**Clinical Research Protocol (S21-147)**

*This research protocol was previously approved as S10-79 and S20-025 by the Institutional Clinical Research Ethics Review Board of Saitama Medical Center, Jichi Medical University. It was subsequently submitted as an additional application and reapproved as S21-147 on February 10, 2022.*

**1. Title and Classification of the Research**

**(1) Title of the Research**

**Association Between Extreme Outliers in Clinical Laboratory Test Results and Prognosis**

**(2) Classification of the Research**

* ☑ Research conducted solely by Jichi Medical University
* ☐ Multi-center collaborative research led by Jichi Medical University
* ☐ Multi-center collaborative research led by another institution

**2. Research Implementation Framework**

**(1) Implementation Framework within the Institution**

* **Principal Investigator:**  
  Hitoshi Sugawara, MD, PhD, Division of General Medicine (Completed Ethical Training)
* **Co-investigators:**
  + Takahiko Fukuchi, MD, PhD, Division of General Medicine (Completed Ethical Training)
  + Tamami Watanabe, MD, PhD, Division of General Medicine (Completed Ethical Training)
  + Keishiro Sueda, MD, Division of General Medicine (Completed Ethical Training)

**3. Objective and Significance of the Research**

**(1) Objective**

To elucidate the relationship between extreme outliers in clinical laboratory tests and short-term prognosis within three days of testing.

**(2) Significance**

In routine clinical practice, patients occasionally exhibit extreme outliers in laboratory test results. However, the short-term prognosis and risk factors for such cases remain unclear. By identifying background factors, this research aims to facilitate short-term prognosis prediction, thereby improving the quality of medical care in primary care settings.

**4. Research Methods and Duration**

**(1) Research Methods**

**1. Study Design**

* **☑ Clinical Research**
* ☐ Epidemiological Research
* ☐ Other Medical Research
* **☐ Interventional Study**
  + ☐ Involves more than minimal invasiveness
  + ☐ No or minimal invasiveness
* **☑ Observational Study**
  + ☐ Involves more than minimal invasiveness
  + ☑ No or minimal invasiveness

**2. Specific Details**

Anonymous biochemical blood test data (complete blood count panel, biochemistry panel, coagulation panel) collected over 14 years from January 1, 2004, to December 31, 2023, will be analyzed. The dataset will exclude duplicates and erroneous samples (e.g., blood drawn from an intravenous route, CPA, cases with unknown short-term prognosis). Data from the top and bottom 0.1% of each test parameter will be extracted. Test items with at least 10% (≥50 cases) of deaths within three days among 500 cases will be analyzed.

**3. Estimated Sample Size and Rationale**

* **Estimated sample size per test item:** 500 patients
* **Rationale:** Previous clinical research indicates that approximately 0.1% of test cases correspond to 500 patients.

**4. Statistical Analysis Methods and Evaluation Criteria**

* **Statistical Analysis Methods:**
  + Retrospective analysis of risk factors such as age, sex, vital signs, Charlson Comorbidity Index (CCI), underlying disease, and biochemical test data.
  + Short-term prognosis defined as death within three days of testing or survival beyond three days.
  + Univariate analysis will be performed with significant factors subsequently analyzed using multiple logistic regression analysis with ROC analysis and stepwise selection.
  + Machine learning analysis will be performed using gradient boosting methods (e.g., XGBoost, LightGBM, or CatBoost) to identify significant predictors and compare predictive performance with traditional statistical methods.
* **Evaluation Criteria:**
  + Identify risk factors predicting short-term prognosis for extreme outliers.
  + Develop a prognostic prediction formula using multiple logistic regression models and machine learning models.
  + Compare regression coefficients, intercepts, adjusted odds ratios, 95% confidence intervals, C-statistics, sensitivity, specificity, and likelihood ratios between models.
  + Determine the simplest model that best predicts short-term prognosis based on the number of risk factors and C-statistics.

**(2) Research Duration**

* Until **March 31, 2025**, following approval.
* Results will be published in conferences and journals within **two years** after analysis completion.

**5. Selection of Research Subjects**

**(1) Number and Type of Subjects**

* **☑ Patients (Approx. 500 per test item)**
  + **No specific disease name designated**
  + **☑ Outpatients from Jichi Medical University Hospital and Saitama Medical Center**
  + **☑ Inpatients from Jichi Medical University Hospital and Saitama Medical Center**

**(2) Inclusion Criteria**

* Data extracted from the **top and bottom 0.1%** of each test parameter in the laboratory database.

**(3) Exclusion Criteria**

* Duplicate cases
* Erroneous samples (e.g., collected from an IV route)
* CPAOA cases
* Cases with unknown short-term prognosis

**(4) Age Limitations**

* **☑ Limited (18 years and older)**

**(5) Gender**

* **☑ Both male and female**

**(6) Recruitment Methods**

* **☑ No recruitment of subjects**

**6. Scientific Rationale**

Large volumes of blood test data are accumulated daily in medical institutions but are not fully utilized to improve healthcare quality. In clinical practice, extremely abnormal laboratory values are occasionally observed. These extreme values could indicate life-threatening conditions if not promptly addressed. Identifying predictive risk factors will enable rapid prognosis assessment and appropriate intervention. Machine learning techniques will enhance the predictive accuracy and provide novel insights into complex relationships among variables.

**7. Informed Consent Procedures**

**(1) Use of Samples and Information**

* **☑ Use of Existing Data**
  + **Data Source:** Blood test database from January 1, 2004, to December 31, 2023.
* **☑ Use of Information**
  + **☑ Opt-out method applied**
  + **Rationale:** As this is a retrospective study utilizing existing medical records, informed consent is waived per the "Ethical Guidelines for Medical and Health Research Involving Human Subjects" (Chapter 5, Section 12.1(2)-i), ensuring subjects have an opportunity to opt out.

**8. Handling of Personal Information**

* **☑ Anonymization with Correspondence Table**
  + Hospital IDs will be replaced with alternative identifiers.
  + **Rationale:** Ensuring data accuracy, allowing corrections, and enabling withdrawal upon request.
* **☑ No exchange of samples or information with other institutions.**

**9. Risk-Benefit Assessment and Mitigation Measures**

* **No financial burden, risk, or discomfort for participants.**
* **Strict management of personal information will be ensured.**