**Proposal for Collecting a Comprehensive Dengue Dataset**

**1. Title** Protocol for the Collection and Integration of Clinical, Laboratory, and Imaging Data for Dengue Fever Diagnosis and Analysis

**2. Background and Objective** Dengue fever, a mosquito-borne viral infection, poses significant public health challenges globally. Effective diagnosis and management of dengue fever require accurate and comprehensive datasets that include patient symptoms, clinical findings, laboratory results, and imaging data. This protocol outlines a systematic approach to collect and curate a dataset focused on dengue diagnosis, integrating clinical features, laboratory test results, and platelet count imaging from pathology departments.

**3. Study Design** This study is observational and cross-sectional, collecting data from dengue-suspected patients in a clinical and pathology setup. The dataset will include both categorical and continuous variables, as well as associated platelet count imaging data.

**4. Data Collection Framework**

**4.1. Inclusion and Exclusion Criteria**

* **Inclusion Criteria:**
  + Patients presenting with symptoms suggestive of dengue fever (e.g., fever, headache, retro-orbital pain, rash, myalgia, arthralgia).
  + Patients undergoing dengue-specific diagnostic tests (NS1 antigen, IgG, IgM).
  + Patients consenting to provide clinical and imaging data.
* **Exclusion Criteria:**
  + Patients with confirmed alternative diagnoses.
  + Patients with incomplete records or refusal to consent.

**4.2. Data Features**

* **Patient Information:**
  + Patient Number (unique identifier)
  + Sex (M/F)
  + Age (in years)
* **Patient-Reported Symptoms:**
  + Fever (duration, intensity)
  + Headache
  + Retro-orbital pain
  + Rash
  + Muscle pain (myalgia)
  + Joint pain (arthralgia)
  + Nausea/vomiting
  + Abdominal pain
  + Bleeding manifestations (e.g., gum bleeding, petechiae)
* **Clinician-Observed Signs:**
  + Temperature
  + Rash distribution
  + Capillary refill time
  + Signs of dehydration (e.g., sunken eyes, dry mucous membranes)
  + Hepatomegaly
  + Tenderness (abdominal, muscle)
  + Hemodynamic parameters (e.g., blood pressure, pulse rate)
* **Laboratory Test Results (Suggested by Doctors):**
  + Complete Blood Count (CBC):
    - Platelet count
    - Hematocrit
    - White blood cell (WBC) count
  + Dengue-specific tests:
    - NS1 antigen
    - IgG/IgM antibodies
  + Liver function tests (e.g., AST, ALT levels)
* **Pathology Imaging Data:**
  + Platelet count images from blood smears (captured using high-resolution pathology imaging systems).
* **Final Outcome:**
  + Diagnosis status (Dengue positive/negative).
  + Severity classification (mild, moderate, severe/dengue hemorrhagic fever).

**5. Data Collection Procedure**

**5.1. Clinical Data Collection**

* Structured case report forms (CRFs) will be used to capture patient-reported symptoms and clinician-observed signs.
* Data will be manually entered by trained personnel and cross-verified by medical officers.

**5.2. Laboratory Test Results**

* Results of diagnostic tests will be obtained from collaborating diagnostic laboratories and pathology departments.
* All test results will be matched to the corresponding patient using unique identifiers.

**5.3. Pathology Imaging Collection**

* High-resolution images of blood smears for platelet count will be captured using microscopy.
* Images will be annotated and linked to patient records for cross-referencing.
* Collaborations with pathology labs will ensure standardized imaging protocols.

**6. Ethical Considerations**

* Informed consent will be obtained from all patients before data collection.
* The study will adhere to ethical guidelines, ensuring confidentiality and anonymization of patient data.
* Approval from the Institutional Review Board (IRB) will be secured before initiating data collection.

**7. Data Quality Assurance**

* Data will be reviewed for consistency and completeness.
* Automated tools will be used to detect and correct errors in data entry.
* A random subset of data will be validated by independent reviewers.

**8. Data Integration and Storage**

* Clinical, laboratory, and imaging data will be integrated into a secure database with access control.
* Data will be stored in compliance with institutional data security policies.

**9. Data Analysis Plan**

* Statistical methods will be used to analyze relationships between symptoms, laboratory results, and imaging findings.
* Machine learning models will be developed to classify and predict dengue diagnosis and severity.

**10. Timeline**

* Month 1-2: Ethical approvals, personnel training, pilot testing of data collection.
* Month 3-8: Full-scale data collection and quality assurance.
* Month 9-10: Data integration, preliminary analysis.
* Month 11-12: Final analysis and report preparation.

**11. Budget and Resources**

* Personnel: Data collectors, medical officers, pathologists, and imaging technicians.
* Equipment: High-resolution microscopes, imaging software.
* Miscellaneous: CRF printing, data storage systems.

**12. Expected Outcomes**

* A comprehensive, multimodal dengue dataset suitable for diagnostic and predictive modeling.
* Insights into clinical and laboratory predictors of dengue severity.
* A resource for future research and algorithm development.

**13. Collaboration Plan**

* Clinical Departments: Patient recruitment, symptom and sign documentation.
* Pathology Departments: Laboratory test results and platelet count imaging.
* Data Science Teams: Data integration, analysis, and modeling.

**14. References**

* WHO Dengue Guidelines for Diagnosis, Treatment, Prevention, and Control.
* Institutional protocols for data collection and storage.