Title: Optimizing Biomarker Models for Biologically-Heterogeneous Cancers: A Nested Model Approach for Lung Cancer Diagnosis

Authors: Laurel Jackson, MS[1]; Michael N. Kammer, PhD[2]; Palina Woodhouse, MD[2]; Caroline M. Godfrey, MD, MPH[2]; Sanja Antic[2]; Yong Zou, MS[2]; **Susan Gawel, PhD[1] (presenting author);** Fabien Maldonado, MD, MSc[2]; Eric L. Grogan, MD, MPH[2, 3], Stephen A. Deppen, PhD[2, 3*]; Gerard J. Davis, PhD[1*].

[1]Abbott Diagnostics Division, Abbott Park, IL
[2]Vanderbilt University Medical Center, Nashville, TN
[3]Tennessee Valley Healthcare System, Nashville, TN
*Gerard J. Davis and Stephen A. Deppen are joint senior authors.

Background: The heterogeneous biology of cancer subtypes, especially in lung cancer, poses significant challenges for biomarker and algorithm development in early detection of cancer. Standard model building techniques often fall short in accurately identifying various subtypes due to their diverse biological characteristics. This study explores a nested biomarker model to address this issue, aiming to improve early detection across different lung cancer subtypes.

Methods: Conducted using a prospective specimen collection and retrospective evaluation design, the study included 337 patients from two clinical sites. Blood biomarkers (CA-125, CEA, SCC, HE4, ProGRP, NSE, CYFRA 21-1, and Ferritin) were analyzed on Abbott's immunoassay & chemistry systems, and various statistical methods, including logistic regression and feature selection, were employed to develop a nested model. This model was designed to account for the biological heterogeneity of cancer subtypes, comparing it against traditional logistic regression models.

Results: The patient cohort included a range of malignant and benign nodules and included different cancer subtypes reflecting lung cancer heterogeneity. The novel nested modeling approach of interest performed well in the training set with an AUC of 77.6 (95% CI 72.2, 83.0) and 77.3 (95% CI 65.8, 88.9) in testing, and incorporated different key biomarkers across different cancer subtypes, underscoring the biological diversity of these subtypes. The nested subtype versus benign model had the best performance in the training set overall and had a particular advantage for the small cell subtype.

Conclusion: This proof-of-concept study highlights the challenges cancer heterogeneity present for biomarker development and the potential for nested biomarker models to improve early cancer detection. By focusing on subtype-specific biomarkers, this approach may assist in developing more effective Multi-Cancer Early Detection (MCED) strategies. Future research with larger cohorts is essential to validate this approach, emphasizing the need for models that can navigate the intricate biological diversity of cancer subtypes.