

Differentiating high- and low-grade serous ovarian carcinoma using radiomics: a pilot study in the MITO group

Francesca Ciccarone¹, Giulia Zinicola¹, Domenica Lorusso^{2,3}, Huong Elena Tran¹, Giulia Baldassari¹, Edda Boccia¹, Federica Pozzati¹, Rosanna Mancari⁴, Robert Fruscio⁵, Fulvio Borella⁶, Francesca Buonomo⁷, Camilla Nero¹, Francesca Moro¹, Tina Pasciuto¹, Anna Fagotti¹, Giovanni Scambia¹, Antonia Carla Testa¹, Gabriella Ferrandina¹

¹Fondazione Policlinico Universitario A. Gemelli IRCCS, Rome, Italy
²Department of Biomedical Sciences, Humanitas University, Milan, Italy
³Humanitas San Pio X, Milan, Italy
⁴Gynecologic Oncology Unit, IRCCS Regina Elena National Cancer Institute, Rome, Italy
⁵Clinic of Obstetrics and Gynecology, Department of Medicine and Surgery, University of Milano-Bicocca, Monza, Italy
⁶Division of Gynecology and Obstetrics 1, Department of Surgical Sciences, City of Health and Science University Hospital, University of Turin, Turin, Italy
⁷Institute for Maternal and Child Health, I.R.C.C.S. "Burlo Garofolo", 34137 Trieste, Italy

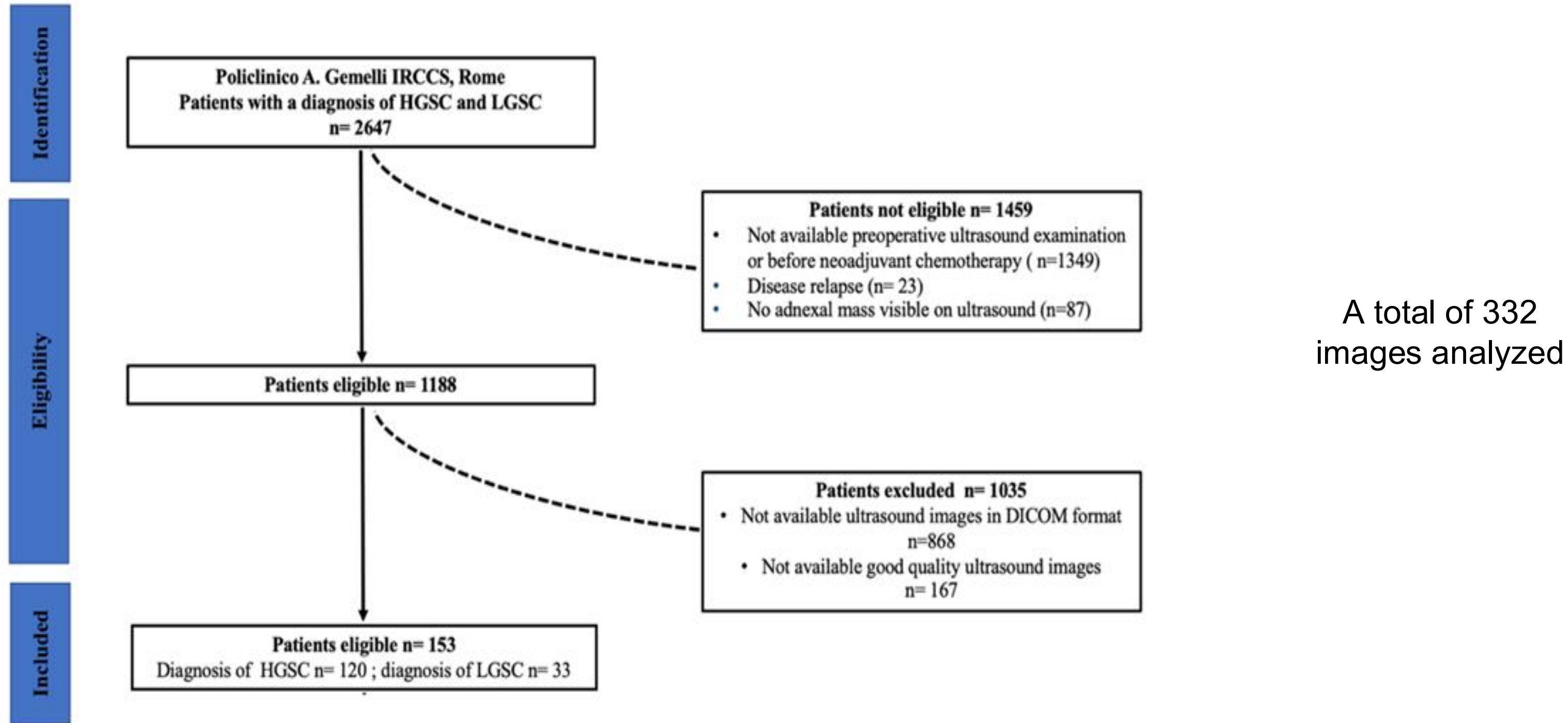
Objective

- ✓ To identify ultrasound-based radiomic features capable of distinguishing between ovarian high-grade serous carcinomas (HGSCs) and low-grade serous carcinomas (LGSCs)
- ✓ To develop machine learning models that include radiomics features to discriminate between the two entities.

Methods

Single-center, analytical, observational, retrospective, pilot study.

All patients with a histological diagnosis of HGSC or LGSC and available preoperative digital ultrasound images were identified from the gynecologic oncology ultrasound unit database between January 2014 and June 2022.



The extracted radiomic features belonged to two different families: intensity-based statistical features and textural features.

Features selection was performed via univariate analysis with the Wilcoxon-Mann-Whitney statistical test, followed by Boruta algorithm, Pearson correlation and Akaike Information Criterion stepwise regression.

Selected features were used to train logistic regression models to distinguish between HGSC and LGSC:

- a Radiomic model based solely on radiomic features
- a Clinical-Radiomic model combining radiomic features with patient age.

For comparison, two additional models were constructed: a Radiomic-Ultrasound model incorporating radiomic features and significant ultrasound variables, and an Ultrasound model based exclusively on significant ultrasound variables.

The area under the curve (AUC) of the receiver operating characteristics (ROC) curve and the AUC of the Precision-Recall (PR) curve were computed.

Modeling results were internally validated via the bootstrap resampling method, according to the TRIPOD guidelines and optimism-corrected metrics were computed.

Results

For each image, a total of 75 radiomic features were extracted.

After feature selection the following models were developed:

- Radiomic model: included three textural features ("F_cm.ifo.corr.2", "F_rlm.r.perc", "F_szm.zs.var")
- Clinical-Radiomic model : included one textural feature ("F_szm.zs.var") and age;
- Radiomic-Ultrasound model: included one textural feature ("F_cm.info.corr.2") and one ultrasound variable (Color Score);
- Ultrasound model: included three ultrasound variables (Color Score, presence of papillary projection, ovarian crescent-sign).

The internally validated (optimism-corrected) AUC was 0.69 (95% CI: 0.57-0.77) for the Radiomic model, 0.80 (95% CI: 0.72-0.90) for the Clinical-Radiomic model, 0.68 (95% CI: 0.56-0.77) for the Radiomic-Ultrasound model and 0.78 (95% CI: 0.68-0.88) for the Ultrasound model.

The Clinical-Radiomic model (ROC AUC 0.80) performed similarly to the Ultrasound model (optimism corrected ROC AUC 0.80 vs 0.78, p= 0.726) and better than the Radiomic (optimism corrected ROC AUC 0.80 vs 0.69, p = 0.042) and Radiomic-Ultrasound (optimism corrected ROC AUC 0.80 vs 0.68, p = 0.048).

Conclusion

The Clinical-Radiomic model performed essentially the same as the Ultrasound model and outperformed both the Radiomic model and the Radiomic-Ultrasound model. In this setting, radiomics did not provide additional value beyond the ultrasound-only model, although it offers the advantage of requiring only patient age in addition to the ultrasound image as input. This suggests that there is promise in continuing to explore radiomic analysis for ovarian masses, with the goal of better defining and predicting specific histologies.

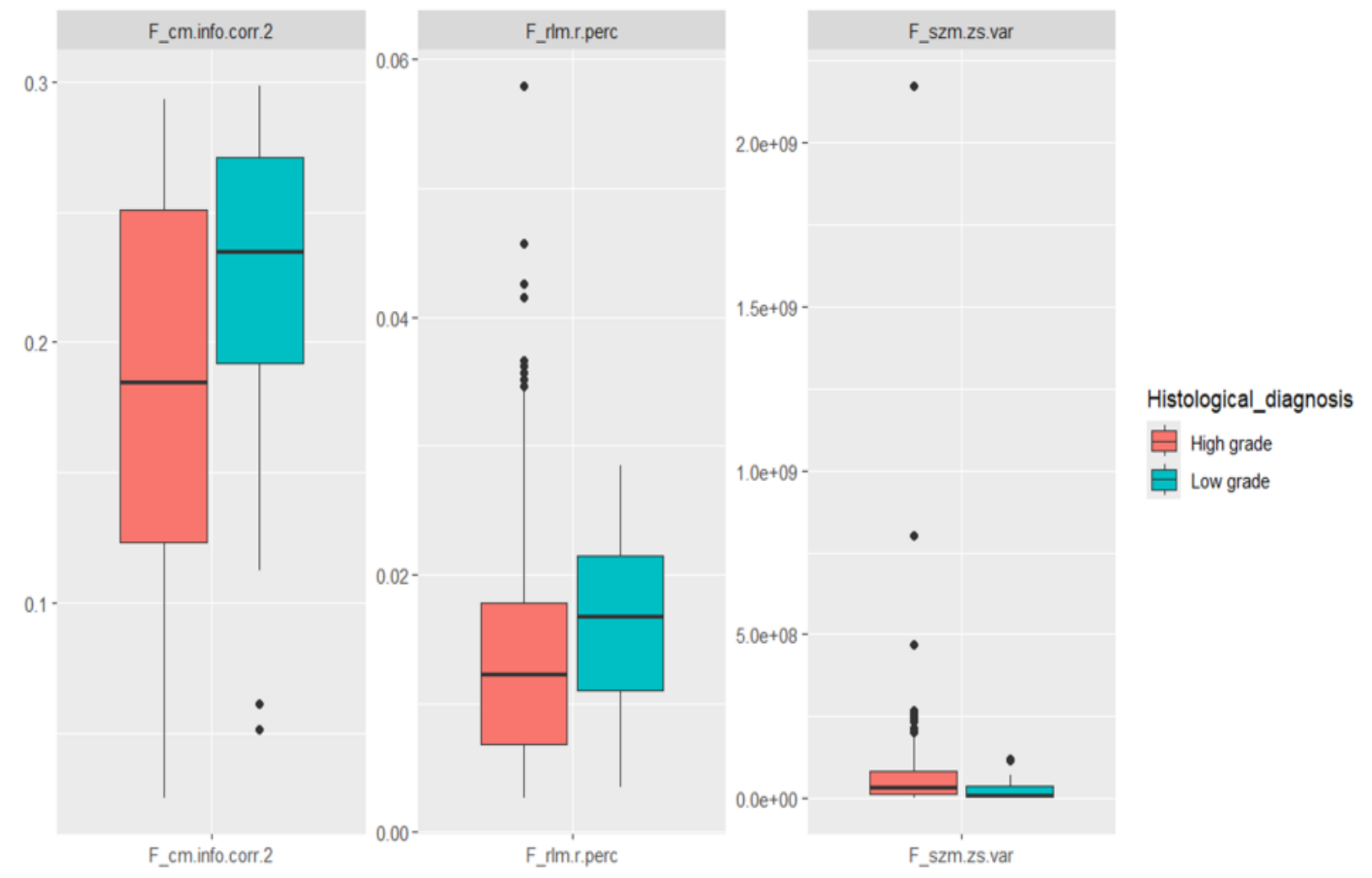


Figure 1. Boxplot of the statistically significant features Class 0 and class 1 of histological diagnosis include patients diagnosed with HGSC and LGSC, respectively.

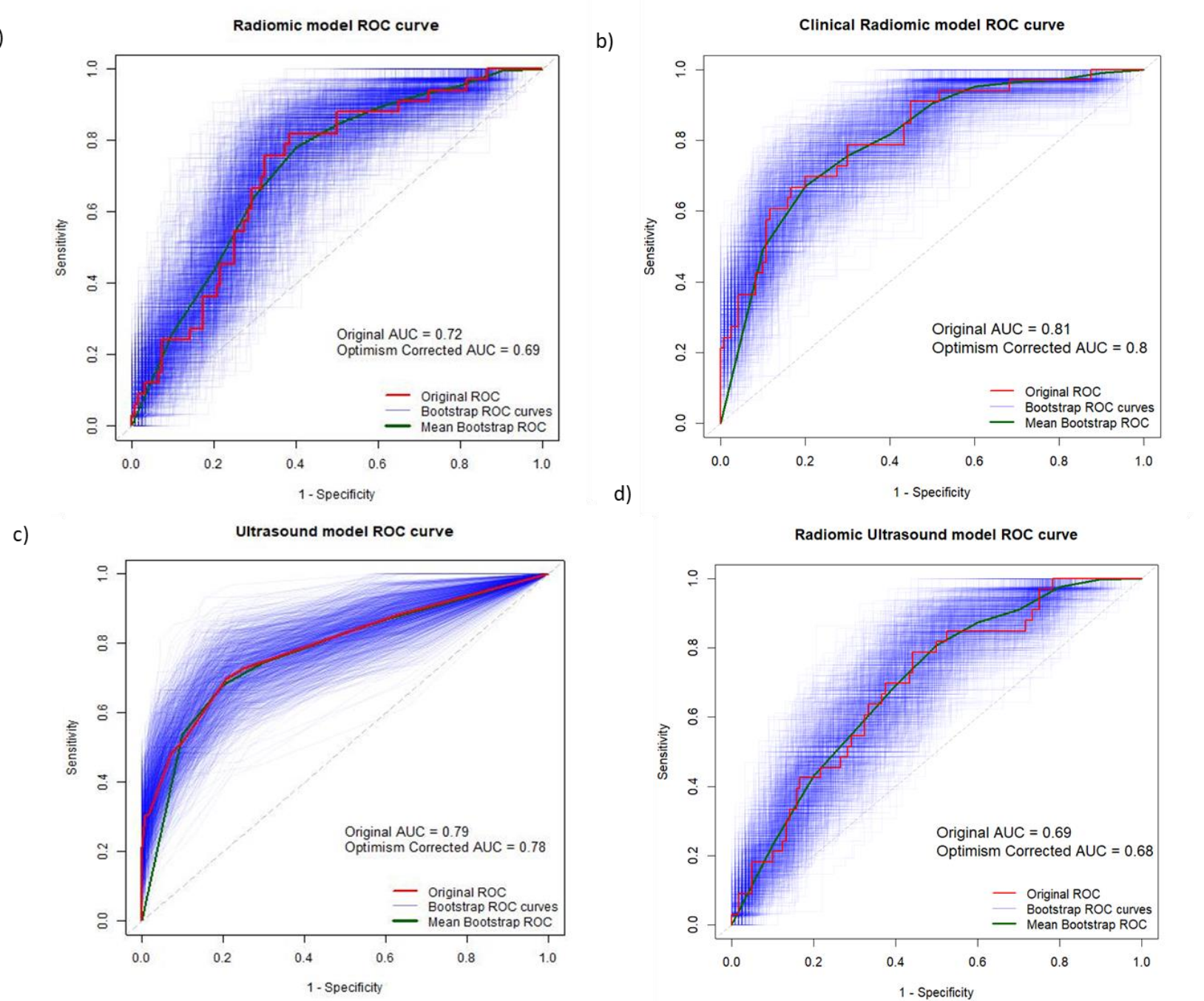


Figure 2. ROC curves of the Radiomic model, Clinical-Radiomic model, Ultrasound model and Radiomic-Ultrasound model during model training (red), Bootstrap internal validation ROC curves (blue) and the mean Bootstrap ROC curve (green).

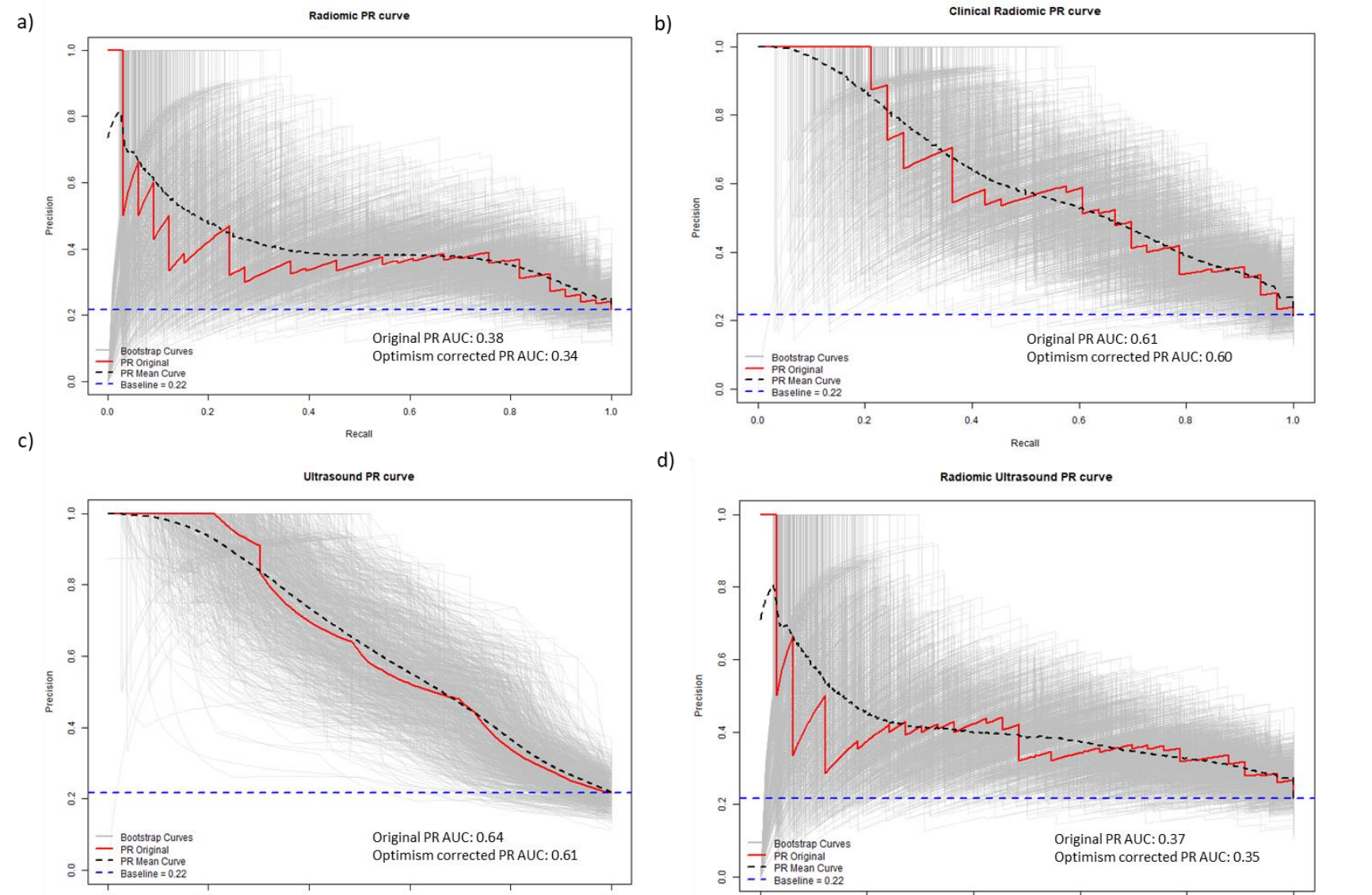


Figure 3. Precision–recall (PR) curves of the Radiomic, Clinical-Radiomic, Ultrasound and Radiomic-Ultrasound model, reporting the PR AUC during model training (red), in the bootstrap samples (gray) and the average of the bootstrap PR curves (black). The dashed blue line represents the baseline – 0,22 and in all the curves, the PR AUC is always higher than the baseline, meaning that the classifier does always better than randomly guessing predictions.