Prospectus

Adapting a Bioinformatics Approach for Use in the Discovery of Novel Endocrine Interactions in Humans
(Technical Topic)

Bioinformatics Approaches as Actants Within Clinical and Research Networks
(STS Topic)

By

Jonathan Blichar

Fall, 2020

Technical Project Team Members:
Felipe Barraza
Emmanuel Edu

On my honor as a University student, I have neither given nor received unauthorized aid on this assignment as defined by the Honor Guidelines for Thesis-Related Assignments.

Signed: _____Jonathan Blichar_____ Date ___11/1/2020___

Approved: _____ Date _____
Rosalyn W. Berne, Associate Professor, Department of Engineering and Society

Approved:

_____ Date _____11/9/2020_____
Warren Anderson, Postdoctoral Fellow, Center for Public Health Genomics

I. Introduction

Bioinformatics as a subfield of life sciences has become an increasingly important subject of study and has undergone significant growth in recent decades with the advent of powerful computers with specialized software that can be applied by researchers for a plethora of medical purposes. This includes the use of large cohort biological data for simulating and quantifying DNA, protein, and RNA interactions (Reches et al, 2019; Iqbal et al., 20202). Inter-organ communication, and consequently homeostasis, is maintained via the secretion and uptake of endocrine factors by central and peripheral organs and is controlled primarily by the functionality of the endocrine system. A breakdown in efficiency of this system will often result in endocrine and metabolic disorders that are present worldwide. There is a high a prevalence and incidence of metabolic and endocrine disorders; common examples include cardiovascular diseases (CVDs), diabetes mellitus, and obesity. The list of endocrine disorders that afflict the population is extensive, however, and it is estimated that at least 5 percent of adults in the United States suffer from one or more endocrine or metabolic disorders (Golden et al., 2009). The goal of our technical project is to analyze existing human transcription data to identify endocrine interactions among organ pairs. This information could provide new knowledge on how some endocrine disorders form and potentially lead to novel therapeutics to combat these diseases.

Related to the technical project, understanding how bioinformatics approaches fit into the networks of researchers and clinicians who could utilize these technological systems could provide useful information for their widespread implementation. Enhanced treatment methods and more accurate information of certain disorders that could be introduced as a result of wide scale use of bioinformatics in the medical fields could benefit clinicians, and more importantly, the patients suffering from various ailments. Bioinformatics can apply to clinical areas that go

beyond just the endocrine system, including clinical genomics, genomic medicine, and pharmacogenomics (Bellazzi et al., 2012). Bioinformatics use in medical research and clinical application is growing rapidly, and so this STS research project seeks to analyze how it fits into and interacts with the actors in these healthcare networks to draw conclusions on its significant benefits and limitations.

II. Technical Topic

Complex multicellular organisms utilize protein signaling between organs to maintain physiologic homeostasis that allows the organism to grow within and react to their environment. Advancement in technologies such as proteomics and RNA sequencing have allowed for the discovery of secreted proteins responsible for the maintenance of homeostasis. Our group's technical project will complete two primary goals that will be met through the adoption and adaptation of a previous study's bioinformatics approach for the analysis of novel endocrine interactions in mice (Seldin et al., 2018).

Seldin et al. (2018) created a bioinformatics approach that utilizes natural variations in transcript levels of genes within a population to identify endocrine pathways and analyze their function. The authors termed this methodology the Quantitative Endocrine Network Interaction Estimation (QENIE). Individuals express differing levels of transcription for most genes. Based on inter-individual differences in transcription, the authors were able to analyze expression of an endocrine factor in a sender organ and correlate it with the responder gene transcript pathways within the recipient organ. Secreted endocrine factors could then be identified by investigating the organ pairs that have coordinated levels of expression and filtering for the sender organ. The researchers were able to identify and support the function of established endocrine interactions that have been the focus of previous scientific works. They were also able to uncover potential

novel endocrine functions for three secreted peptides using their bioinformatics approach. One such novel endocrine function discovery was how elevated expression of LCN5 enhanced mitochondrial function and glucose metabolism within mice muscles. The authors of this article have made the basic code for QENIE available to the public on GitHub. Their study however, failed to apply QENIE to human data and conglomerate their work into a function software package for ease of access and use. Our group's first goal is to alter the R script made available to accommodate human transcript data that is available via the GTEx Consortium which provides transcript information for various organs. Using the human transcript data, we will be able to replicate the analysis of Seldin et al. (2018) to discover and support both novel and established endocrine interactions between human organ pairs.

The second goal of our technical research is to analyze potential differences between endocrine interactions within males and females. It is well established that metabolic homeostasis is regulated differently between the sexes. It is from these differences in homeostasis and adipose tissue interactions that variation in the development of obesity and diabetes between males and females form (Mauvais-Jarvis, 2015). Endocrine factors may be responsible for differences in the location of adipose generation between sexes (Anderson et al., 2020). The differences in adipogenesis between sexes has also been linked to cardiometabolic diseases (Lumish Heidi S. et al., 2020). Our group proposes that by analyzing the endocrine interactions and their pathways between the two sexes may allow for the discovery of other traits and disorders that may be genetically linked to a particular sex. By doing so, novel therapeutics or preventative measures can be discovered to aid people who may be more susceptible to a given disorder based on sex.

III. STS Topic

Bioinformatics applied directly to biomedical research and clinical medicine combating human diseases has been termed clinical bioinformatics or simply medical informatics (Bellazzi et al, 2012; Cheng, p. 201). The STS topic seeks to establish an understanding of bioinformatics approaches' role in healthcare, from both research and clinical settings. An analysis of the interactions between bioinformatics tools and other actants within the established networks of medical research and clinical fields can give significant insight into the potential benefits of widespread implementation of such systems, but also any drawbacks they may pose and how their adoption may be limited. The analysis will be conducted from the STS perspective of Actor-Network Theory (ANT). Despite contention that questions the applicability of this framework, it remains a useful approach for analyzing the role of technology in complex systems containing multiple actors (Cresswell et al., 2010). Since one of ANT's primary features is to focus on inanimate objects and their effect on social processes, this is an optimal framework to apply to bioinformatics for conceptualizing these approaches' effects on the other actors within healthcare (Cresswell et al., 2010).

The ANT framework will identify actors within the networks of interest that are most influenced by the introduction of bioinformatics tools into commonplace usage. Researchers, clinicians, computer systems, and patients are the physical actors who will interact directly or indirectly with information provided by the bioinformatics systems. Large-scale biological data and other bioinformatics software approaches are intangible actors that also play significant roles in the analysis of the implementation of the medical informatics tools. Understanding the paired interaction of bioinformatics systems with each of these actors is a major goal of the research question.

Bioinformatics tools and methods have the potential to influence the work done by researchers in a vast group of biomedical fields that includes DNA and protein sequencing as well as deriving protein structure and function (Gauthier et al., 2019). There are extensive publications on bioinformatics approaches for investigating medical related issues. The article utilized in the technical project is one such example, which uses broad transcript data from various organs to draw conclusions on endocrine interactions that provide many implications. Other research is refined and honed to a specific issue, like that conducted in one study where the authors examined the pathobiology of intracerebral hemorrhage using proteomics to discover novel molecular targets that could promote new treatment options (Desari et al., 2020). Broad bioinformatics research can typically lead to further downstream studies while specific research may have more direct effects on work in clinical settings. The number of bioinformatics approaches and the speed at which these methods are altered and advanced can result in difficulties for researchers trying to maintain the skills and knowledge to utilize bioinformatics software (Gauthier et al., 2019).

Similar to the research aspect of bioinformatics, clinicians in a considerable quantity of medical fields may be influenced by the adoption of clinical bioinformatics. Patients suffering from disorders in these medical disciplines also stand to benefit. Studying and enhancing the delivery and targeting of drugs, sequencing large populations of DNA and proteins to assist in individual sequence analysis, and screening for biomarkers that indicate disease susceptibility are just a few of the ways in which bioinformatics may reshape clinical work (Cheng, p. 204-207). As large-scale human DNA data becomes available for sequencing by increasingly precise bioinformatics methods, clinicians may be able to incorporate individuals genotype into their diagnoses (Reches et al., 2019). If enough data can be gathered from healthcare workers,

phenotype-genotype correlations could be found and enhance diagnosis and treatment as well (Bellazzi et al., 2012). Clinical work done to diagnose and treat cancer is another field that could experience significant change with the implementation of bioinformatics approaches. Bioinformatics would likely enhance discovery of biomarkers that signal a susceptibility of patients to certain types of cancers, but also enhance the efficacy of precision medicines for combating cancer cells (Wu et al., 2012). Lastly, one study attempted to analyze the effects of sequencing on downstream healthcare using genome sequencing of newborns coupled with their treatment related health records (Mackay et al., 2020).

The last group of actors that will be investigated during the STS research will be the intangible actors, including biological data and individual bioinformatics approaches. Despite their importance, it seems that a majority of the limitations of medical informatics is sourced at least partly from its interactions with these intangible actors. At present, there is little to no standardization amongst databases containing the large cohort biological data that is utilized by researchers and clinicians in bioinformatics. Variations in how biological data is stored and presented to researchers seeking to use it results in difficulties that slow down the research processes (Stein, 2002). Creating a common language for communication for biological data is essential for its widespread adoption in healthcare fields (Kouskoumvekaki et al., 2014). Similarly, the development of different bioinformatics approaches that complete similar tasks can create confusion amongst researchers and make communicating and replicating results difficult (Stein, 2002).

The STS Research will be efficiently conducted by analyzing reviews and studies that utilize bioinformatics approaches to make medical observations. This will aid in screening for individual fields of medicine that may promote the most growth of bioinformatics tools in their

individual discipline. By selecting the medical areas that will be most influenced by bioinformatics adoption, better insight into the role of medical informatics as it pertains to both clinical and research networks can be acquired. This includes both how bioinformatics can benefit the actors of both networks, but also how it limits these fields. I will attempt to qualitatively analyze bioinformatics' role within healthcare as a whole and interpret differences in its clinical and research roles to better understand the implementation process of these methods and systems.

IV. Timeline and Expected Outcomes

The technical project will be split into two major deliverables. The first is an R software package that will be able to perform the data manipulation and analysis. The second will be a manuscript of our methodology and findings based on our project's analysis of novel endocrine interactions and a comparison of these interactions between male and female cohorts.

The STS research is anticipated to provide a novel understanding of bioinformatics approaches as a growing tool for use in the networks of clinical research and medical application as it applies to a vast number of patient illnesses and disorders. Significant acumen may be gleaned from the analysis of medical informatics through the ANT framework lens.

The STS research project, technical project, and the technical manuscript will be completed over the course of the 2020-2021 academic year. All three projects will be completed by the end of April 2021.

Word Count: 1996

References

Anderson, W. D., Soh, J. Y., Innis, S. E., Dimanche, A., Ma, L., Langefeld, C. D., Comeau, M. E., Das, S. K., Schadt, E. E., Björkegren, J. L. M., & Civelek, M. (2020). Sex differences in human adipose tissue gene expression and genetic regulation involve adipogenesis. *Genome Research*, *30*(10), 1379–1392. https://doi.org/10.1101/gr.264614.120

Bellazzi, R., Masseroli, M., Murphy, S., Shabo, A., & Romano, P. (2012). Clinical Bioinformatics: Challenges and opportunities. *BMC Bioinformatics*, *13*(Suppl 14), S1. https://doi.org/10.1186/1471-2105-13-S14-S1

Chang, P. (2005). Clinical bioinformatics. *Chang Gung Medical Journal*, *28*, 201–211.

Cresswell, K. M., Worth, A., & Sheikh, A. (2010). Actor-Network Theory and its role in understanding the implementation of information technology developments in healthcare. *BMC Medical Informatics and Decision Making*, *10*(1), 67. https://doi.org/10.1186/1472-6947-10-67

Dasari, R., Zhi, W., Bonsack, F., & Sukumari-Ramesh, S. (2020). A Combined Proteomics and Bioinformatics Approach Reveals Novel Signaling Pathways and Molecular Targets After Intracerebral Hemorrhage. *Journal of Molecular Neuroscience*, *70*(8), 1186–1197. https://doi.org/10.1007/s12031-020-01526-7

Gauthier, J., Vincent, A. T., Charette, S. J., & Derome, N. (2019). A brief history of bioinformatics. *Briefings in Bioinformatics*, *20*(6), 1981–1996. https://doi.org/10.1093/bib/bby063

Golden, S. H., Robinson, K. A., Saldanha, I., Anton, B., & Ladenson, P. W. (2009). Clinical review: Prevalence and incidence of endocrine and metabolic disorders in the United States:

a comprehensive review. *The Journal of Clinical Endocrinology and Metabolism*, *94*(6), 1853–1878. https://doi.org/10.1210/jc.2008-2291

Iqbal, M. N., Rasheed, M. A., Awais, M., Chammam, W., Kanwal, S., Khan, S. U., Saddick, S., & Tlili, I. (2020). BMT: Bioinformatics mini toolbox for comprehensive DNA and protein analysis. *Genomics*, *112*(6), 4561–4566. https://doi.org/10.1016/j.ygeno.2020.08.010

Kouskoumvekaki, I., Shublaq, N., & Brunak, S. (2014). Facilitating the use of large-scale biological data and tools in the era of translational bioinformatics. *Briefings in Bioinformatics*, *15*(6), 942–952. https://doi.org/10.1093/bib/bbt055

Lumish Heidi S., O'Reilly Marcella, & Reilly Muredach P. (2020). Sex Differences in Genomic Drivers of Adipose Distribution and Related Cardiometabolic Disorders. *Arteriosclerosis, Thrombosis, and Vascular Biology*, *40*(1), 45–60. https://doi.org/10.1161/ATVBAHA.119.313154

Mackay, Z. P., Dukhovny, D., Phillips, K. A., Beggs, A. H., Green, R. C., Parad, R. B., Christensen, K. D., Agrawal, P. B., Ceyhan-Birsoy, O., Fayer, S., Frankel, L. A., Genetti, C. A., Gutierrez, A. M., Harden, M., Holm, I. A., Krier, J. B., Lebo, M. S., Machini, K., McGuire, A. L., … Yu, T. W. (2020). Quantifying Downstream Healthcare Utilization in Studies of Genomic Testing. *Value in Health*, *23*(5), 559–565. https://doi.org/10.1016/j.jval.2020.01.017

Mauvais-Jarvis, F. (2015). Sex differences in metabolic homeostasis, diabetes, and obesity. *Biology of Sex Differences*, *6*(1), 14. https://doi.org/10.1186/s13293-015-0033-y

Reches, A., Weiss, K., Bazak, L., Baris Feldman, H., & Maya, I. (2019). From phenotyping to genotyping—Bioinformatics for the busy clinician. *European Journal of Medical Genetics*, *62*(8), 103689. https://doi.org/10.1016/j.ejmg.2019.103689

Seldin, M. M., Koplev, S., Rajbhandari, P., Vergnes, L., Rosenberg, G. M., Meng, Y., Pan, C., Phuong, T. M. N., Gharakhanian, R., Che, N., Mäkinen, S., Shih, D. M., Civelek, M., Parks, B. W., Kim, E. D., Norheim, F., Chella Krishnan, K., Hasin-Brumshtein, Y., Mehrabian, M., … Lusis, A. J. (2018). A Strategy for Discovery of Endocrine Interactions with Application to Whole-Body Metabolism. *Cell Metabolism*, *27*(5), 1138-1155.e6. https://doi.org/10.1016/j.cmet.2018.03.015

Stein, L. (2002). Creating a bioinformatics nation. *Nature*, *417*(6885), 119–120. https://doi.org/10.1038/417119a

Wu, D., Rice, C. M., & Wang, X. (2012). Cancer bioinformatics: A new approach to systems clinical medicine. *BMC Bioinformatics*, *13*, 71. https://doi.org/10.1186/1471-2105-13-71