Artificial intelligence-based image analysis can predict outcome in high-grade serous carcinoma via histology alone

Anna Ray Laury, Sami Blom, Tuomas Ropponen, Anni Virtanen & Olli Mikael Carpén

aiforia® AI for image analysis

HELSINGIN YLIOPISTO
HELSINGFORS UNIVERSITET
UNIVERSITY OF HELSINKI

BACKGROUND

High-grade extrauterine serous carcinoma (HGSC) is an aggressive tumor with high rates of recurrence, mainly found in ovaries. Standard therapy includes debulking surgery and platinum-based chemotherapy, using progression free interval (PFI) as an indicator of treatment efficacy. Treatment response and disease progression are poorly understood.

Artificial intelligence (AI) based image analysis has improved prediction and identification of other tumor types, suggesting a new treatment prediction model for HGSC tumors as well. There is likely an intrinsic difference in HGSC tumors that are platinum sensitive (PFI >12 months) and resistant (PFI <6 months). However, this variation is currently not prospectively identifiable by pathologists. The authors propose that some indication of this underlying difference is detectable in the tumor morphology with the use of AI for differentiation of these two groups of patients.

Fig 1. High confidence pixel mask of digital biomarkers per outcome group.

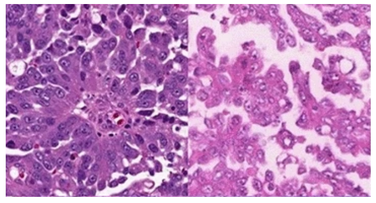


Fig 2. Representative digital biomarker regions identified as associated with PFI-S.

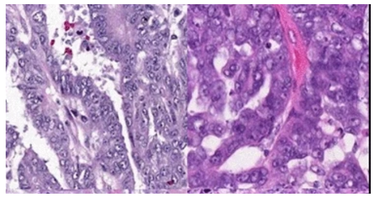


Fig 3. Representative digital biomarker regions identified as associated with PFI-L.

METHODS

Sample selection and preparation

Patients were selected from those diagnosed with HGSC and treated at HUS Helsinki University Hospital between 1982 and 2013. Criteria for selected patients stage III-IV disease at presentation, primary cytoreductive surgery, and at least 6 cycles of adjuvant platinum-based chemotherapy. All slides were prepared from archival formalin fixed paraffin embedded (FFPE) tissue blocks, and stained with hematoxylin & eosin (H&E) to create the whole slide images (WSI). The slides were digitized using a whole slide scanner and uploaded to Aiforia's cloud-based platform.

Neural Network Training (n=205, 30 patients, 2–13 per woman)

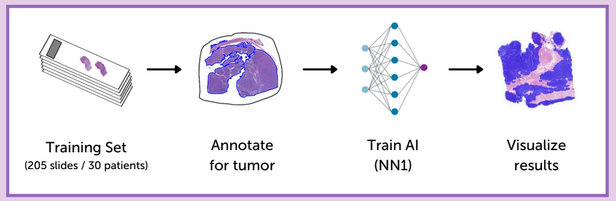
The AI model was trained on 205 WSI of adnexal tumors from 30 women. Patients had evidence of biochemical remission/response at some point during treatment (defined as CA-125 < 35 IU/ml) and split into two groups:

PFI-L: extended progression free survival (≥18 months, n=13)
PFI-S: very short time to tumor progression (≤6 months, n=17)

Validation (n=22)

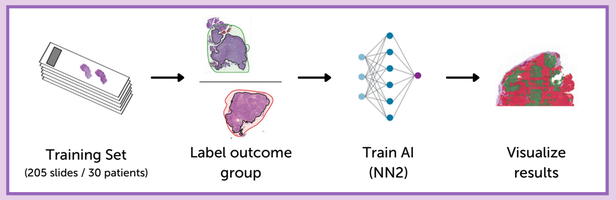
The validation test set consisted of separate group of women with same selection criteria as training cohort and one representative slide from each woman (11 in the PFI-S group and 11 in PFI-L). Combined inference pipeline was applied to the validation test set and WSI were classified by relative percent area of digital biomarkers identified within the tumors.

1



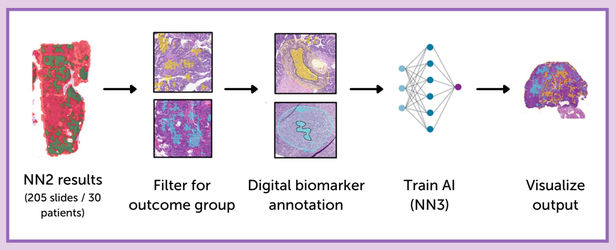
Neural network 1 (NN1): Pathologist teaches AI model tumor segmentation within whole slide image (WSI). Gross tumor regions annotated to exclude background/benign tissue, artifacts (eg folded tissue), whitespace, and extensive necrosis.

2



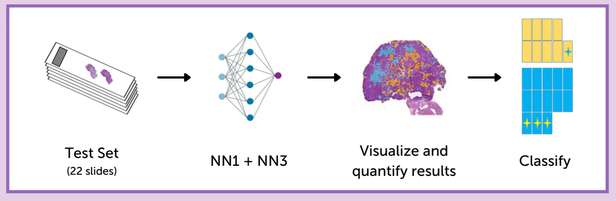
Neural network 2 (NN2): Tumor segments from NN1 relabeled based on patient outcome group. AI model trained to associate tumor features with the outcome group (PFI-S and PFI-L).

3



Neural network 3 (NN3): Output of NN2 filtered by pixel-level confidence for digital biomarkers (what features the AI model identifies per outcome group). High-confidence digital biomarkers reviewed and manually annotated into a data set to training new AI model which tumor features strongly associate with the outcome group.

4



NN1 and NN3 combined to run final inference with output for segmentation of tumor (NN1) and percentage of outcome classes within the tumor region. NN3 inference results are filtered by NN1 inference results.

RESULTS

- In validation test set, all slides had pixel areas identified as digital biomarkers (high-confidence pixel areas) for short PFI, and all but one slide had at least focal digital biomarker regions for long PFI. Representative digital biomarkers are shown in Fig. 1 & 2.
- Classification of validation test set was very similar for both output groups of the AI model.
- 8/11 PFI-S samples and 10/11 PFI-L samples correctly classified
- Sensitivity = 73%; specificity = 91%

No obvious morphologic or artifactual parallels were observed by reviewing the test set WSI visual results alongside the classification results indicating AI model results were not due to a retrospectively apparent similarity between the slides.

DISCUSSION

This study provides evidence that an AI model can use tumor histology alone to predict the biological response of HGSC to adjuvant platinum chemotherapy, using PFI as a proxy with high sensitivity and specificity.

This study is of practical and conceptual importance, as there are currently no validated tissue-based prognostic or predictive markers for primary platinum-based treatment in use for HGSC. AI models have the potential to provide a mechanism for assessing HGSC tumor morphology quickly and in a clinically relevant manner. This provides a solid foundation for future investigations to identify and confirm differences in gene expression between the digital biomarker regions, especially PFI-L and PFI-S.

Need more info or want to book a demo?

Contact Aiforia