

POSITION DESCRIPTION
RESEARCH ASSOCIATE (CLINICAL DATA SCIENCE RESEARCH GROUP)
GS-9 to GS-11
PD# XXXX

INTRODUCTION

Use your data ninja skills for good! This position is a unique opportunity to join a group of passionate physicians and researchers trying to improve health within the Veterans Health Administration and beyond. We are using the latest statistical, data science, natural language processing, and machine learning tools to extract knowledge from vast amounts of clinical data. We are currently focused on multiple cancer-specific applications, but this position could grow to encompass the full spectrum of Veterans' health. You will be expected to interact and collaborate with investigators at the intersection between medicine and informatics to perform data extraction, curation, and analysis for multiple parallel topics of clinical interest. You will have the opportunity to develop new methodologies and derive new insights into both health and disease, and interact with the broader VA and academic community to disseminate scientific findings with direct applications to healthcare.

MAJOR DUTIES AND RESPONSIBILITIES

Factor 1: Research Assignment

Under the direction of the leadership of the Clinical Data Science Research Group (CDSRG), you will serve as a research associate on a variety of research projects primarily using large existing clinical data sets. Below is a detailed breakdown of expected skills and responsibilities.

Data Mining 35%

You will actively collaborate with investigators about topics of clinical interest, and will design queries to extract and tidy datasets accordingly. You will extract and integrate information from both local and national VHA data sources, as well as other external data sources such as public repositories. You will rely on your knowledge of SQL and other querying strategies, potentially including natural language processing techniques, as well as medical ontologies (such as ICD, CPT, LOINC, SNOMED, RxNorm). All data will be subject to strict HIPAA and institutional constraints as relevant. You will perform the bulk of query development and implementation within the VA Informatics and Computing Infrastructure (VINCI), a secure research ecosystem. You should also be familiar with standard healthcare data sets, quality standards, processes, and outcome measures, industry-standard terms of clinical data, and basic medical terminology. Data mining will also include cleaning and curation of data, including systematic quality controls, and random spot checks to ensure data integrity.

Over time, you will build a set of standard queries to frequently requested data elements and attributes, as well as clinical phenotypes. In doing so, you will have a foundational role in building the capacity and capabilities of the Clinical Data Science Research Group (CDSRG).

Data Analytics 35%

You will utilize numerous approaches to summarize and visualize extracted data to identify both simple and complex relationships among covariates and explanatory variables and outcomes. Specific approaches include but are not limited to Kaplan-Meier estimators, propensity score matching, Markov models, regression analysis, cluster analysis, and multiple machine learning techniques.

Scientific Communication 20%

You will frequently present findings and recommendations to internal and external stakeholders, and will further engage in the academic publication process, with numerous opportunities for authorship commensurate with contributions. You will also be expected to generate interim reports with clear and understandable displays of data, as well as reproducible and verifiable workflows. You will actively participate in CDSRG and related research meetings, with opportunities for detailed discussion of research methodologies and findings.

You will also pursue continuous learning of the VINCI capabilities and CDW structure. This will include live and recorded webinars, following and contributing to an active data listserv, and contributing to VA Pulse, an internal social community of like-minded researchers and data scientists. You will summarize changes and new capabilities to the CDSRG.

Administrative 10%

You will maintain up-to-date knowledge of data repository rules and regulations, and help ensure ongoing access to data warehouse resources and data. This will include continuing review of IRB approvals, progress reports for funding agency review as relevant, and both routine and ad hoc reports using internal and external data sources to complete data requests and track access and utilization. You will also be partially responsible for training additional CDSRG personnel as relevant, and will interact with and support multiple VA-based investigators. You will be expected to maintain all training and certification requirements incumbent on your position.

Knowledge Required by the Position

You will be required to have a working knowledge of SQL for data queries, as well as R and Python for scripting, data cleaning, and analysis. You will need to have a demonstrated ability to perform data cleaning and visualization, and should have experience working with VA databases or other large clinical relational databases. You should have a facility with advanced statistical techniques as well as natural language processing and/or machine learning techniques. You should also have prior experience with scientific publishing.

Factor 2: Supervisory Controls

Although much of your time will be spent working independently, this position will directly report to the leadership of the Clinical Data Science Research Group (CDSRG), consisting of Michael Conlin, Kenneth Gundle, and Reid Thompson. You will not have any administrative responsibilities or oversight over other CDSRG positions.

You will meet at least weekly with the leadership of the CDSRG to review the progress, quality, and barriers to your work.

Factor 3: Guidelines and Originality

Guidelines include Study Protocols and Operational Manuals for active studies. Guidelines also include accepted professional standards and ethics, professional literature, standards of The Joint Commission, agency regulations, procedural manuals, and safety regulations of the treatment facility and equipment. You must exercise judgment in selecting the appropriate health science procedures and apply a thorough understanding of health science procedures and techniques in interpreting the guidelines, determining their applicability to situations not specifically covered, and adapting procedural instructions as necessary and appropriate. You must rely upon your initiative and resourcefulness when deviating from traditional methods or researching trends and patterns to develop new methods, criteria, or proposed new policies. You will operate under the rules and regulations within the VINCI ecosystem, and follow hospital and VA national policies on patient confidentiality, such as HIPAA, and Good Clinical Practice guidelines.

Factor 4: Contribution, Impact, and Stature

You will work as part of a multidisciplinary team within the CDSRG, jointly led by three VA investigators. You will interface with multiple other researchers on projects both at the Portland VA and potentially other collaborating institutions. You will be responsible for periodic research presentations within the context of the CDSRG and will help to organize these meetings, communicate via listserv, VA Pulse, and serve as a resource for VA investigators more broadly. You will extract research-related data from VINCI using SQL queries, perform data analysis using basic and advanced statistical techniques and software, and employ machine learning and other advanced approaches to generate generalizable knowledge from VA data. You will also have opportunities to gain recognition for your contributions to this work by communicating research findings to scientific and clinical communities via publications and presentations.

Other Significant Facts

Customer Service

You will meet the needs of multiple stakeholders while supporting VA missions. You will be expected to consistently communicate and treat individuals (Investigators, their representatives, visitors, and all VA staff) in a courteous, tactful, and respectful manner. You will provide consistent information according to established policies and procedures. You should handle conflict and problems in dealing with all individuals constructively and appropriately.

ADP Security

You will be required to protect printed and electronic files containing sensitive data in accordance with the provisions of the Privacy Act of 1974 and other applicable laws, federal regulations, VA statutes and policy, and VHA policy. You must protect the data from unauthorized release, loss, alteration, or unauthorized deletion. You are expected to follow applicable regulations and instructions regarding access to computerized files, the release of access codes, etc., as set out in the computer access agreement that the employee signs.

Privacy

In the performance of official duties, the employee has regular access to printed and electronic files containing sensitive information, which must be protected under the provisions of the Privacy Act of 1974, and other applicable laws and regulations. The employee is responsible for: 1) protecting that information from unauthorized release or from loss, alteration, or unauthorized deletion; and 2) following applicable regulations and instructions regarding access to computerized files, release of access and instructions regarding access to computerized files, release of access codes, etc. as set out in a computer access agreement which the employee signs.