MICHIGAN MEDICINE
UNIVERSITY OF MICHIGAN

December 9, 2023

Department of Mathematics & Statistics
South Dakota State University
Chicoine Architecture, Mathematics and Engineering Hall
Box 2225
Brookings, SD 57007

Dear Search Committee Members:

I am writing to apply for the position of Assistant Professor of Statistics.

I am a statistical geneticist and physician with expertise in biostatistics and statistical computing applied to biobank-scale genotype and phenotype data. This opportunity would support growth of my research programs and continued development of my skills as a team science-based research mentor. Together, we will make great advances in biomedical research and contribute to society, medicine, and public health.

Research training and plans

I am applying for this job because I want to develop world-class research programs in statistical genetics and evidence-based biostatistics teaching. My extensive training has prepared me to excel in these complementary areas. The long-term goal of my research program in statistical genetics is to develop methods and software tools that promote genetics-informed decision-making in clinical medicine and public health. My first steps towards this goal address shortcomings in current polygenic score methods, which include the failure to model SNP-SNP and SNP-environment interactions, the limited ability to incorporate SNP functional annotations, and the absence of methods to quantify uncertainty in polygenic score estimates (for many widely used construction methods).

When we overcome these shortcomings, we will be able to use polygenic scores to improve clinical medicine and public health. For example, our new polygenic score methods will accurately identify individuals at high risk of coronary artery disease. Previous research demonstrates that coronary artery disease polygenic scores complement traditional clinical risk factors, such as diabetes, sex, age, and smoking behavior. Importantly, coronary artery disease polygenic scores provide risk information that is not captured by clinical risk factors, such as those identified in the Framingham Heart Study; thus, risk assessments that use both clinical risk factors and polygenic scores result in more accurate risk predictors. With the availability of low-cost preventions such as statin therapy, our combined risk assessments will improve clinical decision-making and, ultimately, lead to better public health outcomes.

Our complementary research program in evidence-based biostatistics teaching will draw on my experience as a statistics instructor and my years of training in evidence-based teaching,

course planning, and curriculum design. We will develop and assess teaching and mentoring practices that promote student understanding and encourage, welcome, and support students from all backgrounds.

My Ph.D. from the University of Wisconsin-Madison is in statistics, and I wrote a thesis in statistical methods for multivariate mapping of quantitative trait loci in model organism. My recent work at the University of Michigan involves developing statistical methods to construct prediction intervals for polygenic scores with tools from conformal inference. While many scientists report polygenic scores for diverse traits in biobank studies, few account for the uncertainty in these point estimates. My project rectifies this deficiency. My method is both scalable and easy to implement. Additionally, my method is agnostic to the polygenic score construction method; it works for any such method, including clumping and thresholding, DBSLMM, ldpred2, and lassosum. I have examined my method's performance in large-scale simulation studies, and I've applied my method to UK Biobank traits with more than 300,000 study subjects.

Teaching experience and plans

I am determined to develop into a world-class teacher. To progress towards this goal, I taught statistics courses at the undergraduate and graduate levels. Specifically, I served as lecturer for introductory statistics courses seven times, and I served as a teaching assistant for graduate courses in statistics and biostatistics. I also contributed to the redesign of a collection of introductory statistics courses for graduate and undergraduate students.

I am excited by the diverse collection of courses in your department. My training and teaching experiences have prepared me to teach any courses in the graduate and undergraduate curricula. Additionally, I am eager to develop new courses that serve additional student needs. For example, I would like to develop a course entitled "How to teach statistics". My experience with curriculum design and teaching of a similar course has prepared me to excel in this role. Additionally, I would be delighted to design and teach a course on any of my research interests, including statistical genetics, reproducibility for statistics research, and causal inference for observational studies.

Experience in and plans for promoting diversity, equity, inclusion, and accessibility in science

I prioritize the creation of supportive and welcoming communities for all, especially those from populations that have suffered from exclusionary practices in science careers. Representation matters. I've mentored students from underrepresented groups in statistics research, and I've designed and twice facilitated a one-semester graduate course on teaching for a diverse student body. Science advances when many perspectives are present and supported.

Many of the challenges in human genetics research involve inclusion of subjects from underrepresented groups. For example, white European-ancestry subjects are overrepresented in many studies, which ultimately deprives people with other ancestries of the benefits of our scientific

research. We will address this problem by developing statistical genetics methods that ensure that our research applies to subjects from diverse populations.

Thank you for considering my application. I look forward to hearing from you.

Sincerely,

Frederick J. Boehm, III