**Conformal Prediction in Clinical Sciences:**

**Preliminary Analysis in Breast Cancer Risk Prediction**

Julio C Facelli and Ram Gouripeddi

**Introduction:** The use of machine learning (ML) and artificial intelligence (AI) applications in medicine has attracted a great deal of attention in the medical literature [1-7]. Several issues arise from this body of literature. The most pertinent to this project is the lack of methods enabling uncertainty quantification (UQ), generalizability, and reproducibility of clinical machine learning. The current state of the art for evaluating the performance of clinical predictive models is to provide values that measure overall global performance [5 8 9]. However, these global properties do not provide any information about the confidence in individual predictions. It is noteworthy that the model prediction for different individuals will have very different intervals of confidence depending on the representation of similar individuals in the training set. This is particularly concerning when using predictive analytics for individuals of underserved populations that systematically are excluded from the training sets used in parametrizing predictive models [10 11]. A promising approach to provide uncertainty for each prediction is Conformal Prediction (CP) [12 13]. A recent paper [14] has shown the underutilization of CP in medical clinical applications and highlighted the potential of the method to quantitatively expand the use of AI tools to unrepresented populations.

***Aim****: This work will demonstrate to the use of CP for AI methods used to predict breast cancer risk.*

**Methods:** For the data sets and methods targeted in this project will be those review by Ming et al. [15] For the selected ML prediction methods, we will select three CP methods, which include Moridian Predictors, Inductive Confidence Machine, and Non-Conformity Measurements (8), to implement the CP for the models discussed above. With this, we will have multiple combinations of the ML classier-CP method that we will apply to the following outcome predictions (common in the AI literature in medicine):

This is the first time that a comprehensive program is being put together to explore systematically the potential impact of using CP for quantifying the **uncertainty of individual predictions** given for ML predictive clinical models. If successful, this demonstration project will provide a critical avenue to move forward the application of ML methods to medicine and addressing emerging health disparities when using AI in clinical practice.

Students interested in this practicum should contact Dr. Facelli ([julio.facelli@utah.edu](mailto:julio.facelli@utah.edu)). Programming experience in R or Python required.

**References**

1. Shah NH, Milstein A, Bagley P, Steven C. Making Machine Learning Models Clinically Useful. JAMA 2019;**322**(14):1351-52 doi: 10.1001/jama.2019.10306[published Online First: Epub Date]|.

2. Beam AL, Manrai AK, Ghassemi M. Challenges to the Reproducibility of Machine Learning Models in Health Care. JAMA 2020;**323**(4):305-06 doi: 10.1001/jama.2019.20866[published Online First: Epub Date]|.

3. Liu VX, Bates DW, Wiens J, Shah NH. The number needed to benefit: estimating the value of predictive analytics in healthcare. Journal of the American Medical Informatics Association 2019;**26**(12):1655-59 doi: 10.1093/jamia/ocz088[published Online First: Epub Date]|.

4. Stead WW. Clinical Implications and Challenges of Artificial Intelligence and Deep Learning. JAMA 2018;**320**(11):1107-08 doi: 10.1001/jama.2018.11029[published Online First: Epub Date]|.

5. Van Calster B, Wynants L, Timmerman D, Steyerberg EW, Collins GS. Predictive analytics in health care: how can we know it works? Journal of the American Medical Informatics Association 2019;**26**(12):1651-54 doi: 10.1093/jamia/ocz130[published Online First: Epub Date]|.

6. Frey LJ, Bernstam EV, Denny JC. Precision medicine informatics. Journal of the American Medical Informatics Association 2016;**23**(4):668-70 doi: 10.1093/jamia/ocw053[published Online First: Epub Date]|.

7. Hunter DJ. Uncertainty in the Era of Precision Medicine. New England Journal of Medicine 2016;**375**(8):711-13 doi: 10.1056/NEJMp1608282[published Online First: Epub Date]|.

8. Steyerberg EW, Vickers AJ, Cook NR, et al. Assessing the Performance of Prediction Models: A Framework for Traditional and Novel Measures. Epidemiology 2010;**21**(1):128-38 doi: 10.1097/EDE.0b013e3181c30fb2[published Online First: Epub Date]|.

9. Rajkomar A, Dean J, Kohane I. Machine Learning in Medicine. New England Journal of Medicine 2019;**380**(14):1347-58 doi: 10.1056/NEJMra1814259[published Online First: Epub Date]|.

10. Vyas DA, Eisenstein LG, Jones DS. Hidden in Plain Sight — Reconsidering the Use of Race Correction in Clinical Algorithms. New England Journal of Medicine 2020;**383**(9):874-82 doi: 10.1056/NEJMms2004740[published Online First: Epub Date]|.

11. Cirillo D, Catuara-Solarz S, Morey C, et al. Sex and gender differences and biases in artificial intelligence for biomedicine and healthcare. npj Digital Medicine 2020;**3**(1):81 doi: 10.1038/s41746-020-0288-5[published Online First: Epub Date]|.

12. Shafer G, Vovk V. A Tutorial on Conformal Prediction. J. Mach. Learn. Res. 2008;**9**:371-421

13. Conformal Prediction for Reliable Machine Learning. In: Balasubramanian VN, Ho S-S, Vovk V, eds. Conformal Prediction for Reliable Machine Learning. Boston: Morgan Kaufmann, 2014:i.

14. Vazquez J, Facelli JC. Conformal Prediction in Clinical Medical Sciences. Journal of Healthcare Informatics Research 2022 doi: 10.1007/s41666-021-00113-8[published Online First: Epub Date]|.

15. Ming C, Viassolo V, Probst-Hensch N, Chappuis PO, Dinov ID, Katapodi MC. Machine learning techniques for personalized breast cancer risk prediction: comparison with the BCRAT and BOADICEA models. Breast Cancer Research 2019;**21**(1):75 doi: 10.1186/s13058-019-1158-4[published Online First: Epub Date]|.