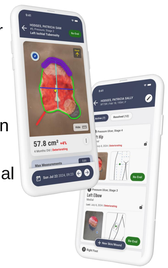
Improving Wound Tissue Evaluation Accuracy: Comparing Clinicians' Assessments and AI-Driven Technology Across Diverse Skin Tones

Heba Tallah Mohammed,[1] Robert D. J. Fraser, [1,2] Sheila Wang,[1,2] Amy Cassata,[1] Justin Allport[1]

[1] Swift Medical, Toronto,[2] Arthur Labatt Family School of Nursing, Western University, [3] Women College Hospital, [4] Temerty Faculty of Medicine, University of Toronto

Introduction

- Chronic wounds have a significant impact on patients and healthcare systems.[1,2]
- Accurate assessment of different tissue types is essential for effective wound management.[3]
- Wound assessment relies heavily on a clinicians' ability to classify different tissue types within the wound bed,[4] which can be inconsistent due to knowledge gaps, visual perception limitation and variations in skin tones.[3]
- AI-driven wound care technologies (AI-WCT) provide potential solutions to these limitations. Swift Medical's SmartTissue™ is trained on hundreds of thousands of images to improve consistency and reduce clinician bias (e.g., experience, skin tone) and enhance segmentation and quantification.



Objective

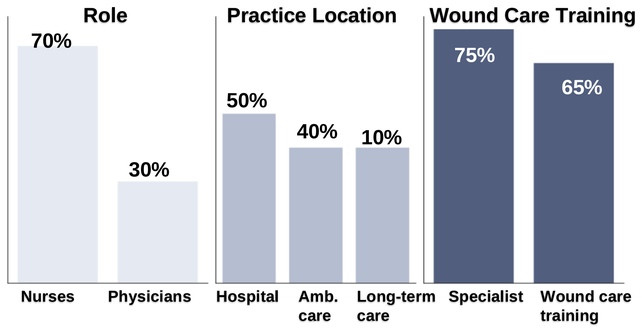
- To compare clinicians' subjective assessment and classification of wound tissue types with those generated by AI-WCT, highlighting case studies on variations influenced by skin tone and wound complexity

Methodology

- In May 2024, an observational study surveyed 20 wound care clinicians of varying expertise online, recruited through purposive sampling at a wound care conference.
- The survey included questions about clinicians' experience, practice, and wound care training. The survey also included de-identified wound images from patients with diverse skin tone and wound complexity.
- Wounds were classified as complex if they displayed combination of diverse tissue types, and non-complex if the wound presented with one or two tissue types.
- Skin tones were categorized using the Fitzpatrick scale (Type I/II for light/fair skin, Type V/VI for dark/darkest skin).
- Each clinician was asked to identify and estimate the different tissue types within each wound, such as granulation, slough, eschar, and epithelialization, and their percentage.
- Estimates were aggregated and averaged to compare to the AI-generated classification.

Results

Clinicians Demographic: Role, Settings and Training

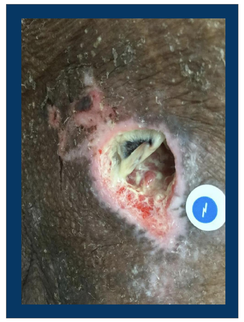


Case 1- Complex Wound in a Light Skin Tone Patient

	Clinician Average Assessment	AI-Generated Scores
Granulation Tissue	35.04% (±12.80)	30.64%
Slough	61.78% (±18.33)	47.94%
Eschar	5.20% (±10.89)	0.00%
Epithelialization	36.94% (±6.38)	21.42%

Clinicians overestimated granulation tissue (+4.4%), slough (+13.84%), and epithelialization (+15.2%) compared to AI. Clinicians also reported eschar, which was not detected by AI.

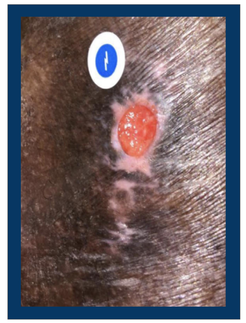
Case 2- Complex Wound in a Dark Skin Tone Patient



	Clinician Average Assessment	AI-Generated Scores
Granulation Tissue	23.30% (±14.26)	13.25%
Slough	52.80% (±18.52)	41.22%
Eschar	10.71% (±9.17)	0.00%
Epithelialization	8.53% (±11.02)	45.53%

Clinicians overestimated granulation tissue (10.05%) and slough (11.58%) compared to AI. They also reported eschar (10.71%), which AI did not detect (0.00%). In contrast, AI identified more epithelialization (36.8%) than clinicians.

Case 3- Non-Complex Wound in a Dark Skin Tone Patient



	Clinician Average Assessment	AI-Generated Scores
Granulation Tissue	81.00% (±26.98)	44.00%
Slough	7.70% (±17.85)	0.30%
Eschar	0.00% (±22.34)	0.00%
Epithelialization	7.15% (±10.90)	55.50%

Clinicians overestimated granulation tissue by 37% and slough by 7.4% compared to AI. Meanwhile, AI detected 48.35% more epithelialization than clinicians. Both assessments agreed on the absence of eschar.

Discussion

- Clinicians frequently identified the presence of eschar when it was not predicted by the AI, overestimating its occurrence in wounds. Further, clinicians consistently noted elevated levels of slough and granulation tissue. underscoring the inconsistencies and variations among clinicians.
- Clinicians reviewing the same images showed variations up to 36% (slough) when estimating tissue types, and frequently reported tissue types (eschar) not identified by by the AI, raising questions about subjective data reliability for wound prognosis.
- Inconsistent tracking and reporting may cause delays in identifying and intervening in non-healing wounds or change in treatment of a wound that is objectively improving.
- Epithelialization, a crucial indicator of wound healing progress often characterized by subtle color changes in the wound bed, was consistently underestimated by clinicians – particularly in darker skin tones.
- AI can transform clinical decision by providing consistency and objective tracking.

References

1. Järbrink K, Ni G, Sönnergren H, Schmidtchen A, Pang C, Bajpai R, et al. The humanistic and economic burden of chronic wounds: A protocol for a systematic review. Syst Rev. 2017;6(1):15.
2. Phillips CJ, Humphreys I, Fletcher J, Harding K, Chamberlain G, Macey S. Estimating the costs associated with the management of patients with chronic wounds using linked routine data. Int Wound J. 2016;13(6):1193-7.
3. Dhoonmoon L, Harikrishna KR, Abbas Z. Wound care and skin tone: Addressing skin tone bias in wound assessment. *Wounds Int.* 2023. Available at: https://www.woundsinternational.com. Accessed March 12 2025.
4. Lazarus GS, Cooper DM, Knighton DR, Margolis DJ, Pecoraro RE, Rodeheaver G, et al. Definitions and guidelines for assessment of wounds and evaluation of healing. Wound Repair Regen. 1994;2(3):165-70
5. Kim RB, Moscoso-Kingsley W, Hirata R, Voronov L, Patel J. Utilization of smartphone and tablet camera photographs to predict healing of diabetes-related foot ulcers. *Comput Biol Med.* 2020;126:104042. doi:10.1016/j.compbiomed.2020.104042.