Project Title: Evaluation of an Automated Information Extraction Tool for Identification of Venous Thromboembolism in Electronic Health Records (EHRs)

CIO Name: CDC / Office of Noncommunicable Diseases, Injury and Environmental Health (ONDIEH) / National Center on Birth Defects and Developmental Disabilities (NCBDDD) / Division of Blood Disorders

Special Eligibility Requirements (If applicable, justification for any special requirements that would relate to eligibility to apply for the funding opportunity. May include requirements related to organizational capacity or unique skills or expertise needed to successfully execute the project. Applicants must substantiate that they meet these requirements.):

Awardees should have:

- Access to and use of their proposed healthcare setting's EHR/EMR data sources including radiology data;
- A method for providing a gold standard data source, or
- An understanding of VTE occurrence and its diagnosis

Project Description (The problem the project addresses; the need for the project; target populations, if applicable; key outcomes to be achieved by the end of the project period; how the project contributes to developing the workforce that contributes to the public's health.):

Problem statement:
Venous thromboembolism (VTE), including both deep vein thrombosis (DVT) and pulmonary embolism (PE), is an important and growing public health issue. The precise number of people affected by VTE is unknown; however, estimates suggest that 300,000 to 900,000 events occur in the United States yearly, with healthcare costs up to $10 billion. VTE is associated with substantial health impacts including death. In the United States, it is estimated that as many as 100,000 people die of PE each year. With many of the risks for VTE, such as obesity, advanced age, and chronic diseases increasing in the US population, we can expect to see increasing numbers of people affected by VTE.

Need for the project:
Surveillance data are needed to quantify the burden, provide a baseline to measure the effectiveness of interventions aimed at preventing VTE and help to better understand where to focus research and prevention efforts. Presently, there is no national surveillance for VTE and current estimates are thought to be underestimates due to the many challenges that make it
difficult to conduct surveillance of VTE. Thus, novel innovative methods are needed to accurately capture and monitor the burden of VTE.

The adoption of electronic health records (EHRs) provides public health with an opportunity to develop innovative and efficient ways to facilitate and improve surveillance methods. A significant challenge to conducting VTE surveillance is case identification; using ICD-9/10 codes alone is not sufficient. Documenting clinically relevant VTE requires information from radiologic imaging or autopsy, evidence of symptoms from clinical notes, and evidence of treatment from pharmacy records. However, many of these data are found only in the clinical notes or text fields of the EHR. Manually extracting these data would require substantial user time and effort and would be very costly. Thus, an electronic tool that can read and extract this information directly from the EHR will greatly enhance CDC’s ability to conduct accurate and efficient surveillance. In the past year, work has been conducted in collaboration with Emory University to evaluate the utility and accuracy of IDEAL-X, a novel open source information extraction system, for identifying VTE diagnosis.

Key outcomes:
The purpose of this WIP project is to scale-up and validate IDEAL-X to further assess the capacity of the tool to successfully perform VTE identification in at least two other hospital settings with different electronic health record systems.

- Obtaining an appropriate Test EHR data source for testing IDEAL-X
- IDEAL-X generated VTE dataset containing VTE cases identified by IDEAL-X
- Gold Standard VTE dataset: Dataset of confirmed VTE cases to use as gold standard
- Evaluation report of IDEAL-X accuracy and usability
- Final Report describing methods and findings

This technology could help accelerate more efficient and accurate VTE surveillance in healthcare settings by providing more accurate estimates of the incidence of VTE within hospital systems and nationally. This innovative project combines the state-of-the-art health informatics technologies, and the need for improved VTE surveillance in healthcare settings. With continued support, this technology could help accelerate VTE surveillance in healthcare by providing more accurate estimates of the incidence of VTE within hospitals and nationally. With better estimates of the burden of disease, improved patient safety, morbidity and mortality, can ensue.

Target population:
The target population for this funding opportunity are Academic Medical Centers/Schools of Medicine and Nursing.

How this project contributes to developing the workforce that contributes to the public’s health:
This project is a community-level activity that supports public health practice, namely population-based disease (VTE) surveillance. In addition, this project will evaluate a prevention or health promotion program or strategy of public health importance. The lessons learned from
this project will provide the needed foundation to guide and evaluate prevention of VTE within communities.

**Awardee Strategies and Approaches** *(Potential strategies or approaches the awardee should address to implement the project; any required or recommended strategic partnerships or collaborations for implementing the project.)*

To successfully implement this project, the awardee should work closely with CDC to implement and evaluate the IDEAL-X software in their setting's Electronic Health Record (EHR) or Electronic Medical Record (EMR) to identify confirmed cases of VTE and compare the VTE cases findings to a gold standard. The gold standard can be either a previously implemented active VTE surveillance system that accurately identifies VTE or through a chart review process. Participate in regular meetings, communications, and other activities with CDC that support successful and collaborative implementation and evaluation of the IDEAL-X tool for VTE identification. Create and deliver to CDC a final dataset (with data dictionary) of IDEAL-X generated data.

In collaboration with CDC conduct an evaluation of the IDEAL-X tool's ability to correctly identify confirmed cases of VTE and develop a report of the results of the project methods and results of project and evaluation.

In collaboration with CDC, draft and disseminate methods and findings in the form of presentations, peer reviewed manuscripts, fact sheets, or other publications and materials (e.g. newsletters, federal documents, social media, etc.).

**CDC-CIO Staff Activities** *(CDC-CIO project staff involvement in the project, including the provision of technical assistance and other program support activities and program monitoring activities.)*

CDC will provide technical assistance to awardees as needed to help achieve goals of the project. This includes providing awardees with access to IDEAL-X, technical support and assistance to implement the IDEAL-X tool and appropriate VTE case definitions and essential data elements and case identification algorithms. CDC staff will work collaboratively with awardees to evaluate the tool's effectiveness and write-up and disseminate findings. All results and associated presentations, abstracts, reports and manuscripts will include CDC as a collaborator and will require CDC scientific clearance before publication/presentation.

**Review Criteria** *(Applications will undergo an objective review process. Criteria are listed in descending order by weight, with the total value equaling 100 points.)*

Eligible applications will be evaluated against the following criteria:

a. **Understanding of the Public Health Impact** *(10 points)*
   Description of public health burden that the project will address, understanding of the need for
this project and the planned activities within the target community. Clear, concise descriptions and understanding of the requirements, objectives, and purpose of this FOA.

b. **Objectives (5 points)**
   Clear and measurable objectives to address the need and activities of this FOA.

c. **EHR Test Data Availability (40 points)**
   Clear description of the EHR or EMR data source including but not limited to: population, time period, data access and any applicable fees. Clear description of the gold standard data source including the method of ensuring accuracy of VTE diagnosis. Documentation of legal authority to access data sources needed to accomplish activities and goals. Methods of accessing data and implementing IDEAL-X specific and feasible within the project timeframe.

d. **Plan and Timeline (30 points)**
   Complete and comprehensive plan to carry out proposed objectives for the entire project period. Clear description of the procedures to be used to evaluate the IDEAL-X accuracy, including a proposed analysis plan and list of potential products. Reasonable and achievable timeline to complete the activities in the specified project period; description of ability to begin immediate implementation of the activities. Description of planned collaboration with CDC and any other awardees and partners. Ability to accomplish the project objectives.

e. **Organizational and Personnel Capacity (15 points)**
   Appropriate expertise and skills. Demonstrated capacity and infrastructure to manage federal grants. Ability to identify and timely hire of needed staff and consultants. Staff roles defined and sufficient to accomplish program goals.

**Other Information (Any additional information that would be useful for applicants.):**

For more information on CDC’s VTE program: [http://www.cdc.gov/ncbddd/dvt/index.html](http://www.cdc.gov/ncbddd/dvt/index.html)

**Length of Project Period (1-3 years):** 1 year

**Approximate Average Award:** $200,000

*(The award for the first 12-month budget period, including direct and indirect costs)*

**Approximate Number of Awards:** 2