



OPEN Investigating the impact of clinical and genetic factors on the post-surgery prognosis of sinonasal squamous cell carcinoma

Meng Lian¹✉, Boxuan Han¹, Jiaming Chen¹, Xixi Shen¹, Yanming Zhao¹, Qian Shi¹, Ling Feng¹, Shizhi He¹, Hongzhi Ma¹, Lizhen Hou¹, Qi Zhong¹, Hongbao Cao² & Jugao Fang¹✉

Sinonasal squamous cell carcinoma (SNSCC) is an aggressive cancer affecting the nasal and sinus regions, with its progression factors, particularly genetic ones, not yet fully understood. Here, we first conducted a retrospective study with 219 SNSCC patients to identify clinical factors affecting SNSCC prognosis. Additionally, we mined a vast literature dataset to uncover genetic factors associated with SNSCC progression. Based on this data, we constructed SNSCC prognosis pathways and performed a gene set enrichment analysis (GSEA). Clear operative margins were linked to a 73.5–86.3% improvement in overall survival and a 73.5–88.9% lower risk of recurrence. Nasal cavity-originated cases exhibited a 67.6–97.4% decrease in mortality and an 80.7–96.7% lower recurrence rate. Patients at T1-2 staging had a 65.0–80.6% reduced risk of death and recurrence compared to those at T3 stage. Additionally, we identified 53 genes associated with SNSCC, with 14 also implicated in primary tumor site, T stage, and operative margin. These genes, including EGFR, PIK3CA, ERBB2, PTEN, BCL2, BRAF, KRAS, and PRL, form a complex SNSCC-prognosis pathway and were significantly enriched in 42 KEGG pathways and Gene Ontology (GO) terms (FDR-corrected p-value < 0.001), influencing cell growth, apoptosis, and oncogenic signaling pathways. Our study suggests that three clinical parameters (operative margin type, primary tumor site, and T-stage) and 14 genetic factors may influence SNSCC prognosis post-surgery. These findings deepen our understanding of SNSCC and offer potential avenues to enhance its treatment and outcomes.

Keywords SNSCC, Overall survival, Recurrence-free survival, Genetic factor, Prognosis analysis

The malignancy of malignant tumors in the nasal and sinus regions is notably elevated, with a global incidence rate ranging from 0.2 to 0.8% of all malignant tumors, and comprising approximately 3% of malignancies in the upper respiratory tract¹. Within this category, sinonasal squamous cell carcinoma (SNSCC) stands out as the most prevalent histological subtype. SNSCC primarily originates from the mucosal lining of the nasal cavity and paranasal sinuses, the air-filled spaces adjacent to the nose. While relatively uncommon, this malignancy is characterized by its aggressive nature².

The precise etiology of SNSCC may not always be definitively established, yet it is frequently linked to risk factors such as occupational exposure to specific hazards (e.g., wood dust, nickel, and formaldehyde), tobacco usage, and the presence of chronic sinusitis^{3–5}. Typical manifestations encompass nasal congestion, epistaxis, facial discomfort, and alterations in olfactory perception. The diagnostic process typically entails imaging modalities such as CT scans and MRI, along with confirmatory biopsy procedures².

Treatment modalities of SNSCC can encompass a combination of surgical procedures, radiation therapy, and chemotherapy. Presently, the National Comprehensive Cancer Network (NCCN) guidelines endorse holistic treatment approaches that integrate surgery with radiotherapy and chemotherapy^{6,7}. The primary objective of surgical intervention is the complete eradication of the tumor. Over the years, head and neck surgeons have been actively investigating surgical treatments, with particular emphasis in the last decade on the advancements in minimally invasive surgery and early screening techniques, which have led to gradual modifications in

¹Department of Otorhinolaryngology Head and Neck Surgery, Beijing Tongren Hospital, Capital Medical University, Beijing 100730, China. ²School of Systems Biology, George Mason University, Manassas, VA 20110, USA. ✉email: lianmeng19861222@163.com; fangjugao2@ccmu.edu.cn

conventional treatment strategies. Nevertheless, there is currently no consensus on the standardized selection of surgical methods.

The outlook for SNSCC remains bleak, with a 5-year survival rate averaging around 50%, and there has been little discernible improvement over several decades⁸. Furthermore, prognosis is subject to variation based on factors such as the cancer's stage at diagnosis, the extent of tumor involvement, and the efficacy of treatment^{9,10}. Recent studies have suggested a connection between certain genetic factors, such as MIR21, PTEN, and EGFR, and the unfavorable prognosis of SNSCC^{11,12}. Nevertheless, due to the rarity of this disease, only a limited number of factors have been investigated and reported to influence the progression of SNSCC.

This study seeks to investigate and assess the influence of clinical and genetic factors pertaining to surgical interventions on the survival and prognosis of individuals with SNSCC. Our research has the potential to offer novel insights into understanding the disease, ultimately leading to enhanced treatment strategies for SNSCC.

Materials and methods

Participants for clinical data collection

From December 2004 to April 2021, a total of 219 patients with nasal and sinus squamous cell carcinoma underwent surgical resection at the Department of Head and Neck Surgery, Beijing Tongren Hospital, affiliated with Capital Medical University. Approval for this study was obtained from the ethics committee of Beijing Tongren Hospital. Written consent was obtained from all participants, indicating their agreement to allow the use of their data for the purposes of this research and any related publications. All methods were performed in accordance with the relevant guidelines and regulations.

The inclusion and exclusion criteria for the participants are described as follows. Inclusion Criteria: Patients diagnosed with nasal and paranasal sinus squamous cell carcinoma in our hospital between December 1, 2004, and April 30, 2021, who underwent pathological confirmation and had complete clinical and pathological data. Exclusion Criteria: (1) Patients with a history of other malignant tumors or a history of severe uncontrolled infections, major organ dysfunction, and other relevant medical histories. (2) Patients with misdiagnosis, mistaken inclusion, or those with incomplete or no recorded research data.

The relevant clinical information were collected, including age at surgery, primary tumor site, type of surgery, margin status, T-stage, N-stage, and prognosis-related information. Pathological assessment of tumors was conducted by physicians with senior-level qualifications in the Pathology Department. Clinical TNM staging was determined based on the American Joint Committee on Cancer (AJCC)'s 2018 staging criteria. All examinations and treatments in this study complied with the requirements of the Helsinki Declaration.

Treatment methods

We describe the treatment methods for the participants as follows, including surgical, radiation therapy, and chemotherapy approaches.

Surgery

Surgical approaches include endoscopic surgery and open surgery. Depending on the location of the primary lesion, extent of involvement, preservation of function, and reasonable repair considerations, techniques such as endoscopic tumor resection, total maxillectomy, lateral nasal resection, craniofacial combined surgical approaches, and appropriate defect repair are employed. For patients with confirmed or suspected neck lymph node metastasis, selective or radical neck dissection is performed.

Radiation therapy

For patients with early-stage lesions (T1-T2N0M0), postoperative close follow-up observation is conducted if the surgical margins are negative. If the surgical margins are positive or inadequate, or if there is nerve or vascular involvement, postoperative radiation therapy or concurrent chemoradiotherapy is necessary. For patients with intermediate to advanced-stage lesions (T3-T4aN0M0) without lymph node metastasis, postoperative radiation therapy or concurrent chemoradiotherapy is administered. For patients with advanced-stage lesions (T4b), postoperative concurrent chemoradiotherapy is indicated. For patients with neck lymph node metastasis (any T stage), postoperative radiation therapy or concurrent chemoradiotherapy is required, and the radiation field includes the tumor bed and the lymphatic drainage area of the neck. Three-dimensional conformal radiotherapy or intensity-modulated radiotherapy (IMRT) is utilized, with conventional fractionation of 2 Gy per fraction, 5 times per week, and a total radiation dose of 50 to 70 Gy. Primary lesions are irradiated with 2 or 3 fields depending on the extent of the lesion. For ethmoid sinus lesions, a combination of 6 MV-X-ray and 9–12 MeV electron beam irradiation is used to reduce the dose to the optic nerve and optic chiasm, and a wedge filter is used to improve coverage for patients with extensive involvement of nasal lesions. Prophylactic irradiation is administered to the upper neck in cases of neck lymph node metastasis, with a dose of 55 to 60 Gy, and prophylactic irradiation is given to the lower neck on the same side at a dose of 50 to 55 Gy. Treatment plans are evaluated using dose-volume histogram (DVH), and the treatment plan is standardized to cover at least 95% of the planning target volume (PTV) range.

Chemotherapy

Concurrent chemotherapy includes platinum-based monotherapy or TP regimen. The dosage of chemotherapy drugs is adjusted appropriately based on the degree of bone marrow suppression. Non-concurrent chemotherapy uses TP or TPF regimens. For patients with a larger preoperative range of primary lesions, induction chemotherapy before surgery may be considered.

Identify clinical parameters relevant to SNSCC progression

SAS (version 9.4) and R (version 4.2) were used for statistical analysis and result visualization. The Kaplan–Meier (KM) method¹³ was used to estimate the median and 95% confidence intervals for overall survival and recurrence-free survival based on factors such as age at surgery, primary tumor site, surgical type, margin type, T stage, and N stage. Kaplan–Meier survival curves were plotted accordingly. The KM curve was presented for each analysis to provide estimated survival rates at various time points (landmark rates), with time on the horizontal axis and corresponding survival rates on the vertical axis. The log-rank test was used to compare differences between groups for various factors, and the Cox proportional hazards model was employed to estimate hazard ratios (HR) and their 95% confidence intervals for overall survival and recurrence-free survival.

Identify genetic factors and pathways relevant to SNSCC progression

We employed natural language processing (NLP)-based literature data mining (LDM) to uncover genetic factors associated with the progression of SNSCC. Initially, we identified genes related to general squamous cell carcinoma (SCC). Subsequently, these genes underwent further analysis using an AI-based NLP tool from AIC LLC (<https://www.gousinfo.com/en/userguide.html>) to assess their connection with SNSCC and its pertinent clinical factors. The NLP analysis encompassed a dataset of 35 million PubMed articles. Genes demonstrating associations with both SNSCC and all relevant clinical parameters related to SNSCC progression were utilized in the construction of pathways linked to SNSCC prognosis. Additionally, we conducted gene set enrichment analysis (GSEA) using DAVID (<https://david.ncifcrf.gov>) to gain insights into the functional roles of the identified gene set in SNSCC progression.

Results

Patient treatment and prognosis

Among the 219 patients, 155 were males and 64 were females. There were 148 primary cases and 71 recurrent cases. The patients' ages ranged from 16 to 82 years, with a median age of 54 years. Tumor stages were distributed as follows: 12 patients with T1 stage, 55 with T2 stage, 77 with T3 stage, and 75 with T4 stage. Lymph node status was categorized as N0 in 199 cases, N1 in 6 cases, N2 in 11 cases, and N3 in 3 cases. The primary tumor sites included the nasal cavity (66 cases), ethmoid sinus (26 cases), maxillary sinus (111 cases), pterygopalatine fossa (2 cases), and other locations (14 cases). Endoscopic surgery was performed in 88 cases, and open surgery in 131 cases. Surgical margins were positive in 66 cases, clear in 131 cases, close in 11 cases, and unknown in 11 cases. There were 127 cases of tumor-free survival, 15 cases with-tumor survival, 68 cases of death, and 9 cases lost to follow-up. Due to variations in surgery dates and different outcome events (such as death and loss to follow-up, which marked the end of tracking), the follow-up durations varied. The range of follow-up time is [2 months, 164 months], with an average of 43.95 months and a median of 30 months. The average overall survival was 92.3 ± 5.21 months, and the average recurrence-free survival was 87.5 ± 4.92 months. Table 1 presents the statistics for patients with different primary tumor sites and the corresponding stage information.

Survival analysis of different types of margins

We defined a negative margin with a safety margin > 5 mm as a clear margin and a negative margin with a safety margin < 5 mm but > 2 mm as a close margin. After excluding 11 cases with unknown margins, a comparison was made of the overall survival of 208 patients with different types of margins (66 cases with positive margins, 131 cases with clear margins, and 11 cases with close margins), and the results are as follows: Compared to positive margins, the median survival period for clear margins and close margins was 147 months and 25 months, respectively, representing an 86.3% increase in overall survival (HR = 0.137, 95% CI: 0.077–0.243, $P < 0.0001$), with significant differences. Compared to close margins, the median survival period for clear margins and close margins was 147 months and 39 months, respectively, representing a 73.5% increase in overall survival (HR = 0.210, 95% CI: 0.076–0.580, $P < 0.01$), with significant differences. The Kaplan–Meier survival curve for overall survival is shown in Fig. 1a.

In the analysis of recurrence-free survival of different types of margins, compared to positive margins, the median recurrence-free survival for clear margins and close margins was 140 months and 12 months, respectively, representing an 88.9% reduction in the risk of recurrence (HR = 0.111, 95% CI: 0.064–0.192, $P < 0.0001$). The median recurrence-free survival for clear margins compared to close margins (median recurrence-free survival not yet reached) was 140 months, with HR = 0.263 (95% CI: 0.088–0.780, $P < 0.01$), representing a 73.7% reduction in the risk of recurrence, and the difference was statistically significant. The Kaplan–Meier survival curve for recurrence-free survival is shown in Fig. 1b.

Primary tumor site	#Patients	T1/T2/T3/T4	N0/N1/N2/N3
Nasal cavity	66	11/27/20/8	52/3/8/3
Ethmoid sinus	26	1/11/6/8	26/0/0/0
Maxillary sinus	111	0/17/51/43	106/3/2/0
Pterygopalatine fossa	2	0/0/0/2	1/0/1/0
Other sites	14	0/0/0/14	14/0/0/0

Table 1. Location and corresponding stage. Note: 'other sites' of squamous cell carcinoma include the nasal cavity and paranasal sinuses, originating from the frontal sinus, sphenoid sinus, lateral orbital wall, and the infratemporal fossa.

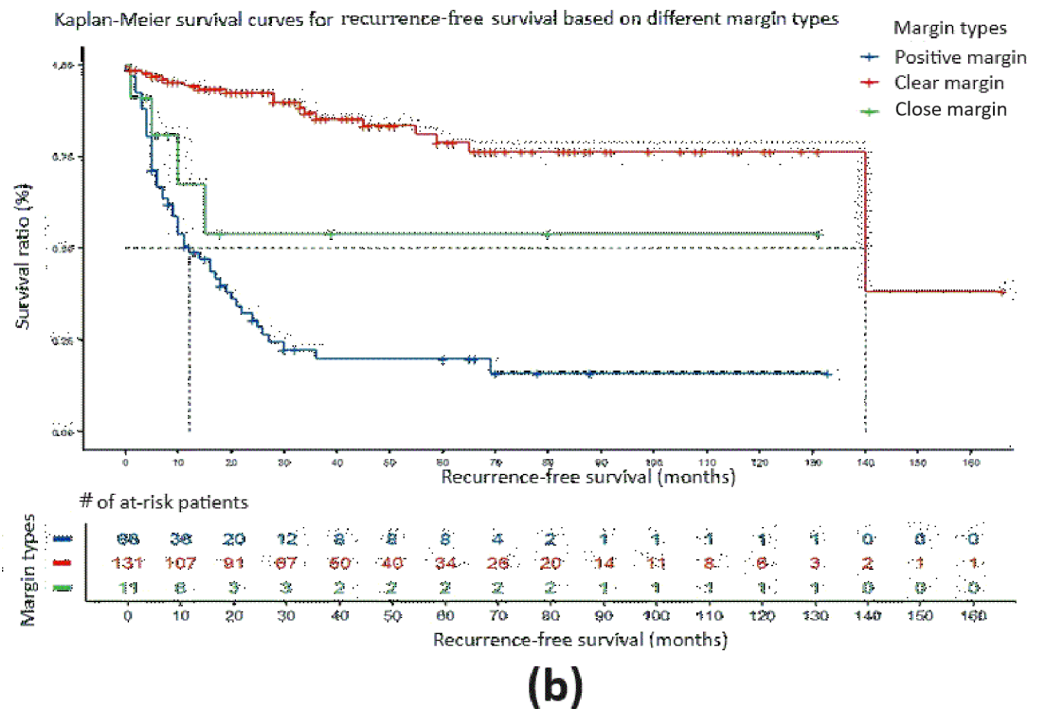
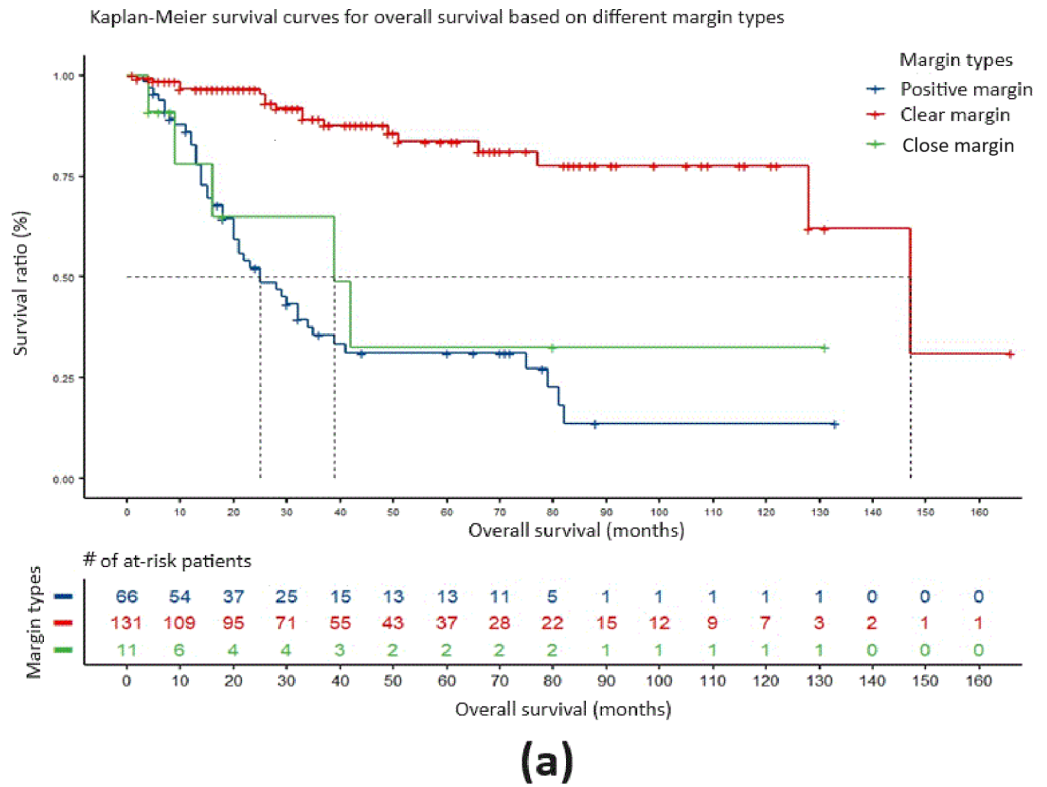


Fig. 1. The impact of types of margins on the overall survival and recurrence-free survival of SNSCC patients. (a) Kaplan–Meier survival curve comparing the overall survival period for different types of margin. (b) Kaplan–Meier survival curve for recurrence-free survival period for different types of margin.

There were no significant differences in overall survival and recurrence-free survival between close margins and positive margins.

Survival analysis for different primary sites

In the analysis comparing different primary tumor sites, for patients with squamous cell carcinoma originating in the nasal cavity, the overall survival (median survival not reached) and disease-free survival (median disease-free survival not reached) were both significantly better than those for patients with squamous cell carcinoma originating in sites outside the nasal cavity. The hazard ratios (HRs) for risk were all significantly less than 1, and the *p*-values were all less than 0.05, indicating that patients with primary nasal cavity squamous cell carcinoma had superior overall survival and disease-free survival compared to patients with tumors originating in sites outside the nasal cavity. Here, the term “nasal cavity” was used to refer to tumors originating from the anatomical structures within the nasal cavity proper. This includes the nasal septum, the floor of the nasal cavity, and the lateral walls, which encompass the inferior, middle, and superior turbinates. Tumors specifically originating from the ethmoid sinus proper, excluding the turbinates, were classified under “ethmoid sinus” squamous cell carcinoma.

Compared to patients with squamous cell carcinoma originating in the nasal cavity, those with tumors originating in the sphenoid sinus (median survival 75 months, median disease-free survival 45 months), maxillary sinus (median survival 81 months, median disease-free survival 69 months), pterygopalatine fossa (median survival 16.5 months, median disease-free survival 14 months), and other sites (median survival 41 months, median disease-free survival 22 months) had a decreased risk of death by 67.6% (HR = 0.324, 95% CI: 0.130–0.806, *P* < 0.05), 70.2% (HR = 0.298, 95% CI: 0.140–0.636, *P* < 0.01), 97.4% (HR = 0.026, 95% CI: 0.004–0.189, *P* < 0.0001), and 77.3% (HR = 0.227, 95% CI: 0.077–0.664, *P* < 0.01), respectively. The risk of recurrence decreased by 80.7% (HR = 0.193, 95% CI: 0.072–0.515, *P* < 0.01), 81.6% (HR = 0.184, 95% CI: 0.079–0.431, *P* < 0.0001), 96.7% (HR = 0.033, 95% CI: 0.005–0.200, *P* < 0.0001), and 88.7% (HR = 0.113, 95% CI: 0.038–0.341, *P* < 0.0001), respectively. Kaplan–Meier survival curves are shown in Fig. 2a and b. Here, the ‘other sites’ of squamous cell carcinoma include the nasal cavity and paranasal sinuses, originating from the frontal sinus, sphenoid sinus, lateral orbital wall, and even the infratemporal fossa. These sites are grouped together because they are relatively rare, often invade the orbital contents and skull base, are difficult to completely remove surgically, and generally have a poor prognosis.

When comparing patients with squamous cell carcinoma originating in the sphenoid sinus, maxillary sinus, and other sites, there were no significant differences in overall survival and disease-free survival.

There were a total of 2 cases of squamous cell carcinoma originating in the pterygopalatine fossa. Both overall survival and disease-free survival were the poorest among these cases, but considering the small number of cases observed, the data may not be representative, and the results are provided for reference only.

The impact of tumor T staging on survival prognosis

In the analysis of T staging, the median survival period for T3 staging was 147 months, and the median recurrence-free survival period was 140 months. For T4 staging, the median survival period was 48 months, and the median recurrence-free survival period was 28 months. In comparison with patients with T1–2 staging, both the total survival period and recurrence-free survival period for patients with T3 and T4 staging had significantly higher hazard ratios, all greater than 1, and *P* values less than 0.01. This indicates that the overall survival period and recurrence-free survival period for patients with squamous cell carcinoma at T1–2 staging are better than those at T3 and T4 staging. Kaplan–Meier survival curves are shown in Fig. 3a,b.

Compared to T3 and T4 staging, the risk of death for T1–2 staging decreased by 71.6% (HR = 3.524, 95% CI = 1.509–8.230, *P* < 0.01) and 80.6% (HR = 5.156, 95% CI = 2.300–11.559, *P* < 0.0001), respectively. The risk of recurrence also decreased by 65.0% (HR = 2.859, 95% CI = 1.381–5.915, *P* < 0.01) and 74.8% (HR = 3.972, 95% CI = 1.972–7.998, *P* < 0.0001), respectively. However, there were no statistically significant differences in total survival period and recurrence-free survival period between T3 and T4 staging.

Impact of surgical type, surgical age, and N staging on survival prognosis

In the analysis of surgical age, it was found that the age group below 65 years had a relatively better overall survival compared to the age group above 65 years (HR = 0.386, 95% CI = 0.233–0.641, *P* < 0.01), while there was no significant difference in recurrence-free survival. Considering that elderly patients often have multiple underlying health conditions and poorer physical health, this result is more likely to be directly attributed to the age of the patients.

In the analysis of surgical type, there was no significant statistical difference in both overall survival and recurrence-free survival between endoscopic surgery and open surgery. In the analysis of N staging, there were no significant differences in both overall survival and recurrence-free survival among the different groups.

Genetic factors related to SNSCC prognosis

At the outset, an AI-driven LDM analysis identified a total of 963 genes associated with general Squamous Cell Carcinoma (SCC), backed by an extensive body of over 1000 references. Within this set of 963 genes, 54 were specifically linked to SNSCC, supported by 83 references. Additionally, 61 genes were found to be correlated with T-stage, supported by a substantial 300-plus references. Furthermore, 79 genes were replicated as associated with the primary tumor site, bolstered by more than 250 supporting references. Lastly, 24 genes were connected to operative margin type, supported by approximately 40 references. Notably, there were 14 genes that were concurrently implicated in SNSCC, primary tumor site, T-stage, and operative margin, as illustrated in Fig. 4.

According to the findings presented earlier, these 14 genes collectively formed an intricate genetic network that impacts the progression, overall survival, and likelihood of recurrence in SNSCC. This network, as depicted in Fig. 5, was created using Cytoscape (version 3.7.2) and comprises 20 nodes and 79 edges, with a shortest path length of 342. The 14 genes include BRAF, COL3A1, COX2, ERBB2, FOXP3, IDH1, KRAS, PCNA, PIK3CA, PRL, PTEN, RHO, BCL2, EGFR, and GSEA.

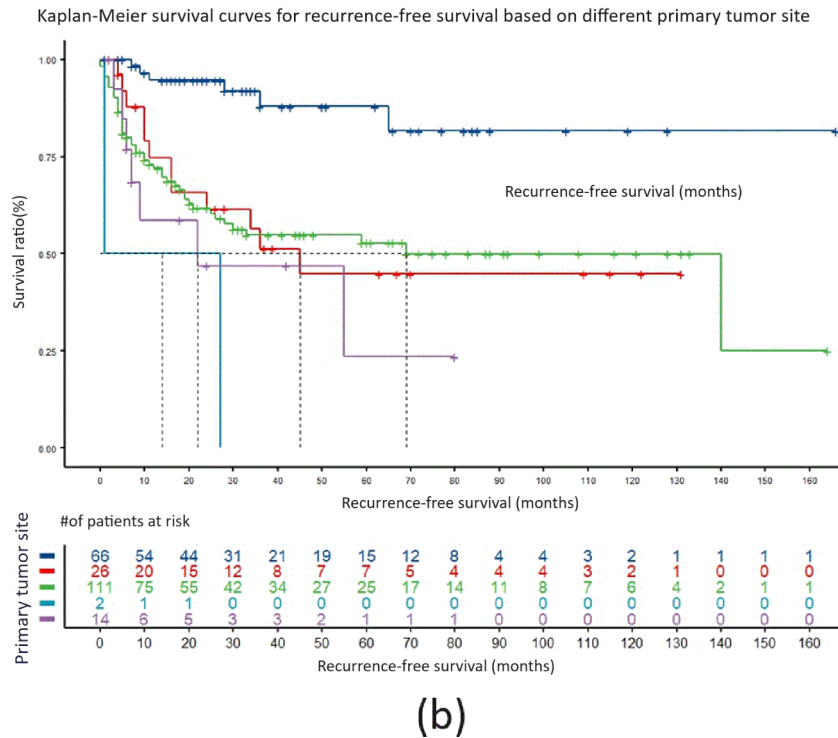
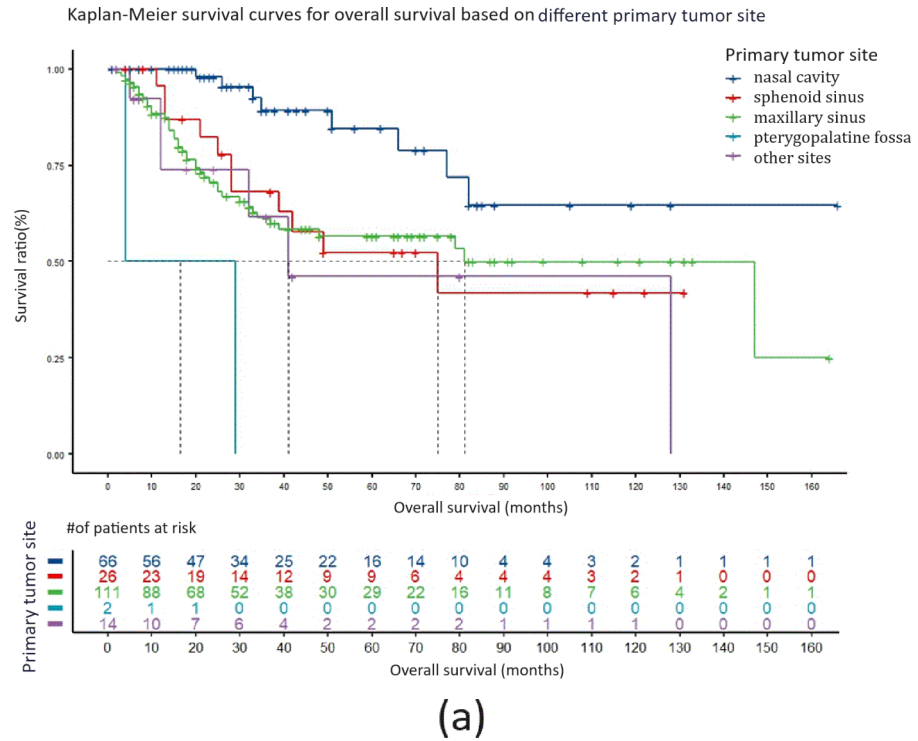
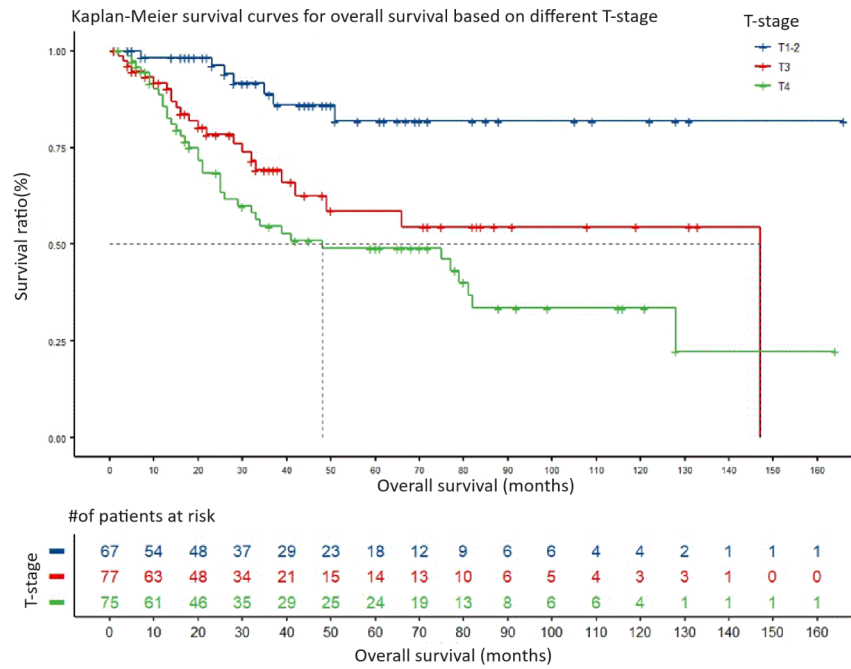


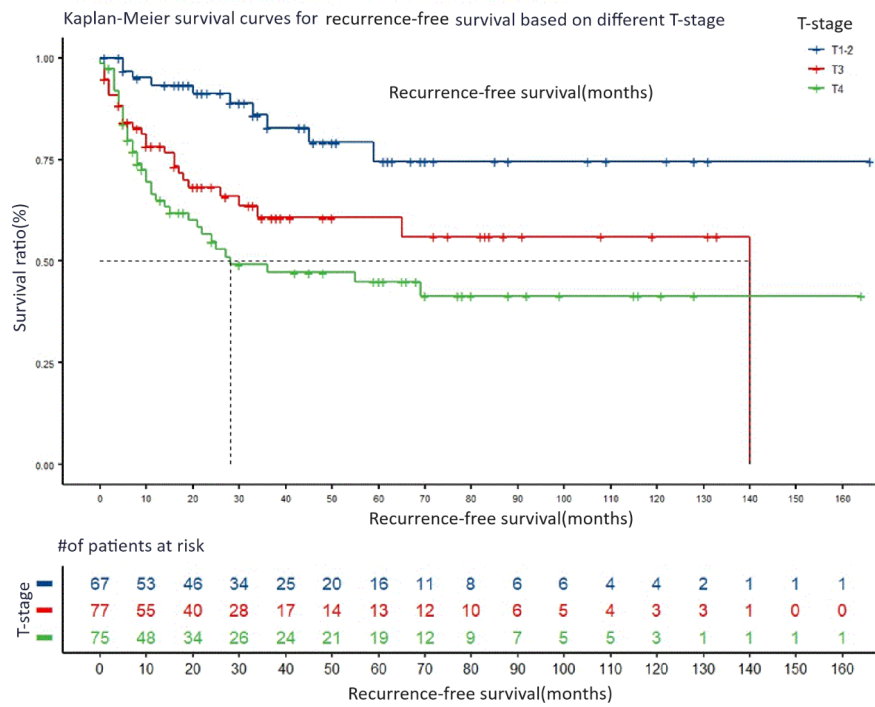
Fig. 2. The impact of primary tumor site on the overall survival and recurrence-free survival of SNSCC patients. **(a)** Kaplan–Meier survival curve comparing the overall survival period for different primary tumor sites. **(b)** Kaplan–Meier survival curve for recurrence-free survival period for different primary tumor sites.

Enrichment analysis of the 14 genes

The GSEA results demonstrated that 8 out of the previously identified 14 genes exhibited significant enrichment across 42 KEGG pathways and Gene Ontology (GO) terms (FDR corrected p-value < 0.001). Figure 6 displays the top 10 pathways, all of which belong to the KEGG category. In summary, the enrichment analysis reveals that these genes collectively play active roles in a wide range of cellular processes and signaling pathways, many of



(a)



(b)

Fig. 3. The impact of T-stage on the overall survival and recurrence-free survival of SNSCC patients. (a) Kaplan–Meier survival curve comparing the overall survival period in different T-stages. (b) Kaplan–Meier survival curve for recurrence-free survival period in different T-stages.

which are closely associated with the development and progression of cancer. When these genes are dysregulated, they have the potential to stimulate cell growth, suppress apoptosis, and initiate oncogenic signaling pathways. These findings suggest that these genes may have a substantial impact on the development and progression of SNSCC. Understanding the functions of these genes is pivotal within the context of SNSCC, and the elucidated

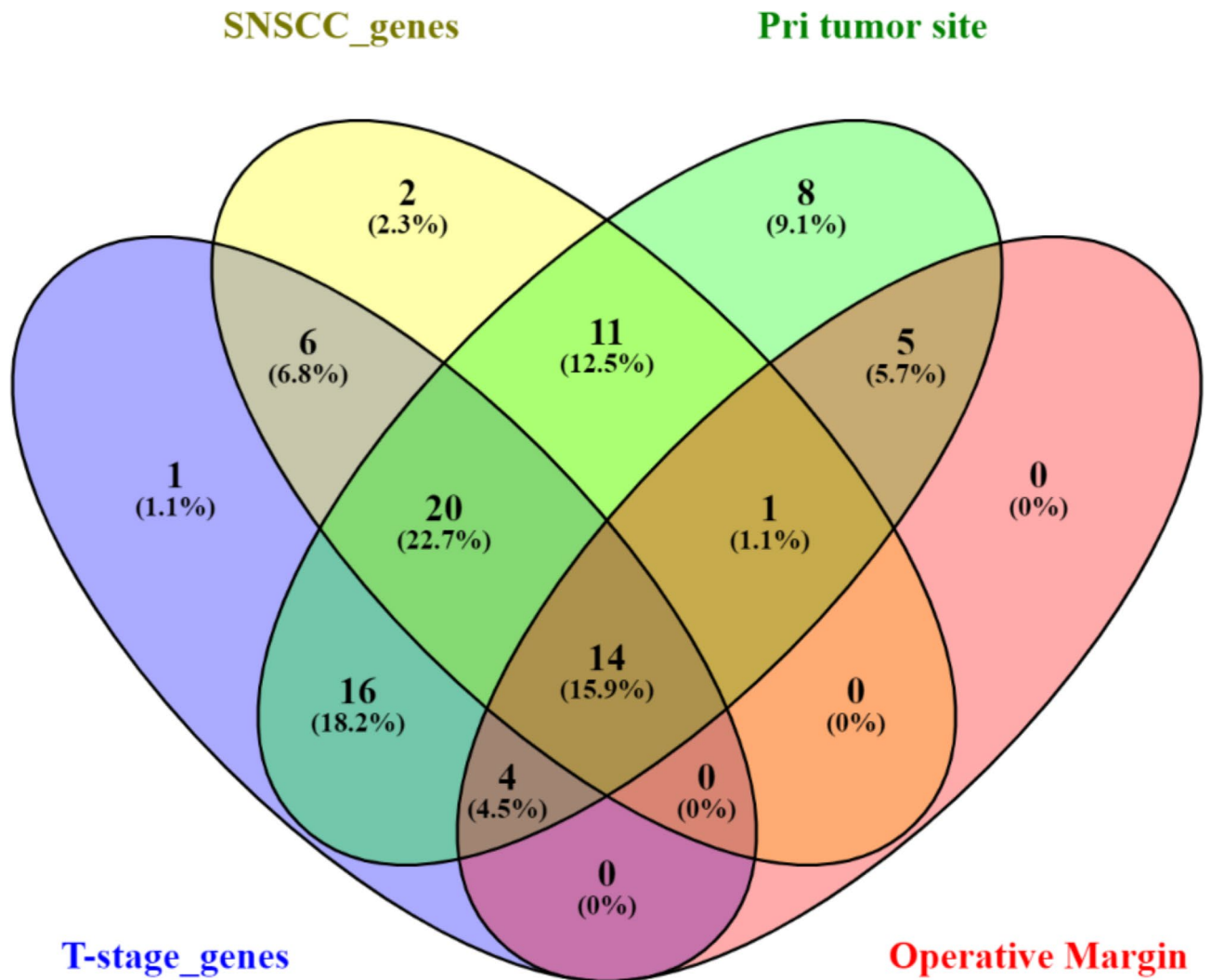


Fig. 4. The genetic factors identified potentially influencing the progression of SNSCC.

pathways provide valuable insights into the underlying molecular mechanisms of various cancers, including SNSCC.

Discussion

Sinonasal squamous cell carcinoma (SNSCC) is an uncommon but highly aggressive cancer that primarily affects the squamous cells lining the nasal and sinus passages. Given its rarity, only a limited number of factors have been documented to influence the prognosis of SNSCC, including the cancer stage at the time of diagnosis⁹, the extent of tumor involvement¹⁴, and the effectiveness of treatment¹⁰. In recent years, research has shed light on the contribution of genetic factors to the development of this uncommon head and neck cancer, such as EGFR and SMARCA4^{15,16}. Identifying and comprehending these factors related to SNSCC prognosis is imperative for enhancing the treatment strategies and overall outcomes associated with this disease.

In this study, we initially conducted a retrospective analysis aimed at identifying clinical prognostic factors related to SNSCC prognosis. The study examined six specific factors, which included the type of surgery, the primary location of the tumor, the surgical procedure performed, the nature of operative margins, the T-stage, and the N-stage. Our findings revealed significant associations between T-stage, the nature of operative margins, and the primary tumor site with both overall survival and the risk of recurrence in SNSCC patients.

In particular, cases with clear operative margins exhibited a notable improvement in overall survival and a reduced risk of recurrence when compared to cases with positive margins and close margins. This observation aligns with a prior study that reported poorer prognoses in SNSCC patients with positive margins¹⁰. Additionally, SNSCC patients classified under T1-2 stages exhibited a substantially reduced risk of death and recurrence, with risk reductions ranging from 65.0 to 80.6%, which is consistent with previous research findings¹⁷.

While maxillary sinus is the most common primary tumor site for SNSCC¹⁴, our study demonstrated that cases originating in the nasal cavity exhibited a decreased risk of death compared to cases originating in the ethmoid sinus, maxillary sinus, pterygopalatine fossa, and other sites, with risk reductions ranging from 67.6

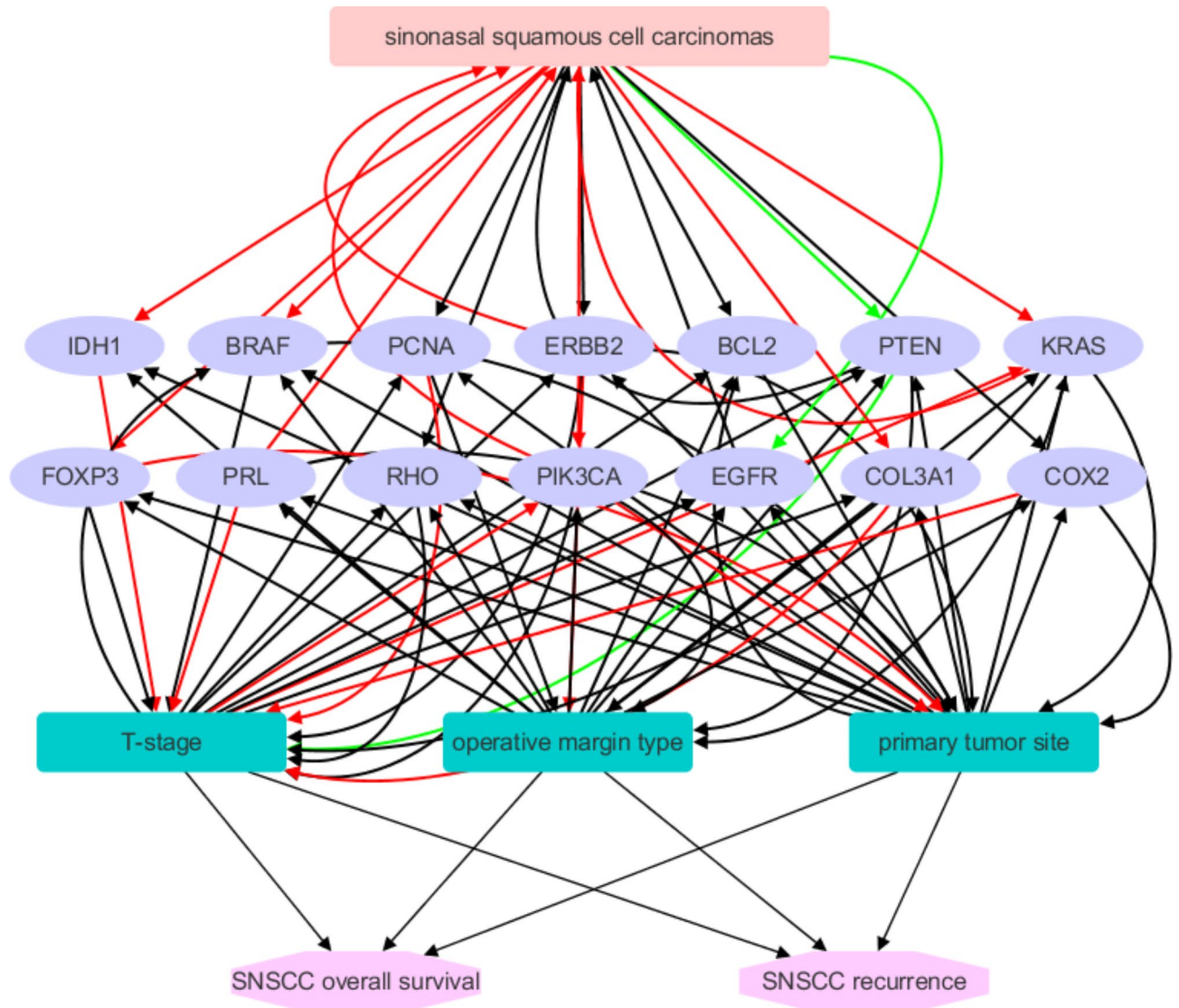


Fig. 5. Illustrates a genetic pathway comprising 14 genes that have been identified through the use of AI-based literature data mining as influential in determining the prognosis of SNSCC. Red edges signify positive effects, green edges denote negative effects, and black edges indicate effects that are currently unknown.

to 97.4%. Similarly, the risk of recurrence significantly decreased for cases originating in the nasal cavity, with reductions ranging from 80.7 to 96.7%. These results underscore the importance of considering the primary tumor site in treatment decisions, especially when it does not originate in the nasal cavity.

Our findings suggest that the type of surgery, the patient's age at surgery, and N staging have a relatively minor impact on the survival prognosis of SNSCC. Notably, N staging showed no significant correlation with overall survival and recurrence-free survival, contrary to previous studies linking it to orbital complications and influencing prognosis and recurrence^{18–20}. This discrepancy may be due to factors such as sample size and distribution, and treatment modalities. Most of the patients recruited in this study were in the N0 stage (199 out of 219; see Table 1). Additionally, the treatment modalities likely influenced the results. For patients with cervical lymph node metastasis ($N \neq 0$), cervical lymph node dissection was performed concurrently, followed by postoperative radiotherapy. The implementation of reasonable, standardized, and comprehensive treatment ensured that the N stage did not have a significant impact on survival prognosis. Thus, more extensive, multi-center studies with larger cohorts are necessary to validate or refute our findings.

Utilizing AI-based Natural Language Processing (NLP), we pinpointed 53 genes associated with SNSCC, and among them, 14 genes exhibited relevance to primary tumor site, T stage, and operative margin, including BRAF, COL3A1, COX2, ERBB2, FOXP3, IDH1, KRAS, PCNA, PIK3CA, PRL, PTEN, RHO, BCL2, EGFR, and GSEA. Our hypothesis posits that these genes may exert an impact on the overall survival and recurrence rates of SNSCC following surgical intervention. Through functional enrichment analysis, it becomes evident that these genes collectively participate in a range of cellular processes and signaling pathways, many of which have direct implications in the development and progression of cancer. When these genes are dysregulated, they have

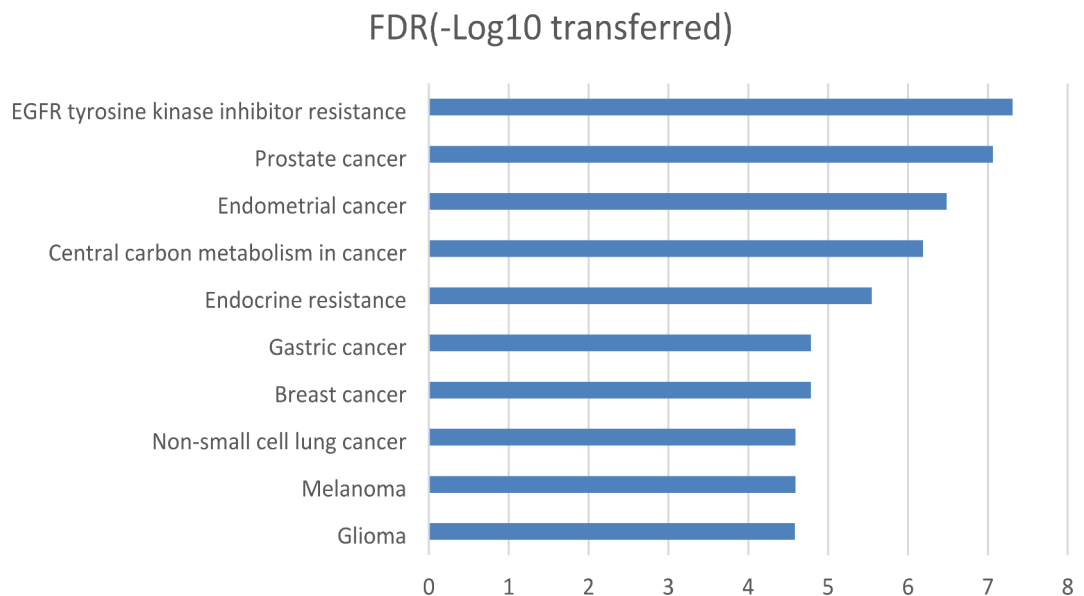


Fig. 6. The top 10 pathways enriched by the 14 genes that have been identified through the use of AI-based literature data mining as influential in determining the prognosis of SNSCC.

the potential to stimulate cell growth, hinder apoptosis, and initiate oncogenic signaling. These discoveries lend support to our hypothesis that these genes could serve as genetic factors influencing the prognosis of SNSCC. The comprehension of the roles played by these genes assumes critical significance in the context of SNSCC and related malignancies.

It is worth noting that 8 out of the 14 genes deserve special attention, as they exhibited enrichment in the top pathways identified through the enrichment analysis. These 8 genes are EGFR, PIK3CA, ERBB2, PTEN, BCL2, BRAF, KRAS, and PRL.

Specifically, EGFR (Epidermal Growth Factor Receptor) is a central player in several pathways, including EGFR tyrosine kinase inhibitor resistance, PI3K-Akt signaling pathway, ErbB signaling pathway, and numerous pathways related to cancer. EGFR plays a pivotal role in regulating cell growth, proliferation, and survival. It has been reported that EGFR-activating mutations, specifically exon 20 insertions, have been identified in SNSCC, indicating that EGFR plays a role in driving the pathogenesis of this rare form of head and neck cancer¹⁵.

PIK3CA (Phosphatidylinositol-4,5-Bisphosphate 3-Kinase Catalytic Subunit Alpha) is another key contributor to multiple cancer-related and signaling pathways. It is an integral component of the PI3K-Akt signaling pathway and is involved in promoting cell growth and survival. It has been shown that SNSCC with PIK3CA mutations have a higher rate of lymph node or distant metastases and shorter survival, indicating the importance of PIK3CA in the metastasis of SNSCC²¹.

ERBB2 (Erb-B2 Receptor Tyrosine Kinase 2), also known as HER2, participates in various cancer-related pathways, including EGFR tyrosine kinase inhibitor resistance, prostate cancer, endometrial cancer, and breast cancer. This receptor tyrosine kinase is instrumental in facilitating cell growth and division. Patients with SNSCC may have alterations in the ERBB2 gene, suggesting that targeting ERBB2 could be a promising therapeutic option for SNSCC²².

PTEN (Phosphatase and Tensin Homolog) is implicated in pathways associated with cancer, such as EGFR tyrosine kinase inhibitor resistance, central carbon metabolism in cancer, endocrine resistance, and PI3K-Akt signaling pathway. PTEN acts as a tumor suppressor by regulating cell cycle progression and inhibiting pathways that promote cell survival. The PI3K-Akt signaling pathway is commonly altered in a subgroup of SNSCC, with loss of PTEN expression and overexpression of pS6 being associated with unfavorable clinical outcomes¹².

BCL2 (B-Cell Lymphoma 2) is found within pathways related to cancer, including EGFR tyrosine kinase inhibitor resistance and endocrine resistance. Its function involves inhibiting apoptosis (programmed cell death) and supporting cell survival. BRAF (B-Raf Proto-Oncogene, Serine/Threonine Kinase) is associated with various cancer pathways, including prostate cancer, endometrial cancer, and melanoma. It operates as a kinase involved in cell signaling and can drive cell proliferation when mutated. Zhao et al. showed that miR-34a could inhibit migration and invasion of SNSCCs by targeting BCL2²³.

KRAS (Kirsten Rat Sarcoma Viral Oncogene Homolog) is a participant in pathways such as EGFR tyrosine kinase inhibitor resistance, central carbon metabolism in cancer, and others. Mutations in KRAS can lead to uncontrolled cell growth and contribute to cancer development. SNSCC and KRAS mutations are frequently associated with sinonasal papillomas, with KRAS mutations being characteristic of oncocytic sinonasal papillomas²⁴. PRL (Prolactin) plays a role in the PI3K-Akt signaling pathway, impacting cell proliferation and survival. It is associated with breast cancer and other signaling pathways.

It should be noted that we applied stringent criteria in our NLP to select the most relevant genetic factors related to SNSCC progression. With less stringent criteria, additional genetic factors could have been included

for further discussion. For example, although the TP53 gene was not found to be associated with operative margin type, it was identified for its role in SNSCC pathology and its association with T-stage and primary tumor site. Some SNSCC patients have been found to have TP53 mutations²⁵, and these mutations are associated with shorter survival²¹.

Conclusion

Our study proposes that three clinical parameters (operative margin type, primary tumor site, and T-stage) and 14 genetic factors may have an influence on the post-surgical prognosis of SNSCC. These findings could contribute to a deeper understanding of the disease and offer potential avenues for improving the treatment and outcomes of SNSCC.

Data availability

The data are available from the corresponding author upon request.

Received: 13 June 2024; Accepted: 13 September 2024

Published online: 27 September 2024

References

1. Oliver, J. R. et al. Human papillomavirus and survival of patients with sinonasal squamous cell carcinoma. *Cancer* **126**, 1413–1423. <https://doi.org/10.1002/cncr.32679> (2020).
2. Lewis, J. S. Jr. Sinonasal squamous cell carcinoma: a review with emphasis on emerging histologic subtypes and the role of human papillomavirus. *Head Neck Pathol.* **10**, 60–67. <https://doi.org/10.1007/s12105-016-0692-y> (2016).
3. d'Errico, A. et al. Exposure to occupational hazards and risk of sinonasal epithelial cancer: results from an extended Italian case-control study. *Occup. Environ. Med.* <https://doi.org/10.1136/oemed-2020-106738> (2020).
4. Bonzini, M. et al. Prevalence of occupational hazards in patients with different types of epithelial sinonasal cancers. *Rhinology* **51**, 31–36. <https://doi.org/10.4193/Rhino11.228> (2013).
5. Nokovitch, L. et al. Oral cavity squamous cell carcinoma risk factors: state of the art. *J. Clin. Med.* **12**. <https://doi.org/10.3390/jcm12093264> (2023).
6. Thawani, R. et al. The contemporary management of cancers of the sinonasal tract in adults. *CA Cancer J. Clin.* **73**, 72–112. <https://doi.org/10.3322/caac.21752> (2023).
7. Caudell, J. J. et al. NCCN guidelines(R) insights: head and neck cancers, version 1.2022. *J. Natl. Compr. Cancer Netw.* **20**, 224–234. <https://doi.org/10.6004/jnccn.2022.0016> (2022).
8. Turner, J. H. & Reh, D. D. Incidence and survival in patients with sinonasal cancer: a historical analysis of population-based data. *Head Neck* **34**, 877–885. <https://doi.org/10.1002/hed.21830> (2012).
9. Lin, N. et al. A clinical-radiomics nomogram based on the apparent diffusion coefficient (ADC) for individualized prediction of the risk of early relapse in advanced sinonasal squamous cell carcinoma: a 2-year follow-up study. *Front. Oncol.* **12**, 870935. <https://doi.org/10.3389/fonc.2022.870935> (2022).
10. Devaraja, K., Sikka, K., Kumar, R. & Thakar, A. Sinonasal malignancies: long term follow up after surgical management-an analysis of outcomes. *Indian J. Otolaryngol. Head Neck Surg.* **67**, 28–33. <https://doi.org/10.1007/s12070-014-0742-2> (2015).
11. Kovarikova, H. et al. Deregulation of selected microRNAs in sinonasal carcinoma: value of miR-21 as prognostic biomarker in sinonasal squamous cell carcinoma. *Head Neck* **39**, 2528–2536. <https://doi.org/10.1002/hed.24930> (2017).
12. Munoz-Cordero, M. G. et al. Predictive value of EGFR-PI3K-pAKT-mTOR-pS6 pathway in sinonasal squamous cell carcinomas. *Acta Otorrinolaringol. Esp. (Engl Ed)* **70**, 16–24. <https://doi.org/10.1016/j.otorri.2017.10.005> (2019).
13. Strum, S. et al. Real-world experience with Cemiplimab in Advanced Cutaneous squamous cell carcinoma. *J. Cutan. Med. Surg.* <https://doi.org/10.1177/12034754241265696> (2024).
14. Tsuji, T. et al. A multi-institutional retrospective study of 340 cases of sinonasal malignant tumor. *Auris Nasus Larynx.* <https://doi.org/10.1016/j.anl.2023.05.002> (2023).
15. Pacini, L., Cabal, V. N., Hermesen, M. A. & Huang, P. H. EGFR exon 20 insertion mutations in Sinonasal squamous cell carcinoma. *Cancers (Basel).* <https://doi.org/10.3390/cancers14020394> (2022).
16. Agaimy, A. Sinonasal neoplasms: update from the WHO 2022. *Pathologie (Heidelberg)* **44**, 233–239. <https://doi.org/10.1007/s00292-023-01202-8> (2023).
17. Nguyen, E. S. et al. Prognostic factors and outcomes of De Novo Sinonasal squamous cell carcinoma: a systematic review and meta-analysis. *Otolaryngol. Head Neck Surg.* **166**, 434–443. <https://doi.org/10.1177/01945998211021023> (2022).
18. Zhao, J. et al. Orbital and periocular complications in patients with sinonasal tumours with orbital invasion. *Br. J. Ophthalmol.* **108**, 465–470. <https://doi.org/10.1136/bjo-2022-322855> (2024).
19. Pala, M., Vrana, A., Novakova, P., Drbohlovova, T. & Podlesak, T. Long-term results of postoperative and definitive (chemo) radiotherapy in sinonasal carcinoma. Adult comorbidity evaluation 27 score as a predictor of survival. *Rep. Pract. Oncol. Radiother.* **28**, 147–158. <https://doi.org/10.5603/RPOR.a2023.0017> (2023).
20. Anschuetz, L. et al. Sinonasal malignancies: histopathological entities, regional involvement and long-term outcome. *J. Otolaryngol. Head Neck Surg.* **52**. <https://doi.org/10.1186/s40463-023-00627-8> (2023).
21. Doescher, J. et al. Epstein-Barr virus infection is strictly associated with the metastatic spread of sinonasal squamous-cell carcinomas. *Oral Oncol.* **51**, 929–934. <https://doi.org/10.1016/j.oraloncology.2015.07.008> (2015).
22. Lopez, F. et al. Gene amplification and protein overexpression of EGFR and ERBB2 in sinonasal squamous cell carcinoma. *Cancer* **118**, 1818–1826. <https://doi.org/10.1002/cncr.26451> (2012).
23. Zhao, Y. & Wang, X. miR-34a targets BCL-2 to suppress the migration and invasion of sinonasal squamous cell carcinoma. *Oncol. Lett.* **16**, 6566–6572. <https://doi.org/10.3892/ol.2018.9427> (2018).
24. Maisch, S. et al. Sinonasal papillomas: a single centre experience on 137 cases with emphasis on malignant transformation and EGFR/KRAS status in carcinoma ex papilloma. *Ann. Diagn. Pathol.* **46**, 151504. <https://doi.org/10.1016/j.anndiagnpath.2020.151504> (2020).
25. Garcia-Inclan, C. et al. Establishment and genetic characterization of six unique tumor cell lines as preclinical models for sinonasal squamous cell carcinoma. *Sci. Rep.* **4**, 4925. <https://doi.org/10.1038/srep04925> (2014).

Author contributions

ML, and JF designed the study, collected and analyzed the data, and wrote the manuscript. All authors reviewed the manuscript.

Funding

This work was supported by the Beijing Municipal Administration of Hospitals' Youth Programme (QML20210203), R&D Program of Beijing Municipal Education Commission (KM202210025014), National Key R&D Program of China (2020YFB1312805), and Capital Health Research and Development of Special (2022-1-2051).

Declarations

Competing interests

The authors declare no competing interests.

Ethics approval

Approval for this study was obtained from the ethics committee of Beijing Tongren Hospital.

Consent to participate

Written consent was obtained from all participants, indicating their agreement to allow the use of their data for the purposes of this research and any related publications.

Additional information

Correspondence and requests for materials should be addressed to M.L. or J.F.

Reprints and permissions information is available at www.nature.com/reprints.

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Open Access This article is licensed under a Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 International License, which permits any non-commercial use, sharing, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if you modified the licensed material. You do not have permission under this licence to share adapted material derived from this article or parts of it. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by-nc-nd/4.0/>.

© The Author(s) 2024