Automatic Referral of Basal Cell Carcinoma Patients: A Natural Language Processing Approach

Parsa Bagherzadeh

Background and Rationale

Basal cell carcinoma (BCC) is the most common skin cancer, but many cases are diagnosed and treated outside cancer centers. As a result, BCC is often under-reported in cancer registries. Pathology reports contain essential information, but reviewing them manually is slow and inconsistent. Artificial intelligence (AI) can help extract this information automatically and support treatment decisions when a tumor is not fully removed.

Objectives

1. **Automatically identify parts of pathology reports that describe BCC tumors**.

2. **Predict the most appropriate next treatment step**, such as follow-up, re-excision, Mohs surgery, or other procedures.

Methods and Results

We collected **387 pathology reports** from **180 patients** and manually labeled **454 BCC tumors**. We trained an AI model (Clinical Longformer) to find BCC text segments and suggest referrals.

The model accurately identified tumor descriptions with a **strict F1 score of 0.90**. For treatment recommendations, it achieved a **weighted F1 score of 0.86**, meaning it made only **22 errors out of 159** test cases. It performed especially well for cases involving complete excision, Mohs surgery, and re-excision.

Conclusions

AI can reliably read pathology reports, extract BCC information, and recommend next steps without manual feature engineering. This provides a consistent and scalable way to interpret pathology findings.

Anticipated Impact

This system can improve the accuracy of BCC reporting in cancer registries and support clinicians by ensuring patients with incomplete tumor removal receive timely and appropriate care. The approach can be expanded to other skin cancers and combined with imaging or genomic data to enhance personalized treatment.