

Validation of a novel artificial intelligence-based tool for Ki-67 quantification in breast cancer: enhanced accuracy, reproducibility, and efficiency in resource-limited setting

A.P. Nellihela¹, R.M.N.D. Rathnayake², N.J. Asanthi³, V.C. Bandaranayake¹, A. Ramanan², T. Kokul², and M. Senarathna⁴

¹Teaching Hospital Anuradhapura, Department of Surgical Oncology, Anuradhapura, Sri Lanka

²University of Jaffna, Department of Computer Science, Jaffna, Sri Lanka

³Teaching Hospital Anuradhapura, Department of Histopathology, Anuradhapura, Sri Lanka

⁴University of Peradeniya, Postgraduate Institute of Science, Kandy, Sri Lanka

Abstract:

Background & Objectives: Ki-67 is an important proliferation marker in guiding breast cancer treatment. International Ki-67 Working Group (IKWG) recommends hot-spot, and the Global Scoring (GS) methods to quantify Ki-67 and levels are categorized as low (<5%), intermediate (6–29%), and high(>30%). These manual methods are time-consuming and prone to inter-observer variability and potential sampling bias based on selected microscopic fields. This study aimed to validate a novel Artificial Intelligence (AI)- based tool, evaluating its accuracy and efficiency compared to manual scoring. The AI tool captures digital images of the entire Ki-67-immunohistochemistry-stained slide in the high-power microscopic field in real-time and automatically calculates the Ki-67 index.

Methods: An AI model was developed using EfficientNetV2 architecture, incorporating distance transformation and watershed algorithms for precise cell segmentation. The model was trained on 350 digital images annotated by pathologists and validated with an additional 350 images. Twenty Ki-67-stained slides from three distinct histopathology units in Sri Lanka were used to minimise variations in slide preparations. AI-derived Ki-67 indices were compared with blinded manual scores by three pathologists using Hot Spot and Global Scoring (GS) methods. Agreement was analysed using Intraclass Correlation Coefficient (ICC) and Bland-Altman plots. Assessment efficiency was compared by measuring evaluation time.

Results: The AI model showed 94.3% accuracy in cell-to-cell identification against pathologist annotations. ICC analysis indicated excellent reproducibility between AI and manual methods for Hot Spot (ICC=0.946, 95% CI:0.672–0.984) and Global Scoring (GS) method (ICC=0.963, 95% CI:0.879–0.987). Bland-Altman analysis confirmed strong agreement without significant systematic bias. The AI method significantly reduced evaluation time (seconds) compared to manual scoring by pathologists (Hot Spot: AI mean=25.5s, Pathologists mean=152.2s, $p<0.01$, $t=-7.38$; GS: AI mean=250.8s, Pathologist mean=455.1s, $p<0.01$, $t=-6.96$).

Conclusion: The AI-based tool offers accurate, reproducible, and significantly more efficient Ki-67 quantification, representing a reliable, standardised alternative to manual methods. The AI tool eases workload and reduces diagnostic turnaround times in resource-limited settings. Limitations include dependence on digital microscopy and high-quality slide preparation.