

Health Statistics and Data Science Core

General Guidelines on Statistical and Data Analytic Collaboration and Consultation (September 28th, 2023)

Overview

The Health Statistics and Data Science (HSDS) Core assists researchers with all sizes and types of research projects. Specific services include:

- **Pre-award grant applications**, including study design, sample size calculation, statistical analysis plan, and data analytic strategies.
- **Funded research projects**, including NIH, NSF, foundations, and other extramural funding agencies.
- **Unfunded research projects**, including DUSON pilot or Duke internal grants, and other small research projects that need guidance on data analysis, results interpretation, and/or manuscript preparation.

The aim of these general guidelines is to foster an effective partnership between principal investigators (PIs) and HSDS Core’s faculty (as co-investigators [HSDS Co-Is]) and biostatisticians, and to ensure the quality of data, the rigor of statistics and data analysis, the appropriateness of interpretation of results, and the completion of the project in a timely manner. The following sections address the responsibilities of an HSDS Co-I and a biostatistician, the responsibilities of a PI from an HSDS Co-I’s and biostatistician’s point of view, and a brief summary of the ethical principles that guide statistical and data analysis collaboration and consultation.

A. Responsibilities of an HSDS Co-I and a biostatistician on a research project (see Table 1):

Table 1. Responsibility and Roles of HSDS Core Members

Type	HSDS Co-I	Biostatistician
Pre-award grant applications	<ul style="list-style-type: none"> • May attend the color-coded (blue, pink, green, and red) teams review meetings • Draft method section related to research design, statistical analysis, and/or data analytics 	<ul style="list-style-type: none"> • Not required

Funded research projects	<ul style="list-style-type: none"> • Work with PI to create a long-term and short-term timeline for the project • Supervise biostatistician to meet deadlines • Finalize analysis plan and report • Provide suggestions for more advanced statistical methods and data analytic strategies 	<ul style="list-style-type: none"> • Draft statistical analysis and data analytic plan (SAP) with help from HSDS Co-I • Conduct analysis • Draft report
Unfunded research projects*	<ul style="list-style-type: none"> • Not required 	<ul style="list-style-type: none"> • Work independently through the Data Analysis Consulting Services (DACs) hours • Provide guidance on data analysis • Help draft method and results sections for the manuscript

*People who need help with quality improvement and educational research projects should email Julie.Thompson@duke.edu for consultation.

I. For pre-award grant applications

1. A submission of a Statistical and Data Analytic Collaboration Request Form from a PI is expected **9-12 months** prior to the grant application deadline. The PI is expected to invite the HSDS Co-I to attend the **blue** team review meeting before or after the submission of the statistical and data analytic support request form.
2. The HSDS Co-I will work closely with the PI on drafting the study design, data analysis plan, and power and sample size calculations **3-4 months** prior to the grant application deadline. The PI is expected to invite the HSDS Co-I to attend the **pink** team review meeting.
 - a. It is important and typically most effective to involve the HSDS Co-I in designing the study from the beginning to ensure that the study design, sample size, etc. are adequate to meet the project's specific aims.
 - b. Further, the HSDS Co-I can assist with and/or provide randomization or blinding procedures, intervention strategies, timing of procedures and visits as well as other study design issues related to data collection.
3. The HSDS Co-I will work closely with the PI on final edits for the grant proposal **1 month** before the grant application deadline. The PI is expected to invite the HSDS Co-I to attend the **red** team review meeting.

II. For funded research projects:

1. We may adopt a team approach or an individual approach based on the scope of the project, the PI's preference, and the availability of HSDS Core members. The team approach

includes an HSDS Co-I and a biostatistician. The individual approach includes only one HSDS Co-I or biostatistician who will take complete responsibility. The detailed responsibilities are described in Table 1.

2. The scope of work and timeline should comply with the statistical and data analytic plan on the original proposal that was approved by the funding agency. In addition, it is not uncommon to have deviations from the original analysis plan. In these instances, the HSDS Co-I (or biostatistician) and the PI must discuss the deviations from the original analysis plan as well as any secondary analysis the HSDS Co-I (or biostatistician) deems necessary to perform in conjunction with the main analysis. Additionally, the revised final scope of work and timeline needs to be agreed upon between the PI and the HSDS Co-I (or biostatistician).

III. For unfunded research projects:

1. A reservation of a consulting session with a biostatistician from a PI through the HSDS Core's DACS is expected.
2. The level of involvement and timeline on an unfunded research project should be discussed and an agreement reached between the PI and the biostatistician at the start of the consulting session.

B. Responsibilities of a PI from an HSDS Co-I's and biostatistician's point of view

1. Collaboration between a PI and an HSDS Co-I and consultation for a PI by a biostatistician are most effective if good communication practices are established. The PI is responsible for ensuring that the HSDS Co-I and the biostatistician have a clear understanding of the objectives of the project by providing sufficient, relevant information. It is not uncommon for the HSDS Co-I and the biostatistician to be unfamiliar with the terminology and key components related to the PI's content area of research. The PI should provide the HSDS Co-I and the biostatistician sufficient time to become comfortable with the content and nature of the data to be analyzed.
2. The PI is responsible for providing a complete and accurate description of the data collection procedures, eligibility criteria, issues during data collection, and protocol deviations.
3. Any publications or reports produced as a result of the project should acknowledge the contribution made by the HSDS Co-I and the biostatistician, consistent with the level of involvement. It is also appropriate to include the HSDS Co-I and the biostatistician as co-authors on manuscripts on which the HSDS Co-I and the biostatistician put forth significant effort, including writing the analysis section, providing analytic results, and reviewing the manuscript for correctness.

C. Financial consideration (see Table 2)

Table 2. Suggested Efforts Allocation for Research Projects

Type	Team Approach		Individual Approach
	HSDS Co-I	Biostatistician	HSDS Co-I or Biostatistician
Pre-award	Courtesy efforts by the HSDS Co-I as scholarship	NA	NA
Funded	<i>NIH R01 or equivalent</i> Year 1-2: 5% Year 3-5: 5-10% <i>NIH R21 or equivalent</i> Year 1-2: 5%	<i>NIH R01 or equivalent</i> Year 1-2: 5-10% Year 3-5: 15-30% <i>NIH R21 or equivalent</i> Year 1: 5% Year 2: 10-20%	<i>NIH R21 or equivalent</i> Year 1: 5-10% Year 2: 10-20% <i>Foundation Grants (25-100k)</i> 5-10%
Unfunded*	PIs should use DACS for general statistical and data analytic support		

*People who need help with quality improvement and educational research projects should email Julie.Thompson@duke.edu for consultation.

1. Table 2 provides a summary of the different types of projects. Effort allocation variation depends on budget allowance and the scope of work in the grant.
 - a. Funding should be matched to the size, scope, and complexity of the data analysis and study design. Key determinants include the number of primary and derived study variables that will be collected and analyzed, the quality and completeness of the data to be supplied for analysis, and the complexity of the programming necessary to assemble input data and implement descriptive and analytical methods.
 - b. For multi-year projects, effort commitments may vary throughout the study timeline, according to the needs in various phases, including randomization schemes for sampling and experimental assignment (early), the development and implementation of data and safety monitoring plans (during the middle phases of prospective studies), and the implementation of statistics and data analytics and communication of study results (later).
 - c. Any changes due to unforeseeable reasons, such as budget cuts, in the percent of the HSDS Co-I's and/or the biostatistician's efforts made during proposal writing or after research has been funded should be negotiated between the PI and the HSDS Co-I and the biostatistician, and approved by the HSDS Core Director and the Vice Dean for Research.

2. The HSDS Core is supported partially by funding from the Center for Nursing Research. This funding allows the HSDS Core to provide some services without charges, such as statistical and data analysis support by DACS for unfunded research projects.

D. Ethics in statistical and data analysis collaboration and consultation

1. HSDS Co-Is, biostatisticians, and PIs should be guided by professional and scientific ethics, which promotes the integrity of the data analysis and conclusions. HSDS Co-Is and biostatisticians, together with PIs, have the responsibility to ensure data quality, conduct rigorous analysis, and appropriately interpret results.
2. HSDS Co-Is and biostatisticians are to use methodology and data without prejudice or favoritism that are relevant and appropriate to produce valid results.
3. All HSDS Co-Is and biostatisticians are to be held accountable to the ethical guidelines from the American Statistical Association (<http://www.amstat.org/asa/files/pdfs/EthicalGuidelines.pdf>).