Name: Joel Livingstone Kofi Ackah

MedTrack Ghana – Data Quality Assessment Report

Organization: MedTrack Ghana

Role: Junior Developer (DevOps Perspective)

Date: February 25, 2026

Data Quality Findings (By Dimension)

Dimension	Issue Identified	Violation / Impact
Accuracy	Patient P002 has two different phone numbers recorded: 244789012 and 0244789012.	Factually incorrect data; at least one number is wrong, causing SMS delivery failure.
Completeness	Patient P004 has a blank PatientName field.	Missing required identity field; complicates clinical documentation and billing.
Consistency	Doctor names (Dr. Osei vs dr. osei) and Payment status (Paid vs paid) use inconsistent casing.	Same logical values stored in multiple formats; causes duplicate reporting counts.
Timeliness	Appointment dates use mixed formats like 2025-10-15 and 10/16/2025.	Ambiguity (MM/DD vs DD/MM) leads to wrongly scheduled reminders or missed appointments.
Validity	Phone number 244789012 is only 9 digits.	Does not follow Ghana's 10-digit mobile format (must start with 0); system rejects the record.
Uniqueness	Patient P001 and P002 appear as duplicate records.	Duplicate IDs inflate patient reports and cause double billing.

Business Impact Analysis

- **Operations:** Failed SMS reminders due to invalid phone numbers and ambiguous dates lead to high "no-show" rates for appointments.
- **Clinical:** Missing names (P004) and inconsistent doctor records create "Identification Errors," making it hard for doctors to track patient history accurately.
- **Finance:** Duplicate IDs and inconsistent PaymentStatus lead to failed billing reconciliations and overcharging patients.

Recommended Solutions (Top 3 Critical Issues)

Critical Issue	Technical Solution (DevOps/Backend)	Responsibility	Verification
Uniqueness (Duplicate IDs)	Enforce **Primary Key constraints** on PatientID in the DB schema.	Database Administrator	Run a SQL query to confirm COUNT(DISTINCT PatientID) matches total rows.
Validity (Phone Numbers)	Implement **Regex validation** (10 digits, starts with 0) at the API/Input layer.	Backend Developer / QA	Monitor SMS gateway logs for "Success" vs "Invalid Number" errors.
Consistency (Date Formats)	Standardize all dates to **ISO 8601 (YYYY-MM-DD)**; use dropdowns for UI inputs.	Backend Developer / Data Engineer	Verify that Cron jobs for reminders trigger exactly 24 hours before appointments.

Biggest Risk of Poor Data Consistency

From a **DevOps and Systems perspective**, the biggest risk is **System Integration Failure**. When data formats are inconsistent, automated pipelines (ETL), APIs, and third-party

integrations (like SMS gateways or billing systems) fail silently or produce "garbage in, garbage out" results. This increases **technical debt** and makes the system nearly impossible to scale reliably.

Conclusion

The MedTrack dataset fails across all six dimensions. By implementing strict validation rules and database constraints, we can restore "Single Source of Truth" integrity, ensuring the platform remains trustworthy for both clinicians and patients.