iCliniq Data Services – Quality Assurance (QA) Process

At iCliniq, we follow a rigorous, multi-step QA process to ensure that all medical datasets meet the highest standards for clinical accuracy, privacy, and regulatory compliance. Below is a detailed overview of our QA workflow:

1. Source Validation

- All data originates from real doctor–patient interactions on the iCliniq platform, ensuring clinical authenticity.
- Only content from certified doctors and verified patient profiles is included to maintain data reliability.

2. Medical Expert Review (Double-Layer Review Process)

Each data point is reviewed by a medical domain expert for quality, followed by a senior reviewer to ensure consistency.

Breakdown of Review Focus:

- **Clinical Relevance**
  Checks whether the content aligns with the patient's complaint, condition, and expected response standards.

- **Diagnostic Accuracy**
  Ensures the diagnosis or guidance aligns with accepted medical guidelines or protocols (e.g., WHO, CDC, ICMR).

- **Contextual Completeness**
  Validates whether the response or case summary covers necessary follow-ups, risk factors, or missing context.

- **Terminology Correctness**
  Confirms the use of appropriate medical terms (e.g., ICD-10, SNOMED) for accuracy and standardization.

3. Dataset Structuring QA

- All structured/semi-structured data is checked for schema integrity, valid field mappings, and format compliance.

- Manual spot-checks are performed to ensure field-level accuracy and maintain logical flow (e.g., SOAP notes, timestamps).

4. De-Identification & Privacy QA

- All records are de-identified using HIPAA Safe Harbor guidelines.
- Personally identifiable information (PII) and sensitive phrases are removed, especially in free-text responses or attachments.

5. Client-Specific QA Customization

- QA criteria are adjusted based on the client's needs (e.g., oncology focus, training dataset, rare disease coverage).
- We provide pre-delivery QA reports, format conversions, metadata tagging, or filtering based on specialty.

6. iCliniq Evals (AI Dataset Validation Framework)

- This tool benchmarks AI-generated answers against doctor-verified gold responses across multiple quality dimensions.
- Evaluates each response for safety, hallucination risk, clinical reasoning, and relevance to query context.

7. Documentation & Reporting

- For every delivery, a QA Summary Report is shared covering error counts, reviewer notes, and de-identification status.
- If data lacks specific context or input, we attach a placeholder report file and mark it for future metadata update.

8. Ongoing QA & Recertification

- We perform content re-audit every 6 months to ensure clinical relevance is up-to-date with evolving guidelines.
- Content flagged as outdated or needing updates is reviewed and re-annotated from the active dataset pool.

9. Metadata Contribution and Validation

- For each dataset, iCliniq adds rich clinical metadata such as:
    - Specialty tags (e.g., Oncology, Dermatology)
    - Clinical categories (symptoms, diagnosis, treatment types)
    - Contextual markers (acute/chronic, follow-up required, etc.)

    - Temporal and demographic attributes (age group, gender, case urgency)
- Metadata tagging is performed through in-house medical reviewers. They tag metadata for accuracy and clinical relevance
- Each metadata layer undergoes:
    - Tag consistency checks across cases
    - Specialty-wise distribution analysis to avoid data imbalance
    - Error flagging for missing or irrelevant tags
- Metadata is aligned with search and downstream modeling goals (e.g., fine-tuning LLMs, training a model to triage cases based on urgency, supporting specialty-specific use cases, classifying medical conversations by disease)
- Final datasets are delivered with metadata index files or embedded fields, depending on client format requirements (JSON, CSV, XML, FHIR, etc.)