NSF | CHOT

COLLABORATIVE PROPOSAL FORM

PROJECT TITLE: Building Heterogeneous Health Databanks: Costs, Benefits, Limitations, and Best Practices

PROPOSAL NUMBER: 13-06191.TAM.PSU	PIs: Ferris, Lee, Tucker	
RESEARCH THEME: Analytics & Innovative Technologies		
BUDGET:	UNIVERSITIES: TAMU, PSU, GIT	PROJECT YEAR: 1

PROJECT DESCRIPTION:

This effort will involve identifying challenges to creating and maintaining an accessible database that includes expanded heterogeneous sources, as well as the benefits available to accessing such a database. Our work not only includes health data commonly stored in EHR systems, but also integrates new types (qualitative, descriptive) and sources (socioeconomic, demographic, mental health) of data. The researchers will analyze and compare costs and benefits for multiple populations, and provide foundation of best practices for integrating heterogeneous data for greater patient acuity and improved modeling of patient risk.

HOW THIS IS DIFFERENT THAN RELATED RESEARCH:

To the knowledge of CHOT members, there does not exist any published work that documents and seeks to define patterns among challenges and best practices in healthcare database curation, integration, and providing access to health data at the point of care. Furthermore, this research attempts to integrate multi-site, multi-source, and multi-type medical data into a database that will be accessible to the healthcare industry. This effort will bring together perspectives on this process from several members of industry in healthcare systems that are varied in geographic location and population demographics.

COLLABORATIVE PLAN:

Each university site will coordinate with IAB members to compile and contribute to a shared list of types/variables of health data which have value for their research and IAB member interests, noting in particular when multiple variables from separate databases are required for decisions/models. This may be facilitated by a survey. The variables will be classified and characterized according to costs, benefits, limitations, and best practices for compiling and storing these data in a way that allows for controlled accessibility (e.g., CHOT members & researchers). Later efforts will seek to apply the best practices in an exploration of creating a CHOT-accessible heterogeneous data repository.

EXPECTED MILESTONES:

(1) Literature review
(2) IRB approval for member survey
(3) Survey of IAB members/broad healthcare population on key challenges & best practices in building and managing heterogeneous health databanks
(4) Publication of results of survey
(5) Implementation of best practices

BENEFITS TO INDUSTRY:

IAB members benefit from the sharing of challenges and best practices that are implementable in the building/management of heterogeneous health databanks. Initial steps can be taken toward shared databanks that combine input across several entities/health data sources.

EXPECTED COLLABORATIVE DELIVERABLES:

(1) Provide best practices for data storage, consolidation, and management
(2) Form extensive list of publicly available data for use
(3) Implement lessons learned/best practices at IAB member institutions
(4) Create CHOT database for member utilization