Research Metrics Africa check-in workflow and is not inferred retrospectively.

CONTROL

No automatic deletion

A flagged row stays in the audit trail until an authorised reviewer records the verification and corrective action.



02 · FINDINGS AND ISSUE REGISTER

Signals were prioritised, not sensationalised.

15	6	2	3
REVIEW-SIGNAL RECORDS	RANGE OR LOGIC	DURATION ANOMALIES	POSSIBLE STRAIGHT-LINE

ISSUE	COUNT	INTERPRETATION	NEXT ACTION
One or more review signals	15	Record requires human review; categories may overlap.	Compare with source form and field notes.
Range or logic violation	6	Value or response path conflicts with a documented rule.	Verify coding, skip logic and correction authority.
Interview-duration anomaly	2	Elapsed time differs materially from the expected pattern.	Review interview context; do not infer intent.
Possible straight-lined interview	3	Repeated response pattern merits contextual review.	Check instrument design and supervisor evidence.
GPS Layer 2	N/A	Not assessed from the static demonstration upload.	Use consented live check-ins when required.

ANALYSIS GATE

Planned chi-square test: WITHHELD

Expected-cell assumptions were not met. No p-value was forced. The result remains descriptive until categories, sample support and the substantive question can be reviewed.

Why this matters

A donor-facing report should not convert a software warning into an accusation, or convert a failed assumption into a confident statistical claim. The audit documents both boundaries.



03 · CORRECTIVE ACTION PLAN

Every signal is assigned to a decision owner.

The aim is a traceable remediation cycle: who verifies the issue, what evidence they use, what action they take and what must be re-run before approval.

OWNER	IMMEDIATE ACTION	EVIDENCE / OUTPUT
M&E; lead	Confirm indicator definitions and material reporting risks.	Approved indicator map and decision log.
Data manager	Review the 15 flagged records; correct only with documented authority.	Issue register and reproducible cleaning log.
Field supervisor	Provide context for duration, repetition or location signals.	Supervisor verification notes linked to record IDs.
Analyst	Re-run quality rules and the analysis assumption gate.	Validated descriptive output; justified test or documented hold.

Recommended next cycle

- Align the Kobo/ODK form, codebook and donor indicator definitions before the next deployment.
- Run the same documented quality rules on each new batch instead of waiting for final reporting.
- Set named review thresholds and turnaround times for the M&E; lead, data manager and field supervisor.
- Keep corrections, exclusions and unresolved signals in an auditable cleaning log.

GPS NOTE

Marked not assessed - by design

Static uploads cannot reconstruct historical network or live check-in evidence. A consented check-in layer must be agreed before fieldwork.



04 · SAMPLE DELIVERABLE

What an NGO receives from a paid pilot.

A Dataset Quality Audit is scoped around one authorised dataset. Fees depend on dataset size, documentation, turnaround and reporting requirements.

01	Audit summary Material risks, decisions, limitations and priorities in management-ready language.
02	Issue register Record-level or rule-level findings with status, owner, evidence and action.
03	Cleaning log A reproducible account of authorised corrections, exclusions and unresolved questions.
04	Codebook notes Variable roles, labels, value definitions, missing codes and control recommendations.
05	Preliminary report Descriptives and assumption-gated analysis labelled for analyst validation.

KSh 15k-30k	20 min	Custom
ONE-DATASET AUDIT	INTRODUCTORY DEMO	MONTHLY NGO QA

NEXT STEP

Book a 20-minute synthetic-data demonstration

See the workflow before discussing access to an authorised dataset. Do not email participant-identifiable data. researchmetricsafrica.com/services/ngo-data-quality-audit

Final evidence note

Automated flags are screening tools, not proof of fraud. All findings require validation against the questionnaire, field context and the programme's data-management plan. This document is preliminary and requires analyst validation.

Research Metrics Africa · Kenya · researchmetricsafrica.com