

AI-Assisted Ki67 Evaluation in Solid Tumors: Consistency and Reliability Compared to Expert Pathologists

Rania Gaspo¹, Xavier Pichon², Maroua Tliba², Sabine Iglesias², Darshan Kumar³, Renaud Burrer², Amanda Finan-Marchi², Marie Gêrus-Durand²

¹Cerba Research, Laval, QC, Canada, ²Cerba Research, Montpellier, France, ³Aiforia, Helsinki, Finland



Poster 5495

Cerba Research



Background

Ki67 is a key proliferation marker in solid tumors, particularly relevant for HR+/HER2- breast cancer when guiding adjuvant therapy. Despite its clinical importance, Ki67 immunohistochemistry (IHC) scoring lacks standardization. International guidelines aim to reduce variability among pathologists^{1,2}, and AI-driven image analysis solutions have recently emerged as rapid and reliable alternatives. This study compares Ki67 scoring using Aiforia® (AI platform) and Halo® (supervised image analysis software) against three independent pathologists across a large solid tumor cohort.

Method

We stained 192 tumors from various tissue origins (Fig 1), including breast and prostate, using the CONFIRM anti-Ki67 [30-9] monoclonal primary antibody (IVD) on the Ventana Benchmark Ultra platform. Three pathologists were trained in accordance with the International Ki67 Working Group (IKWG) recommendations and scored the tissues following these guidelines.³

Using a deep-learning-based approach, the Aiforia® platform automatically quantified Ki67-positive tumor cells (Ki67+) within minutes. In parallel, the Halo® software random forest classifier was applied to segment the images into tumor, non-tumor, and background regions; these classifications were independently verified by a pathologist. Following cell segmentation, Ki67 positivity was determined by thresholding.

A matched-pairs statistical analysis was performed using JMP® software.

Workflow

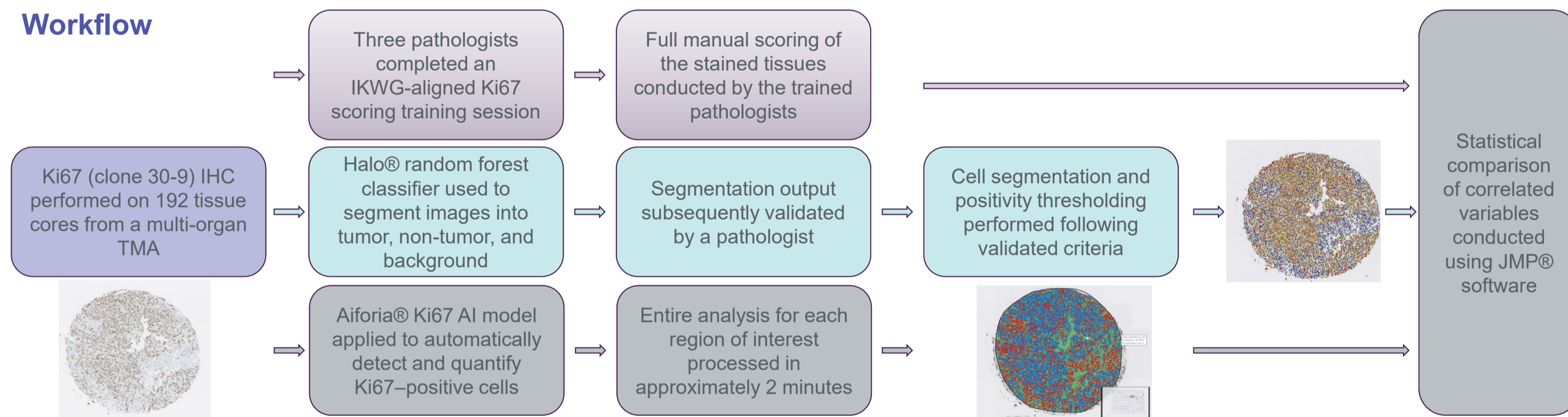


Fig 2. Workflow description with an example of an IHC Ki67 staining using a lung cancer specimen (papillary adenocarcinoma).

Results: Matched-pairs analysis of Ki67 quantification in solid tumors

Correlation analyses were performed to compare Ki67 scoring outputs generated by two image-analysis platforms (Aiforia® and Halo®) and three pathologists.

The strongest correlation was observed between Aiforia® and Halo® (r=0.95). Correlation between Aiforia® and individual pathologists (A, B, C) ranged from fair to strong (r=0.83, 0.82, and 0.94). In contrast, Halo® compared with pathologist scoring showed mostly fair agreement (r=0.76, 0.80, and 0.89). Inter-pathologist correlation was weaker overall (r=0.78 for B-A, r=0.86 for C-A, and r=0.85 for C-B). **The lowest correlation observed in the dataset was between Pathologist A and Halo® (r = 0.76) (Fig 5).**

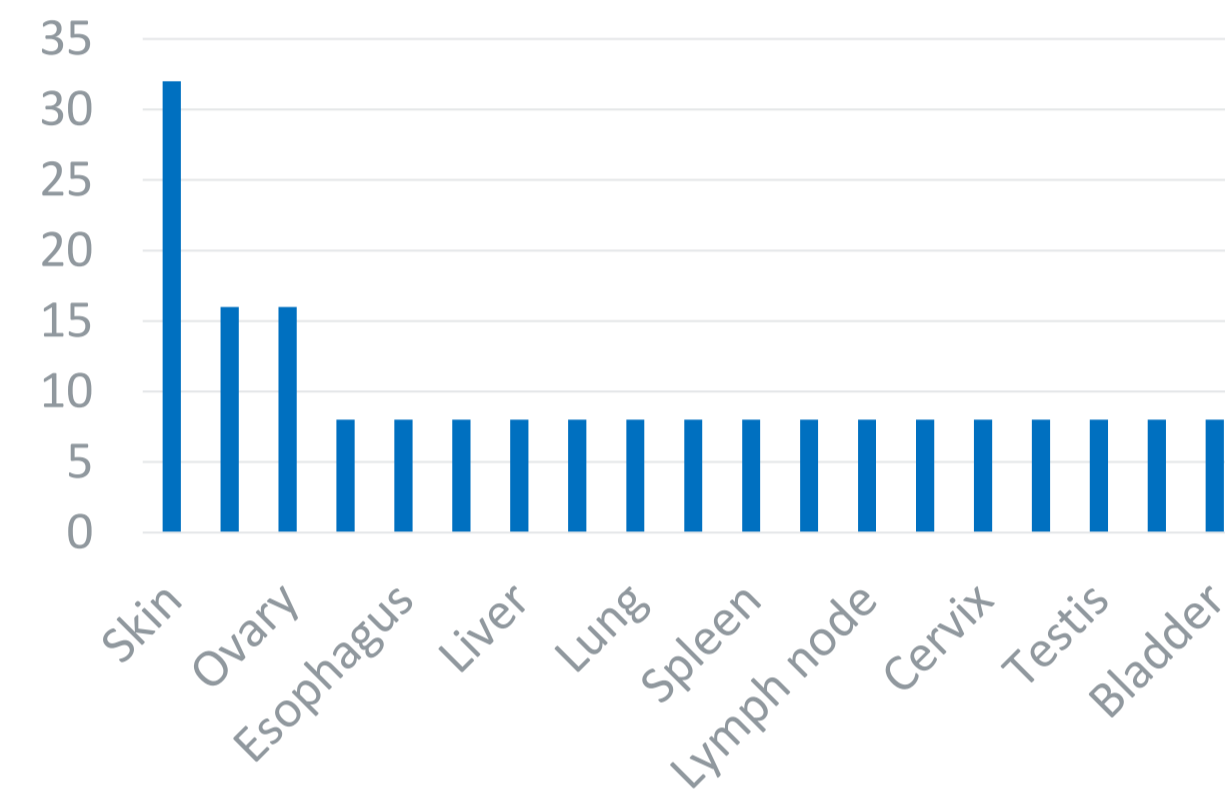


Fig 1. Sample size by solid tumor type from the multiple organ tumor tissue microarray (TMA) (n=192).

Results: Ki67 quantification by solid tumor type

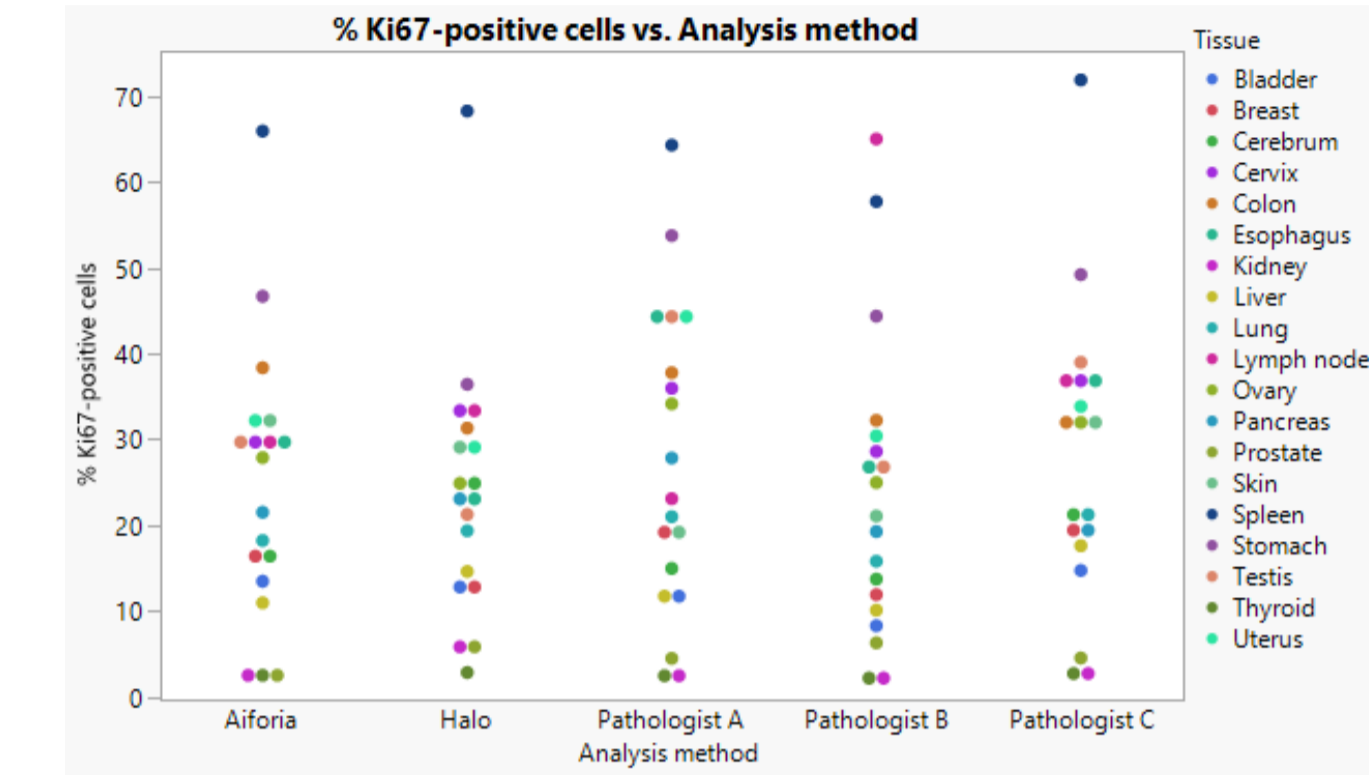


Fig 3. Mean percentage of Ki67-positive cells by tissue type analyzed by different methods. Graph prepared using JMP®.

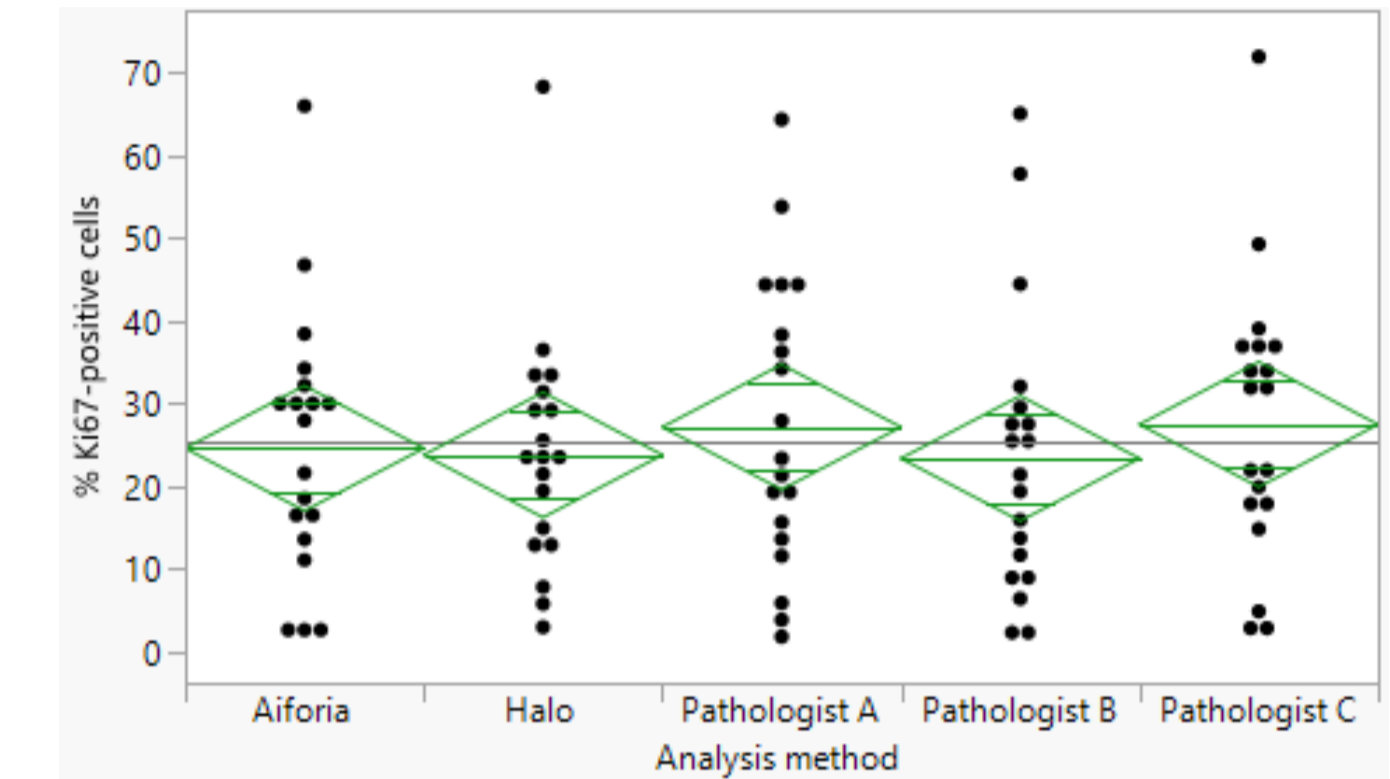


Fig 4. Difference between mean percentage of Ki67-positive cells. No significant difference in overall mean % of Ki-67 positive cells between different analysis methods. Graph prepared using JMP®.

Results: Matched pairs analysis of Ki67 quantification results in various solid tumors (n=158)

Halo-Aiforia	0.95
C-Aiforia	0.94
C-Halo	0.89
C-A	0.86
C-B	0.85
A-Aiforia	0.83
B-Aiforia	0.82
B-Halo	0.8
B-A	0.78
A-Halo	0.76

Fig 5. Aiforia® and Halo® demonstrated the highest correlation in Ki67 quantification (green), while the poorest agreement was observed between Pathologist A and Halo® (red). R values indicated on the graph bars. Matched pair analysis performed using JMP®.

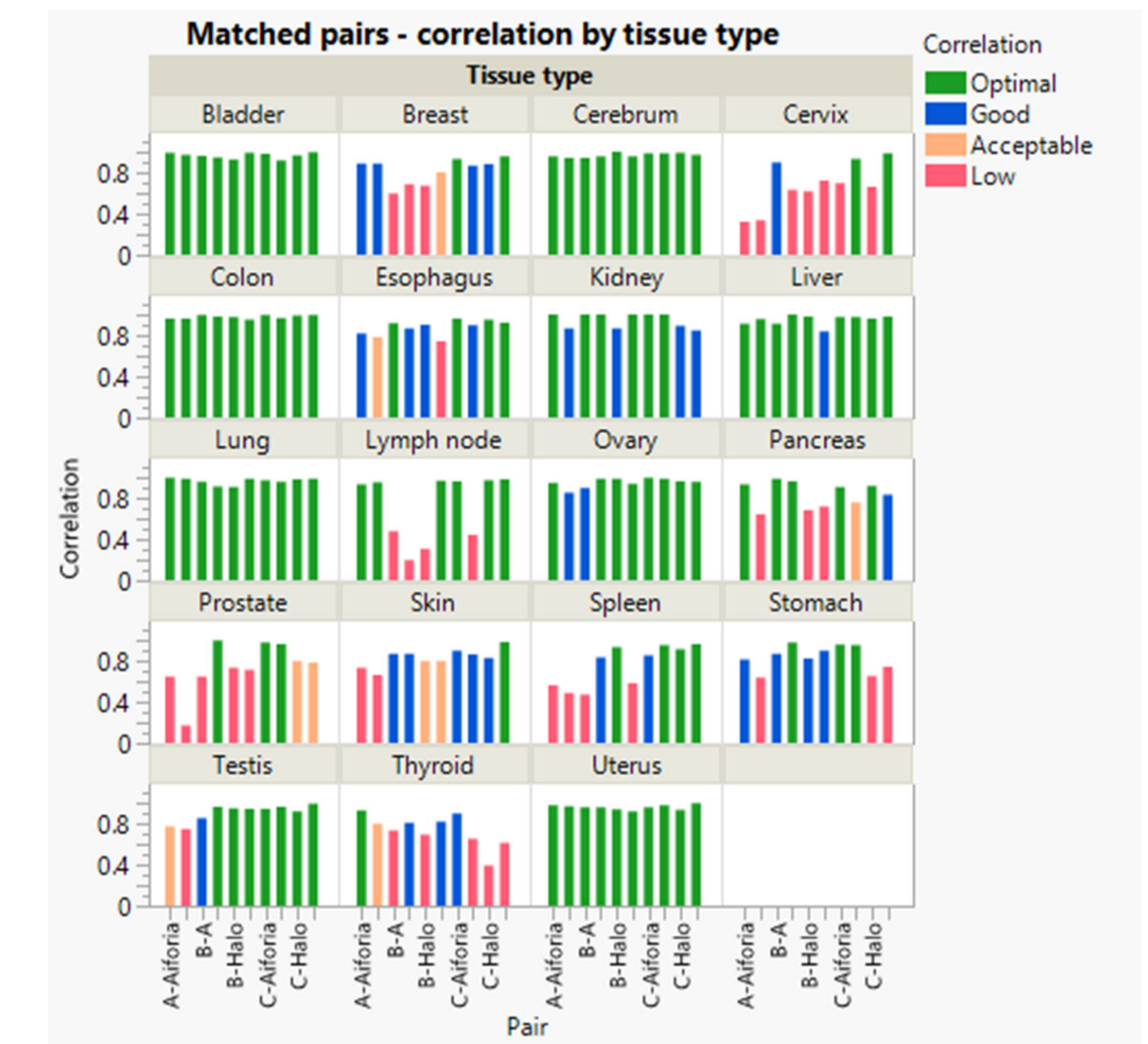


Fig 6 (On the right). Results of Ki67 quantification by solid tumor type with corresponding matched-pairs analysis. Cell color coding for r-values: green (optimal) ≥ 0.90 ; blue (good) = 0.90–0.80; orange (acceptable) = 0.80–0.75; red (low) = below 0.75. Matched pair analysis performed using JMP®.

Conclusion

Across indications, correlations between Aiforia® and Halo® as well as inter-pathologist comparisons were consistently high ($r > 0.90$). While inter-pathologist agreement varied notably by tumor type, the two image-analysis platforms produced highly concordant results regardless of tissue origin. E.g. In lymph node Ki67 scoring, Aiforia® and Halo® showed strong correlation ($r = 0.98$), whereas the corresponding inter-pathologist correlation was markedly lower ($r = 0.44$ for C-B). Among the 19 primary tumor types evaluated, only stomach demonstrated a correlation below 0.75 between the two software platforms ($r = 0.74$), despite inter-pathologist agreement remaining high ($r = 0.86$ – 0.95). These findings underscore the consistency of automated image-analysis tools relative to the variability observed in manual pathologist scoring. **AI-based platforms like Aiforia® and supervised image analysis tools such as Halo® provide robust, reproducible Ki67 scoring and significantly reduce inter-observer variability. These technologies offer valuable assistance for IHC-based clinical analysis and may serve as arbitration tools or standardize Ki67 evaluation in solid tumors.**