Developing a post market surveillance audit methodology for an Artificial Intelligence as a Medical Device (AIaMD) deployed autonomously in Urgent Suspected Skin Cancer Pathways

Dr Karan Punjabi[1], Dr Joshua Luck[1], Dr Dilraj Kalsi[1], Dr Dan Mullarkey[1]

1 Skin Analytics Limited, Smithfield Business Centre, 5 St John's Ln, London EC1M 4BH

Purpose

The post-market vigilance and surveillance (PMS) process enables early detection of undesirable effects and identifies opportunities for improvement or safety relevance, as per EN ISO 13485 or EN ISO 14971. We developed a statistically-led audit methodology to monitor negative predictive value (NPV) in predictable disease prevalence populations, when deploying our AIaMD in urgent suspected skin cancer (USSC) pathways to autonomously discharge benign skin lesions.

Methodology Rationale

Following feedback from a clinical advisory group, a recent NHS England-commissioned independent evaluation of autonomous AIaMD in USSC pathways recommended monitoring NPV for excluding malignant melanoma (MM). This audit methodology reports AIaMD NPV when deployed autonomously.

AIaMD

NPV: **99.95%**
[95% CI 99.85-99.98]

Dermatologist Face-to-Face

NPV: **98.90%**
[95% CI 98.2-99.5]

Figure 1: Real-world NPV for melanoma of AIaMD across 15 NHS sites with sample of >33,000 lesions, compared to a meta-analysis of consultant dermatologist face-to-face NPV[1] in a matched prevalence population.

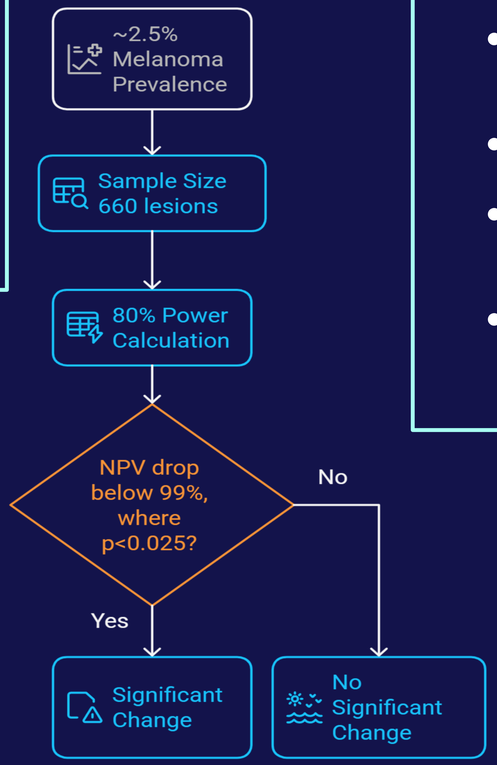


Figure 2: Flowchart to assess whether NPV drop is of significance

Recommendations

- A sample of 660 lesions (~570 patients) provides 80% power to detect a NPV drop to below 99% in a population with MM prevalence ~2.5%, with a one-way 2.5% significance level (Figure 2).
- 570+ discharge cases will therefore be redirected for a second audit read
- Consecutive patient sampling from multiple sites ensures ongoing AIaMD sensitivity and specificity reporting, reflecting a real-world casemix.
- Simulation suggests more frequent audits could speed up signal detection, but given the time required to collect histology outcome data the initial audit will be 6 monthly.

Conclusion

We present a structured and statistically-led methodology for PMS performance monitoring of an autonomous AIaMD in line with its intended use. Quantitative vigilance approaches to demonstrate real-world performance against clinician benchmarked standards should be integrated into post-market clinical follow-up activities for autonomous AIaMD systems.

References

1 *Edge Health. Evaluating Pathways for AI Dermatology in Skin Cancer Detection. 2024 Aug. Available from: https://www.edgehealth.co.uk/wp-content/uploads/2024/08/Evaluating-Pathways-for-AI-Dermatology-in-Skin-Cancer-Detection.pdf*