Guidelines for Collaboration and Consulting Services Provided by Center for Surgical Outcomes Research (CSOR) Epidemiologists and Biostatisticians

Investigators are expected to adhere to the following policies when they make a project request

Process for Initiating Study Collaboration

1. Carefully review the following guidelines for working with CSOR Biostatisticians and Epidemiologists.
2. Fill out the CSOR protocol form (available on our website).
3. Send the finished protocol form to CSOR's directors, Dr. Van Arendonk (kyle.vanarendonk@nationwidechildrens.org) and Dr. McLeod (daryl.mcleod@nationwidechildrens.org) for review to verify that the study proposal is both robust and achievable. Please allow 2-3 weeks for this review process. Multiple revisions may be necessary before your protocol is approved.
4. After approval, your study will be assigned to a CSOR statistician or epidemiologist.

Services Offered

- Database queries
- Assistance with study design
- Statistical analyses
- Preparation of tables and figures
- Power and sample size calculations
- Writing of statistical analysis sections of grant proposals, manuscripts, and abstracts
- Writing and editing results and methods sections of manuscripts and abstracts
- Reviewing and editing final abstracts, manuscripts, presentations, and grant proposals prior to submission
- Assistance with responding to grant and manuscript reviewers' comments

Timeline Expectations

Abstracts: 6 weeks minimum after initial consultation. Must sign off on the final, approved study protocol and provide a clean dataset (if not using a database we will query) no less than 4 weeks before the submission deadline.

Manuscripts: 8 weeks minimum after initial consultation. Must sign off on the final, approved study protocol and provide a clean dataset (if not using a database we will query) no less than 6 weeks before the desired submission date.

Grant proposals: Contact us to determine reasonable deadline.

Collaboration Expectations

We recommend involving one of our epidemiologists/biostatisticians from the very beginning of the project (prior to data collection or extraction). This will facilitate optimal study design and analysis.

The Principal Investigator (clinical or research faculty) on the project must be present at the first in-person consultation meeting. Some level of PI engagement, remote or otherwise, is expected through project completion.

A thorough review of the published literature on the research question of interest should have already been performed at the time of the initial consultation.

An opportunity for final review before submission of any abstract, presentation, or paper is expected.

Co-authorship is expected. However, the epidemiologist/biostatistician's name must not be placed on an abstract, presentation, or paper without their consent.

Although there is no maximum number of projects on which a PI may collaborate with us, we expect projects in which we have invested time to be completed. If a PI is unresponsive or has not completed a manuscript or grant proposal within a reasonable time period, we will not initiate any new projects with this PI until all ongoing projects are completed.

Completion of prior projects and level of PI engagement will be considered when reviewing new projects from previous collaborators.

Epidemiologists/biostatisticians will periodically contact the PI/study team for pertinent updates. After six months of non-responsiveness, a study will be considered "inactive" and removed from the biostatistician/epidemiologist's work queue. At that time, if the PI/study team wish to continue the project, an interim consultation is recommended to address any changes in protocol or study team.

Externally funded projects are prioritized over other projects. Therefore, CSOR epidemiologists/biostatisticians may have limited time to work on collaborative projects at times when they are heavily engaged in work for grant-funded projects.

When requesting a study, we strongly recommend the addition of table shells to best communicate with the analytics team what information and relationships you would like to investigate (see example below).

Table 1. Cohort Description

	Cohort n=XX (XX.X%)	**Simple Appendicitis** n=XX (XX.X%)	**Complicated Appendicitis** n=XX (XX.X%)	**p-value**
Race				
Asian				
Black				
White				
Other				
Unknown/refused				
Ethnicity				
Hispanic				
non-Hispanic				
Primary language				
Arabic				
English				
Nepali				
Spanish				
Other				
Used interpreter				
Age at encounter				
Admission day				
Weekday				
Weekend				
Insurance type				
Private				
Public, PFK				
Public, non-PFK				
Uninsured				

Data Guidelines

If you are not using an established administrative dataset, please adhere to the following guidelines when collecting/creating your data.

- RedCAP or similar is preferred to Microsoft Excel.
- When formatting data:
  - Keep variable names to a minimum length.
  - Units of measurement should not be included in a data point.
  - Units of measurement should be kept consistent for each variable (i.e., height in centimeters/inches or weight in kg/lbs/g). Similarly, the formatting of dates should remain consistent (i.e., YYYY/MM/DD, MM/DD/YYYY).

weight (kg)
10
12
15

body weight of patient
12kg
10205g
15kg

- Missing values should be given a specified value, such as '-9999' for continuous variables and 'MISS' for categorical variables, rather than left blank to allow for differentiation between what is known to be missing information and what was unintentionally not filled out.
- Ensure consistent spelling and capitalization for categorical variables.

race
White
Black/AA
Black/AA
Asian
MISS
Asian

race
White
black/AA
Black
Asian
asain

- When using binary variables, use 1=Yes/Presence of variable and 0=No/Absence of variable.
- A data dictionary should be included for each dataset created. This should include, at minimum: a short description of each variable, the units (if applicable), and a list of all possible levels of the variable if categorical.

variable	description	levels/units
MRN	mrn of patient	
RACE	patient race	White, Black, Asian, Other, MISS
WEIGHT	patient weight	kgs
ADMIT_DATE	date of first admission at NCH	DDMMYYYY (10MAR2023)
ED_VISIT	patient in ED during index visit	1=Yes, 0=No, -9999=MISS

- While your assigned Biostatistician/Epidemiologist can assist with more complicated data cleaning and manipulation, it is expected that basic data cleaning as outlined above is completed before sending over your data.
- In terms of timeline for completion, the day we receive a complete and clean dataset is considered to be Day 1.