Title: Basal cell carcinoma risk prediction among long-term survivors of childhood cancer: a report from the Childhood Cancer Survivor Study (CCSS)

Authors: Cindy Im,[1] Christina Boull,[1] Zhe Lu,[2] Hasibul Hasan,[1] Yadav Sapkota,[3] Rebecca M. Howell,[4] Michael A. Arnold,[5] Miriam Conces,[6] Louis S. Constine,[7] Yutaka Yasui,[3] Kirsten K. Ness,[3] Melissa M. Hudson,[3] Gregory T. Armstrong,[3] Joseph P. Neglia,[1] Yan Yuan,[2] Lucie M. Turcotte[1]

Affiliations
1. University of Minnesota
2. University of Alberta
3. St. Jude Children's Research Hospital
4. MD Anderson Cancer Center
5. University of Colorado
6. Nationwide Children's Hospital
7. University of Rochester Medical Center

Background: Basal cell carcinomas (BCCs) represent >50% of subsequent neoplasms among long-term survivors of childhood cancer. While general population BCC risk prediction models are available, models considering clinical features relevant to survivors are lacking.

Methods: To develop risk prediction models for experiencing any subsequent BCC by age 40 years, we used data from five-year survivors in CCSS. BCCs occurring ≥5 years after childhood cancer diagnosis were confirmed via pathology report or medical record review. Predictors included: demographics; childhood cancer diagnosis/age; treatment era; body region-specific cumulative radiotherapy (RT) doses; hematopoietic cell transplantation (HCT); chemotherapy drugs/doses; prior non-BCC skin neoplasms. Among the evaluated statistical learning algorithms, XGBoost was the best-performing algorithm. To assure generalizability, we assessed nested cross-validated prediction performance metrics (AUROC/AUPRC: area under the receiver operating characteristic curve/precision-recall curve). We compared our model with Children's Oncology Group Long-Term Follow-Up Guidelines (COG LTFU, v6.0) recommendations indicating yearly dermatologic screenings following any RT or HCT.

Results: Among 23,172 survivors, 1,164 developed BCCs (63% occurred <40 years of age). The XGBoost model showed good discrimination (AUROC=0.76, 95% CI: 0.74-0.78) and precision (AUPRC=0.28, 95% CI: 0.24-0.32). The most important predictors included diagnosis age and cumulative doses of cranial/neck RT, alkylators, and anthracyclines. The cross-validated BCC prevalence by age 40 was 4%, 17%, and 59%, respectively, in low (<10%), medium (10-49%), and high (≥50%) model-defined predicted risk groups. In comparison, the COG LTFU-based classification prediction performance metrics were significantly worse (AUROC=0.63, P<0.001; AUPRC=0.10, p<0.001). The BCC prevalence by age 40 years among survivors recommended for screening per COG LTFU was 11%; 63% of them were re-classified as low risk using our XGBoost model.

Conclusion: Once validated in the St. Jude Lifetime Cohort, implementation of our subsequent BCC risk prediction models will more accurately inform risk-/age-based screening recommendations which are currently unavailable for this high-risk population.