Biostatistics Scope of Work Agreement

General Information

Investigator	Alice White, Elaine Scallan alice.white@ucdenver.edu	**Date**	May 2, 2018
Project Number	P1330White		
Project Title	Predictors of food source in foodborne outbreaks		

Project Cost and Milestones

Project Type: Average Data Analysis/Publication

Billing Phase and Milestone	Cost
Phase 1: Project Start Up Discuss and review project materials, establish timelines, deliverables, and data structures with biostatistician.	$ 450
Phase 2: Exploratory Analysis Establish preliminary analysis dataset, run descriptive statistics and graphics, and create a report.	$ 3525
Phase 3: Comprehensive Analysis Complete comprehensive analysis and present a report. Additional changes to analysis are anticipated and part of the Project Complete phase.	$ 3525
Phase 4: Project Complete Complete final analysis and publication quality figures.	$ 450
Customization:	0
Total Due	**$ 7950**

Understanding of Project

Project Description

Goal: The goal of this project is to evaluate characteristics of foodborne outbreaks and develop a predictive model to identify food sources of future outbreaks. Using data available from the CDC national surveillance for foodborne outbreaks, we will examine characteristics associated with food source for *e. coli* outbreaks.

Project History: Co-investigator Alice White first examined predictors of *e. coli* outbreaks as part of her MS thesis work. Later, a biostats MS student continued work on this project. Ed Bedrick, a former faculty member in the Dept. of B&I advised on Alice's committee and some on the subsequent work. For these analyses, food sources were categorized as leafy greens, beef, and dairy, and predictors were evaluated for prediction of each food source using logistic regression and decision trees.

Data: Data from the CDC national surveillance for foodborne outbreaks between 1998 and 2016 will be used, excluding outbreaks from which the food source was not identified. An outbreak is defined as illness from a food source in 2 or more people. Additional variables contained in the surveillance data include percentage of females, age categories, number of cases in outbreak, season, setting (restaurant or home), and whether the outbreak was limited to a single state or multiple states.

Analysis Plan: The CIDA biostatistics team will evaluate surveillance data on *e. coli* outbreaks. Logistic regression and decisions tree models were selected to analyze a prior version of the data; conversations with Dr. Bedrick indicate that alternative model choices are not unreasonable. We will select the most reasonable predictive modeling approach, and can compare findings to either the logistic regression models or the decision tree models.

Notes: The end goal of this project is to develop a tool to share with investigators early on in an outbreak to give a starting point for investigating food sources. Our analysis is limited in that we do not have "live" outbreak information, and thus cannot evaluate model performance on a continuously updating set of information. However, findings from the prior analyses suggest that season and age will be the most relevant factors for predicting food source, and thus we do not anticipate a more complex model will be feasible, even if real-time data were available for model development. Additionally, the focus of this tool is to point investigators in a reasonable direction (i.e. should they investigate leafy greens first, or beef, etc.), so our goal will be to understand the predictive capabilities of the model based on varying factors and sample sizes. We will provide the investigators with the mathematical framework / equations / algorithm for prediction of food source, along with assistance with interpretation of these findings, for the development of this tool.

Timelines/Deadlines

Anticipated timeline: Funding for this project ends July 31, and thus the majority of work and all milestones are to be complete beforehand. The kickoff meeting should happen no later than May 31. During the kickoff meeting, additional meetings should be scheduled, with an anticipated project completion date of July 31, 2018.

Note: The investigators submitted a data request to the CDC at the end of April and hope to have data in the month of May. However, they do have a prior set of data (1998 – 2014), and thus we should be able to begin exploring model options even without the data.

Approval of Agreement

By approving this Scope of Work Agreement, you are acknowledging that you have read and agree to the project costs and milestones, timelines, project details, and terms and conditions outlined in this document.

To approve this Scope of Work Agreement click the button below.

Approve Scope of Work Agreement ▸

(If you don't agree with this Scope of Work Agreement or would like to withdraw your request for CIDA services, please send us an email to cida@ucdenver.edu with a brief explanation.)

Terms and Conditions

Clean data requirements - ready for analysis

The data are assumed to be cleaned and ready for analyses unless otherwise agreed upon, and a data dictionary should be provided to the analyst. We strongly encourage the use of REDCap as a data collection and management tool.

Report writing, abstract and manuscript preparation and revision

A final report will be created with an introduction, statistical methods, and results section. These sections will be close to publication ready. The CIDA biostatistician will edit the methods and results section for publication and read the final version of the manuscript prior to submission. Assuming the biostatistician has provided significant contribution to the manuscript in terms of performing analyses and contributing to the results and methods sections, the biostatistician shall be a co-author on the publication, acknowledging the intellectual contribution of the work.

Assuming no substantial new analysis is needed, the CIDA biostatistician will assist with writing a response to reviewer's statistical questions, make revisions to the paper and review the final version of any revised manuscript. If substantial new analysis is required, a new scope of work will be created and with costs agreed upon by both parties.

CIDA Authorship Guidelines

The CIDA abides by the International Committee of Medical Journal Editors (ICMJE) guidelines concerning authorship. Visit our CIDA website to learn more about CIDA's authorship policies.

Specific CIDA guidelines include:

- The biostatistician performing the analysis will be a co-author on the publication to acknowledge the intellectual contribution to the work. Statistician co-authors will use their primary appointment affiliation on manuscripts and abstracts.
- To maintain study and statistical integrity, data collected for publication and abstracts will only be analyzed after study completion.
- The CIDA biostatistician performs the analysis, collaborates in the structuring of the presentation of the results, and writes the "statistical methods" section of the paper.
- The biostatistician reviews the publication and any revisions prior to submission.
- The biostatistician will assist with revisions, keeping in mind your revision deadlines.

CIDA's right to cancel or close out a project

Please approve the Scope of Work (SOW) within 15 days (or prior to anticipated start of work, if less). SOWs not approved within 30 days will be closed. Projects which remain inactive for over 60 days will be closed unless prior arrangements have been made, and a final bill will be sent for work completed.

CCTSI subsidized projects

If the project cost is subsidized by the Colorado Clinical and Translational Sciences Institute (CCTSI), you are required to cite the CCTSI grant in posters and publications. Please review the CCTSI's Citation and CTSA grant language.