Metrological Issues Facing Validation of Multi-parametric Imaging Biomarkers

Nancy Obuchowski, PhD

Cleveland Clinic Foundation

Four Multi-parametric Use Cases:

Use Case	Description	Example
Multi-dimension descriptor	Panel of individual but related QIBs each of importance	CTA atherosclerosis biomarker panel
Phenotype classification	Multiple QIBs used in a decision tool to classify cases into phenotypes	CT biomarkers used in model to classify liver lesions (e.g. cyst, hemangioma, mets)
Risk prediction	Multiple QIBs used in a decision tool to predict patient outcome or risk	CT lung biomarkers used in model to predict progression of disease
Radiomics	Computer extraction of potentially large numbers of derived metrics for prediction	Development of a radiomics signature that predicts immunotherapy response

Example (use case #3)

- Consider a prognostic tool that uses quantitative CT imaging
  - lung volume, lung density, atherosclerosis biomarkers, nodule(s) volume
- We need to establish the performance of the model
  - Cross-validation and bootstrap validation are great for finalizing model (i.e. selecting variables, tuning model) but not for validation of the locked-down tool
- What are the issues in (external) validation of such a tool?

Issues for (External) Validation of Locked-Down Tool

- Availability of independent (spatially and temporally), generalizable validation datasets Altman and Royston, 2000
- Missing data
  - Imputation ok but can't be based on locked-down tool and never for the outcome
- Less than perfect reproducibility of biomarker measurements
  - It helps to vet different scanners/scanning protocols in training dataset
- "Batch effects"
  - Differences in distribution of QIBs between sites can confound results
- Timing of imaging
  - Are biomarker measurements acquired from different scans? Different timepoints?
  - Are scanning parameters altered depending on other biomarkers' values?