# DermEquity: A Universal Safety Framework for Medical AI Using Worst-Case Bias Certification and Influence-Based Attribution Validated in Skin Cancer Screening

I. BACKGROUND/OBJECTIVE

- Skin cancer: 112,000 new US cases/year - early detection is life-saving.
- Fewer than 1 dermatologist per 100,000 patients; AI fills the gap.

Fig 1: DermEquity Integrated System Overview

- Critical problems: <10% of dermatology AI training images represent darker skin tones. No pre-deployment worst-case safety standard exists for medical AI.
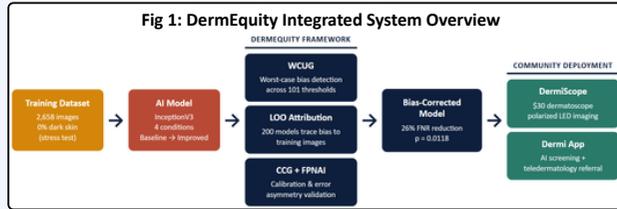
Objective: Build **a stress-test framework with Worst-Case Underdiagnosis Gap (WCUG) and Influence-Based Bias Attribution** for pre-deployment medical AI testing and an **accessible screening system** for equitable skin cancer detection.

II. METHODOLOGY

- **DermEquity Framework:**
  - **Benchmark:** 1,658 skin lesion images balanced across Fitzpatrick skin tones (Fitzpatrick17k)
  - **Models**: InceptionV3 trained at 0%, 5%, 10% dark skin representation + augmented images
  - **WCUG**: Worst-case melanoma miss rate gap across all 101 thresholds; 1,000-iteration bootstrap confidence intervals
  - **Influence Attribution**: 200 leave-one-out models, measuring each training image's contribution to worst-case bias
- **DermiScope:**
  - 3D-printed low-cost dermatoscope equipped with polarized LED with universal smartphone fit.
- **Dermi Mobile App:**
  - Integrates DermEquity framework with DermiScope for optimal community deployment
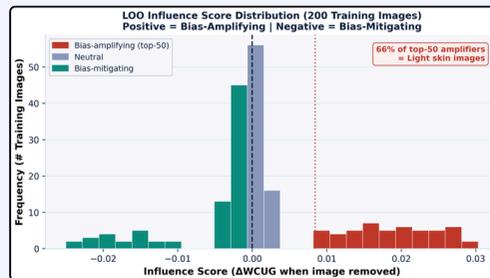
III. RESULTS

- Baseline model hid a **6.8% melanoma miss rate gap on dark skin at threshold 0.4**
- WCUG revealed a **32-point worst-case gap** vs. 15-point standard average gap
- **Ablations**: WCUG dropped 0.068 → 0.047 → 0.027 but bias never eliminated by representation alone
- **Augmented model reversed bias direction entirely:** dark skin no longer disadvantaged at any threshold, and caught **16 more melanomas (↓26% FNR)**
- Removing top-50 bias amplifying images **improved WCUG by 15.6%**
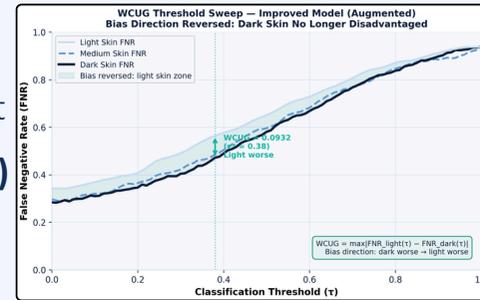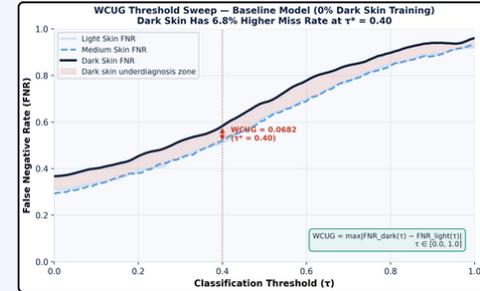
IV. CONCLUSIONS/IMPACT

- **DermEquity:** Pre-deployment safety testing framework for medical AI (possible applications to radiology, pathology, retinal imaging, etc).
  - **DermEquity Benchmark** = stress-test dataset optimized for worst-case fairness measurement
  - **WCUG =** first metric to expose worst-case deployment risk by skin tone across all thresholds
  - **Influence-Based Attribution** = first application to fairness certification in medical AI
- **DermiScope (<$30) + Dermi App:** accessible dermoscopic imaging for diverse data collection and real-world deployment.
- **Future Work**: propose WCUG as FDA-style pre-deployment standard; prospective clinical validation.