QBIO490: Final Project Research Proposal–Ovarian Cancer

Rebecca Aviles Barahona, Jeanne Michelle Revilla, Andrea Velazquez

The cancer we are researching is ovarian cancer. Ovarian cancer affects women in which tumors line the ovaries– a reproductive organ that produces eggs throughout a women's lifespan. This cancer has a 1 % causation rate, and a lifetime risk of 1 in 78 people. This cancer occurs more frequently as you get order–typically after menopause occurs. There are knowledge gaps in longitudinal data as the cancer is always found later. This is due to improper ovarian and cancer screenings as it is more likely to occur later on. As a result, this makes the subtypes of the ovarian cancer and the terms to be generally defined; however, the parameters within each subtype are unclear due to not a lack of research in how the cancer progresses.

Hypothesis: Can we determine how many significant subtypes of ovarian cancer are? How do their molecular signatures differ (specifically in terms of gene expression or clinical features)?

One analysis we are interested in using is machine learning to identify the different subtypes. We would feed the algorithm clinical, genomic, and transcriptomic data. We may do this analysis multiple times, with different clustering methods. While previous analyses exist, where ovarian cancer subtypes are designated and samples are classified based on quantitative methods, they don't provide the reader with enough information on how the classification of the specimens studied occurred. Additionally, they fail to analyze the effectiveness of their classification methods. In our computational analysis, we aim to accomplish the same objective, but in a way that allows us to analyze the model we will use to classify the subtypes through the use of statistical methods and tools to assess the bias and the variance of our model. With this, we will eliminate the possibility of having a complex machine learning model that is overfitted and produces more subtypes than actually exist, or inversely, a machine learning model that is underfitted and yields us with fewer subtypes than those existing. We also hope to use the pre-existing classifications to guide what genes/clinical data we look at to determine subtype; though it is important to emphasize that there are limitations concerning the known biomarkers due to the gap in early detection of ovarian cancer.

Identifying ovarian cancer subtypes based on multi-omic analysis could provide us with insight on the significance of each subtype and their implications on the broader biological processess affected; thereby identifying the extent of the impact the ovarian cancer has on a patient. With this, we could better understand the meaning of the biomarkers identified and their usefulness when it comes to prognosis and therapeutic targets and outcomes. Consequently allowing us to take advantage of precision medicine and provide patients with treatment options that would better suit them– targetting the cancer in specific ways that are bound to progress positively at an earlier stage, improving the chances of survival for patients.

References

Cleveland Clinic Medical. “Ovarian Cancer: Symptoms, Diagnosis & Treatment.” *Cleveland Clinic*, my.clevelandclinic.org/health/diseases/4447-ovarian-cancer.

Helland, Å., Anglesio, M. S., George, J., Cowin, P. A., Johnstone, C. N., House, C. M., Sheppard, K. E., Etemadmoghadam, D., Melnyk, N., Rustgi, A. K., Phillips, W. A., Johnsen, H., Holm, R., Kristensen, G. B., Birrer, M. J., Australian Ovarian Cancer Study Group, Pearson, R. B., Børresen-Dale, A. L., Huntsman, D. G., deFazio, A., … Bowtell, D. D. (2011). Deregulation of MYCN, LIN28B and LET7 in a molecular subtype of aggressive high-grade serous ovarian cancers. *PloS one*, *6*(4), e18064. https://doi.org/10.1371/journal.pone.0018064

Wang, Y.A., Neff, R., Song, W.-m., Zhou, X., Vatansever, S., Walsh, M.J., Chen, S.-H. and Zhang, B. (2023), Multi-omics-based analysis of high grade serous ovarian cancer subtypes reveals distinct molecular processes linked to patient prognosis. FEBS Open Bio, 13: 617-637. https://doi.org/10.1002/2211-5463.13553