

Usability of an Artificial Intelligence-enhanced Multispectral Imaging Tool for Burn Assessment: A Retrospective Heuristic Evaluation

Hajar S. Abdulla, MBBS, MRCS - Royal Victoria Infirmary, Newcastle

Poh (Leslie) Tan, MBBS, MRCS(Eng), MSc - Royal Victoria Infirmary, Newcastle

Zeeshan Sheikh, MBChB, MRCS, MRCS(Eng) - Manchester Foundation NHS Trust, England, UK

Zeeshan Sheikh, MBChB, MD, MRCS - Manchester NHS Foundation Trust

Christopher J. Lewis, MD, PhD - Northern Regional Burn Centre

Introduction: Artificial intelligence (AI)-enhanced multispectral imaging (MSI) is an emerging technology for burn assessment. These devices offer objective predictions of burn healing potential however, their usability in clinical practice remains poorly understood. Key factors such as device handling, workflow integration, and interpretability are critical to successful adoption. This study evaluated the usability of an AI-enhanced MSI device, drawing on nursing and surgeon feedback mapped to consolidated usability heuristic domains.

Methods: We conducted a multicentre retrospective evaluation across four Burn Centres using the device. Clinicians completed a structured survey comprising twelve Likert-scale statements (1 = strongly disagree, 5 = strongly agree), each mapped to six heuristic domains: visibility & recognition, match & standards, control & freedom, error prevention & recovery, flexibility & efficiency, and minimalism & aesthetics. Usability outcomes were reported as domain-specific means with standard deviations, medians with interquartile ranges, and proportions of positive responses (%). Between-group differences by clinical role were assessed using Mann-Whitney U tests, while site-level variability was examined with Kruskal-Wallis tests. Free-text responses were thematically analysed & mapped to heuristic domains to complement the quantitative findings.

Results: Forty-seven clinicians responded (60% nurses, 40% surgeons). Overall usability ratings were high, with domain means ranging from 3.95 to 4.36. Flexibility and efficiency scored highest (mean 4.29; 87.9% positive) while match and standards scored lowest (mean 3.95; 65.9% positive). Surgeons rated match and standards slightly higher than nurses (mean 4.11 vs 3.86; 73.7% vs 60.7% positive), though this difference was not statistically significant ($U = 302.0, p = 0.42$). No statistically significant differences were found between hospitals across any domain (all $p > 0.05$). Free text comments emphasised ease of use & expressed interest in future refinements, including scanning hands & feet and providing more granular prognostic outputs.

Conclusions: This study demonstrates that the AI-enhanced MSI device is both usable and acceptable across multidisciplinary burn teams, with consistent evaluations supporting its integration into routine practice. Its strengths in workflow efficiency & ease of handling highlight its potential as a practical adjunct to existing clinical assessment

methods. Continued refinement & prospective validation will be essential to establish its role in optimising burn care delivery.

Applicability of Research to Practice: These findings support the integration of AI enhanced MSI devices into burn care, with high usability across roles and sites, suggesting potential for broad adoption with minimal training. Targeted refinements could further streamline workflow and improve user experience.

Funding for the study: None

Usability of an Artificial Intelligence-Enhanced Multispectral Imaging Tool for Burn Assessment: A Retrospective Heuristic Evaluation

Hajar Abdulla¹, Poh Tan¹, Sarah Bache², Alexandra Murray³, Zeeshan Sheikh⁴, Christopher Wearn⁵, Christopher J. Lewis¹

1. Northern Regional Burns Centre, Royal Victoria Infirmary, Newcastle upon Tyne, United Kingdom; 2. Department of Burns and Plastic Surgery, Queen Beetham Hospital, Birmingham, United Kingdom; 3. Department of Burns, Plastic & Reconstructive Surgery, Stoke Newington Hospital, Bournemouth, United Kingdom; 4. Manchester Burns Centre, Manchester Royal Infirmary, United Kingdom; 5. Department of Burns, Plastic and Reconstructive Surgery, Salford Hospital, Salford, United Kingdom

Background and Objectives

AI-enhanced multispectral imaging (AI-MSI) offers objective, image-based predictions, but evidence on day-to-day usability is limited. We evaluated usability of an AI-MSI system in routine practice using a multicentre, modified Nielsen heuristic framework.

Methods

- A retrospective, mixed-methods study was conducted across five UK burn services (September 2024 to August 2025).
- Doctors and nurses completed a structured survey comprising twelve Likert-scale statements (1 = strongly disagree, 5 = strongly agree), each mapped to six heuristic domains.
- Usability outcomes were reported as domain-specific means with standard deviations, medians with interquartile ranges, and proportions of positive responses (scores ≥4).
- Between-group differences by clinical role were assessed using Mann-Whitney U tests, while site-level variability was examined with Kruskal-Wallis tests.
- For domains showing large effect sizes ($\eta^2 \geq 0.15$), Dunn's post hoc pairwise comparisons were conducted.

Free-text responses were thematically analysed to complement quantitative findings.

Results

Domain	Mann-Whitney U		Euskal-Wallis Test	
	U	p	H (df=4)	η^2 (effect size)
Visibility & Recognition	296.5	0.767	-0.01	(negligible)
Match & Standards	323.5	0.537	-0.10	(negligible)
Control & Freedom	243.0	0.266	0.17	(small)
Error Prevention & Recovery	257.5	0.436	0.12	(small)
Flexibility & Efficiency	243.0	0.315	0.17	(small)
Minimization & Aesthetics	214.0	0.077	0.27	(small-moderate)

Forty-nine clinicians participated (43% doctors, 57% nurses). Flexibility and Efficiency scored highest (mean 4.29; 87.9% positive), while Match and Standards scored lowest (mean 3.95; 65.9% positive). Mann-Whitney analyses revealed no statistically significant differences between doctors and nurses across any domain (all $p > .05$). Kruskal-Wallis tests showed no statistically significant differences across hospitals (all $p > .05$), though Match and Standards demonstrated a large effect size ($\eta^2 = 0.15$). Dunn's post hoc test confirmed that no specific hospital pairs differed significantly after Bonferroni correction (all p not $> .05$). Qualitative responses highlighted ease of use, visual clarity, and workflow integration, while suggesting refinements such as enhanced limb scanning and more granular predictive outputs.

Conclusions and Applicability to Practice

AI-MSI showed high usability across roles and centres, supporting routine use in outpatient and inpatient settings. As the first multicentre heuristic usability evaluation of AI-MSI in burns, the study provides implementation-focused evidence to guide training, pathway design, and targeted software and hardware refinements.