

Job Title: Data Analyst

Date: April 10, 2020

Job Code:

Grade:

FLSA Status:

GENERAL SUMMARY/ OVERVIEW STATEMENT: Summarize the nature and level of work performed.

The Pulmonary and Critical Care Medicine is a large group of physician-scientists, epidemiologists, basic scientists, and geneticists with major research initiatives in lung diseases and chronic disease epidemiology. In this position, a highly motivated biostatistician/bioinformatician will be engaged in multiple projects related to disease subtyping and the genomics of COPD, with specific attention to analysis and analytical problems arising with next-generation sequencing assays. The successful candidate will work primarily with Dr. Yohannes Tesfaigzi and collaborators in the following specific areas: 1) application of statistical and machine learning methods for COPD subtyping, 2) analysis and methods development for RNAseq analysis, 3) regulatory network analyses.

PRINCIPAL DUTIES AND RESPONSIBILITIES: Indicate key areas of responsibility, major job duties, special projects and key objectives for this position. These items should be evaluated throughout the year and included in the written annual evaluation.

The candidate is expected to engage with Dr. Tesfaigzi and collaborators to establish priorities for integrative analysis. Perform continuous research on bioinformatic approaches to data integration to ensure that best practices are known and emulated where appropriate; develop understanding and use of emerging "deep learning" strategies. Establish workflows for archiving, preprocessing, and analyzing outputs of diverse next-generation sequencing experiments, including WGS, RNA-seq, and ATAC-seq, with particular attention to emerging statistical methods of quality assessment and improvement. Contribute to methodological and substantive publications on data integration and mechanistic studies in general bioinformatics and protein-protein interaction. 6) Interact closely with systems administrators to ensure that hardware and software environments are appropriate.

QUALIFICATIONS: (MUST be realistic, neither overstated nor understated, and related to the essential functions of the job.)

MSc in Bioinformatics or Biostatistics, with 1+ year experience in biomedical research environments. Acquaintance with concepts of molecular biology, and immunology, or next-generation sequencing technology essential.

SKILLS/ ABILITIES/ COMPETENCIES REQUIRED: (MUST be realistic, neither overstated nor understated, and related to the essential functions of the job.)

Strong command of R programming and package creation. Familiarity with Bioconductor, and with continuous integration methodology for software testing and maintenance. Experience with cloud and cluster computing concepts essential. Must have acquaintance with next-generation sequencing pipeline structures, alignment processes, quality assessment and improvement procedures, coverage and feature count processing, inference on differential abundance. Some experience with machine learning methodology and practice.

HOSPITAL WIDE RESPONSIBILITIES: These are required of all staff, regardless of position. Do not remove these standards.
Works within legal, regulatory, accreditation and ethical practice standards relevant to the position and as established by BWH/Partners; follows safe practices required for the position; complies with appropriate BWH and Partners policies and procedures; fulfills any training required by BWH and/or Partners, as appropriate; brings potential matters of non-compliance to the attention of the supervisor or other appropriate hospital staff.

APPROVAL:

(NAME)
Department Mgr. _____ Title: _____ Date: _____

(NAME)
Other, As Appropriate _____ Title: _____ Date: _____

The above is intended to describe the general contents and requirements of work being performed by people assigned to this classification. It is not intended to be construed as an exhaustive statement of all duties, responsibilities or skills of personnel so classified.