"AI- and LLM-Driven Drug Discovery Pipeline for Predictive Evaluation of Nimbolide and Salannin in Oral Squamous Cell Carcinoma"

Nikita Walunjkar[1], Dr. Pravin Badhe[2], Dr. Jaymala Kumawat[3], Prof. Snehal Hase[4]

[1] Final Year Student, School of Pharmacy, Vishwakarma University, Pune, Maharashtra, India

[2] Swalife Biotech Pvt. Ltd.

[3, 4]Assistant Professor, School of Pharmacy, Vishwakarma University, Pune, Maharashtra, India

Corresponding author Email: nikita.walunjkar9390@gmail.com
Doi: https://doi.org/10.5281/zenodo.20180079
Received: 22 April 2026
Accepted: 28 April 2026

ABSTRACT

Background: Oral squamous cell carcinoma (OSCC) is a major malignancy associated with complex molecular pathways involving inflammation, oxidative stress, immune dysregulation, and abnormal cell proliferation. Conventional drug discovery approaches are often time-consuming and less effective for multi-target diseases. Recent advances in artificial intelligence (AI) and large language models (LLMs) have enabled systematic exploration of phytochemicals such as neem-derived compounds.

Objective: To develop an AI- and LLM-driven integrative research pipeline using the Swalife platform to evaluate the therapeutic potential of neem bioactives, particularly nimbolide and salannin, in OSCC.

Methods: A multi-stage AI-assisted workflow was implemented using literature mining from PubMed, Google Scholar, and ScienceDirect, followed by LLM-assisted data extraction using Perplexity AI. Six modules were applied: target identification, lead optimization, in vitro design, in vivo design, clinical pharmacovigilance, and market/IPR analysis. Outputs were structured into HTML datasets and integrated into predictive and preventive medicine frameworks.

Results: Nimbolide and salannin showed multi-target activity across major OSCC pathways, including NF-κB, STAT3, PI3K/Akt, and EGFR. Approximately 70–80% target overlap was observed between neem bioactives and OSCC molecular pathways. Predictive modeling confirmed strong anti-inflammatory, pro-apoptotic, anti-angiogenic, and immunomodulatory effects.

Conclusion: The AI- and LLM-driven Swalife pipeline effectively bridges mechanistic discovery and predictive medicine. Nimbolide and salannin demonstrate strong potential as multi-target adjunct therapies in OSCC, supporting further experimental and clinical validation.

Keywords: Oral squamous cell carcinoma, Neem, Nimbolide, Salannin, Artificial intelligence, Predictive medicine

INTRODUCTION

Oral squamous cell carcinoma (OSCC) accounts for nearly 90% of oral malignancies and remains a major global health burden due to late diagnosis, high recurrence rates, and poor survival outcomes[1,2]. The pathogenesis of OSCC is multifactorial, involving genetic mutations, chronic inflammation, oxidative stress, and dysregulation of key signaling pathways such as EGFR, PI3K/Akt, MAPK, NF-κB, and STAT3[3,4].These pathways collectively promote tumor growth, angiogenesis, immune evasion, and resistance to apoptosis[4].

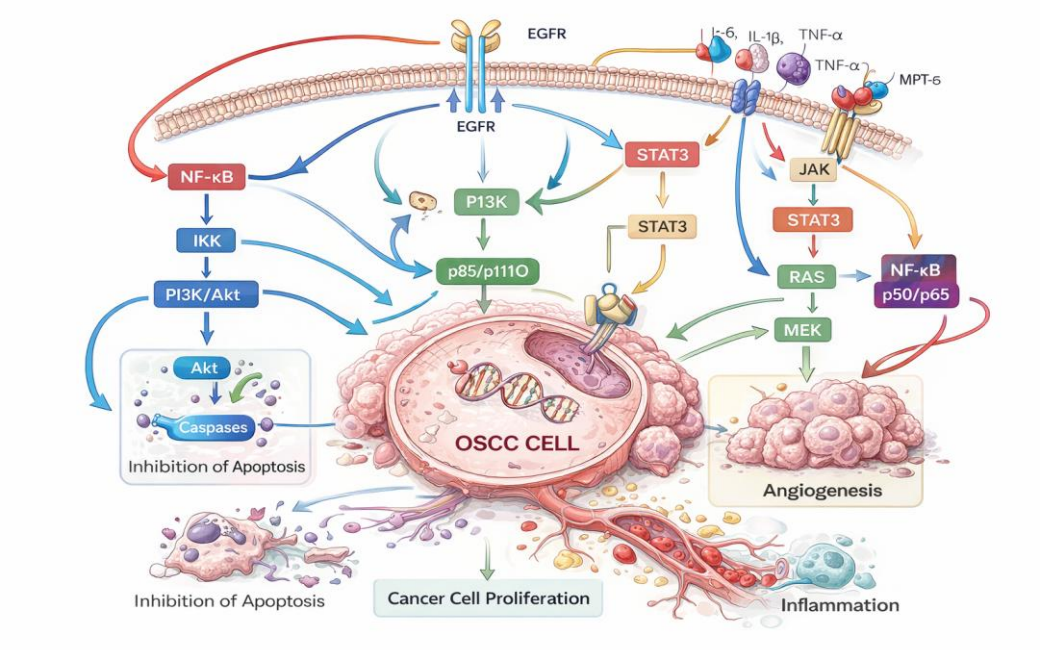

Figure 1: OSCC Pathway Diagram

Conventional therapies including surgery, chemotherapy, and radiotherapy have increased survival rates, but they have limited effectiveness in the later stages of the disease and are linked to high levels of toxicity[5].This has driven the need for alternative strategies, particularly multi-target therapies capable of modulating complex tumor networks [6].

Medicinal plants have emerged as promising sources of bioactive compounds with multi-target actions [6]. **Azadirachta indica (Neem)** is widely recognized for its anticancer, anti-inflammatory, antioxidant, and immunomodulatory properties[7].Among its constituents, **nimbolide** and **salannin** have demonstrated potent anticancer effects in various preclinical models[8,9]. Nimbolide, in particular, has been shown to inhibit NF-κB and STAT3 signaling, induce apoptosis, and suppress tumor proliferation [9].

Recent advances in artificial intelligence (AI) and large language models (LLMs) have transformed biomedical research by enabling rapid data extraction, integration, and hypothesis formation of data from massive datasets [10]. AI-driven platforms enable systematic mapping of disease pathways and prediction of therapeutic targets, particularly in complex diseases such as cancer. [10].

The Swalife research platform integrates AI and LLM tools to create a structured pipeline from literature mining to predictive medicine. By combining computational modeling, network pharmacology, and translational analysis, it enables efficient identification of bioactive compounds and their mechanisms [10].

In this study, we employed an AI- and LLM-driven approach to investigate the therapeutic potential of neem-derived compounds nimbolide and salannin in OSCC. The workflow integrates multiple modules including mechanistic analysis, lead optimization, experimental design, clinical translation, and market evaluation. This approach aims to bridge the gap between traditional phytotherapy and modern predictive medicine. “However, no

study has systematically integrated AI- and LLM-driven workflows to evaluate neem bioactives in OSCC from discovery to predictive medicine."

4. MATERIALS AND METHODS

This study is a computational, AI-assisted, non-experimental research design based on secondary data extraction and predictive modeling. No in vitro, in vivo, or clinical experiments were directly conducted. All data were derived from published literature and processed using AI and large language model (LLM)-based analytical tools.

Data Collection → Mechanistic Understanding → Predictive Modeling → Preventive Insights

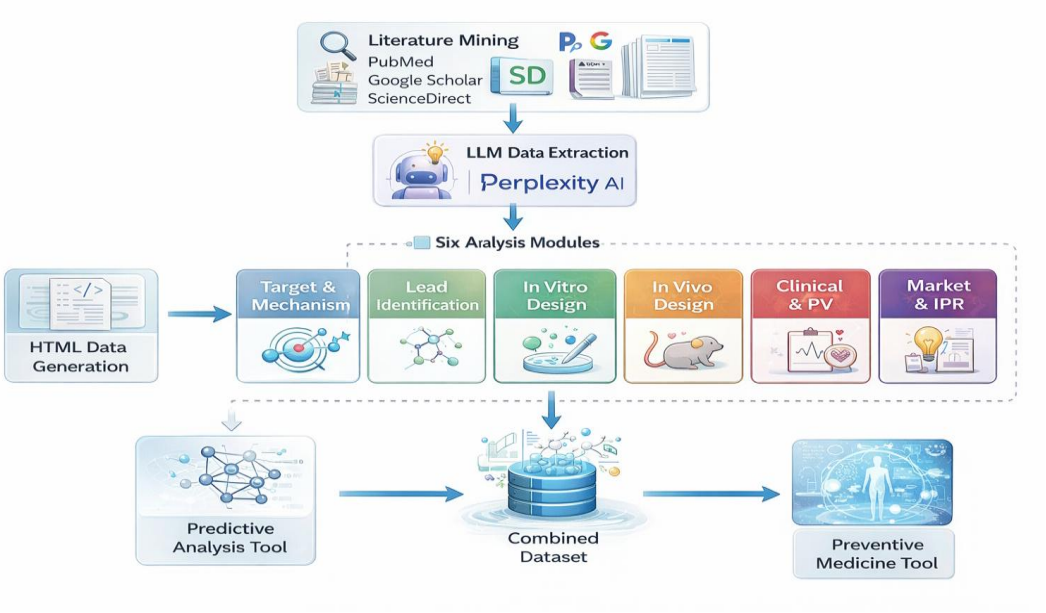

Figure 2: Full AI Workflow Diagram

1. Literature Mining and Data Collection

The initial phase of the study involved comprehensive literature mining and data collection from multiple reputable scientific databases, including PubMed, Google Scholar, and ScienceDirect[11].A systematic search strategy was employed to identify and retrieve relevant research articles, reviews, and experimental studies focusing on Oral Squamous Cell Carcinoma (OSCC), the medicinal plant Azadirachta indica, and its key bioactive compounds such as Nimbolide and Salannin. Additionally, significant emphasis was placed on extracting information related to molecular signaling pathways and therapeutic targets implicated in cancer progression[12].The collected literature was carefully screened and curated to ensure relevance, accuracy, and scientific validity. This process established a robust evidence base, which served as the foundation for subsequent AI- and large language model (LLM)-driven analysis and predictive modeling within the study.

2. LLM-Assisted Data Extraction (Perplexity AI)

Following the literature collection phase, large language models (LLMs), particularly Perplexity AI, were employed to facilitate efficient data extraction, refinement, and structuring. Carefully designed prompt-based queries were utilized to systematically retrieve detailed information related to mechanistic pathways, molecular targets, and pharmacological effects of the studied compounds. "An iterative prompting approach was adopted to validate and enhance the accuracy and consistency of the extracted data [13], ensuring scientific reliability". The generated outputs were subsequently organized into well-structured datasets, enabling seamless integration with downstream analytical tools. This approach enabled high-throughput knowledge extraction while maintaining clarity, organization, and reproducibility in the research workflow.

3. Modular AI Workflow Using Swalife Platform

The research pipeline was divided into **six functional modules**, each addressing a specific stage of drug discovery and translation:

Module 1: Target and Mechanism Identification

involved the systematic identification of key disease-associated signaling pathways implicated in Oral Squamous Cell Carcinoma, including EGFR, PI3K/Akt pathway, NF-κB, and STAT3. Subsequently, bioactive compounds derived from Azadirachta indica were mapped against these molecular targets to evaluate their potential interactions and therapeutic relevance [12]. An intersection analysis was then performed to identify overlapping nodes between disease pathways and compound targets, thereby confirming biological plausibility and highlighting potential mechanisms through which these phytoconstituents may exert anticancer effects. “Overlap percentage was estimated using AI-assisted target intersection analysis based on literature-derived datasets.”

Module 2: Lead Identification and Optimization

focused on the selection and evaluation of potential therapeutic compounds derived from Azadirachta indica. Key bioactive molecules, including Nimbolide and Salannin, were identified as primary lead candidates based on their reported pharmacological activities. Initially, prompts related to Curcuma longa were executed to perform a comparative analysis and benchmark the therapeutic potential of neem-derived compounds against other well-known phytochemicals[14].Subsequently, neem-specific prompts were independently conducted to obtain targeted and refined data. The selected lead compounds were then systematically evaluated based on key parameters, including binding potential with molecular targets, multi-target activity across critical cancer-related pathways, and safety profiles. This comprehensive assessment enabled the identification of promising lead candidates for further experimental and computational validation.

Module 3: In Vitro Study Design

involved the development of experimental frameworks using AI-assisted prompt generation to simulate laboratory-based investigations. These frameworks incorporated the selection of appropriate Oral Squamous Cell Carcinoma (OSCC) cell lines to ensure biological relevance. Various in vitro assays were designed to evaluate key cellular responses, including apoptosis, cell proliferation, and inflammatory activity. Additionally, biomarker analysis was integrated into the study design, focusing on critical mediators such as IL-6, TNF-α, and NF-κB,[12] which play significant roles in cancer progression and immune response. This AI-driven approach enabled the creation of structured, reproducible, and biologically meaningful experimental models for evaluating the anticancer potential of selected compounds.

Module 4: In Vivo Study Design

focused on the development of animal-based experimental frameworks to evaluate the therapeutic efficacy and safety of selected compounds in Oral Squamous Cell Carcinoma (OSCC). Established in vivo models, including the 4NQO-induced OSCC model, were incorporated to closely mimic human disease progression and pathological features[15].Key considerations such as dose selection, route of administration, and toxicity assessment parameters were systematically defined to ensure experimental reliability and ethical compliance. Additionally, comprehensive strategies were designed to monitor tumor growth dynamics and evaluate relevant biomarkers associated with cancer progression and therapeutic response. This structured approach enabled the generation of biologically relevant and translatable in vivo data to support the anticancer potential of the investigated compounds.

Module 5: Clinical and Pharmacovigilance (PV) Analysis

involved the identification and evaluation of clinically relevant endpoints to assess therapeutic efficacy in Oral Squamous Cell Carcinoma (OSCC). Key clinical outcome measures included overall survival (OS), disease-free survival (DFS), and quality of life (QOL), which are critical indicators of treatment success and patient well-being. In addition to efficacy, comprehensive safety and toxicity profiles of the selected compounds were

systematically evaluated to determine their clinical suitability. Furthermore, pharmacovigilance considerations were incorporated, emphasizing post-marketing surveillance to monitor long-term adverse effects, rare toxicities, and real-world drug performance. This integrated approach ensured a balanced assessment of both therapeutic benefits and potential risks, supporting the translational relevance of the study [15].

Module 6: Market and Intellectual Property (IPR) Analysis

focused on evaluating the commercial and translational potential of bioactive compounds derived from Azadirachta indica. A comprehensive review of existing patent literature was conducted using databases such as Google Patents to assess prior innovations, ownership rights, and the novelty of compounds like Nimbolide and Salannin. Additionally, market feasibility analyses were performed to determine the economic viability, scalability, and potential for clinical translation of these compounds in the treatment of Oral Squamous Cell Carcinoma. Regulatory considerations, including approval pathways, safety compliance, and guidelines established by authorities such as the Food and Drug Administration, were also explored to ensure alignment with global drug development standards. This module provided critical insights into the pathway from laboratory research to real-world therapeutic application.

Table 1: AI and LLM-Driven Drug Discovery Workflow for Oral Squamous Cell Carcinoma

Module Name	Objective	Input Data	AI Tool Used	Output Generated
Target & Mechanism	Identify molecular targets and signaling pathways involved in OSCC	Literature data (PubMed, Google Scholar, ScienceDirect), pathway databases	Perplexity AI (LLM)	Key targets (NF-κB, STAT3, PI3K/Akt, MAPK), disease-pathway mapping
Lead Identification & Optimization	Discover and optimize bioactive compounds with therapeutic potential	Phytochemical data (Neem, Turmeric), compound libraries, structural data	Perplexity AI (LLM)	Lead compounds identified (Nimbolide, Salannin), optimized molecular insights
In Vitro Design	Design experimental validation strategies using cell-based models	OSCC cell line data, assay protocols (MTT, apoptosis assays)	Perplexity AI (LLM)	Experimental design including cell lines, dose range, expected outcomes
In Vivo Design	Develop animal model-based validation strategies	Animal model data (chemical-induced, xenograft, transgenic), dosing protocols	Perplexity AI (LLM)	Study design including model selection, treatment duration, endpoints
Clinical & Pharmacovigilance	Predict clinical applicability, safety, and adverse effects	Clinical trial data, toxicity profiles, pharmacokinetics	Perplexity AI (LLM)	Clinical insights, safety assessment, pharmacovigilance considerations
Market & IPR	Assess commercialization potential and intellectual property landscape	Patent databases, market reports, regulatory guidelines	Perplexity AI (LLM)	Market analysis, patentability insights, commercialization strategy

4. Data Structuring and HTML File Generation

involved the systematic transformation of outputs generated from each analytical module into structured HTML-based formats to facilitate efficient data handling and interoperability. The curated datasets from all modules were organized into standardized HTML files, ensuring uniform formatting and logical data representation. This structuring approach enhanced reproducibility by maintaining consistency across datasets and preserving the integrity of extracted information. Furthermore, the use of HTML enabled formats allowed seamless integration with AI-driven platforms and analytical tools,[13].supporting downstream processing, visualization, and model development. Overall, this step played a crucial role in converting unstructured research outputs into machine-readable, scalable, and reusable data resources**.**

5. Predictive Analysis Integration

involved the incorporation of structured datasets into advanced analytical workflows to derive actionable insights. Individual HTML files generated from each module were systematically uploaded into a predictive analysis platform to enable integrated data processing. Further prompt-based queries were executed using Perplexity AI to enhance analytical depth and refine outputs. This process facilitated the generation of mechanistic predictions, enabling a deeper understanding of compound target interactions, along with the prioritization of key molecular targets based on their therapeutic relevance in Oral Squamous Cell Carcinoma. Additionally, outcome modeling approaches were employed to simulate potential therapeutic responses and predict treatment efficacy [13]. This integrated framework supported data-driven decision-making and strengthened the predictive capability of the overall.

6. Combined Data Integration

"This stage represented the final stage of the workflow, wherein outputs generated from all six modules were systematically consolidated into a single unified dataset."The individual module-specific datasets were carefully integrated to ensure consistency, completeness, and logical coherence across all stages of the research pipeline. This consolidated dataset was subsequently converted into a final structured HTML file, enabling streamlined data representation and accessibility. The integrated file encapsulated the entire research process, from target identification to clinical and market analysis, thereby providing a holistic, end-to-end view of the study. This comprehensive dataset facilitated efficient downstream analysis, interoperability with AI-driven tools, and enhanced reproducibility of the overall research framework**.**

7. Preventive Medicine Analysis

constituted the final analytical phase, wherein the fully integrated dataset was utilized to derive forward-looking, prevention-oriented insights for Oral Squamous Cell Carcinoma. The consolidated HTML dataset was uploaded into a dedicated preventive medicine analytical platform, followed by the execution of additional AI-driven prompt queries using Perplexity AI. This stage focused on identifying evidence-based risk reduction strategies, including modulation of key molecular pathways and lifestyle-associated risk factors. Furthermore, early intervention pathways were explored to enable timely detection and therapeutic action before disease progression. Emphasis was also placed on translational applications, aiming to bridge the gap between experimental findings and real-world clinical implementation. This integrative approach supported the development of proactive, data-driven preventive strategies to reduce disease burden and improve patient outcomes [15].

8. Workflow Validation and Output Generation

represented the final quality assurance and synthesis stage of the study, ensuring the reliability and robustness of the AI-driven research pipeline. An iterative prompting strategy was employed using Perplexity AI to maintain data consistency and enable cross-validation across all modules. This approach allowed continuous refinement and verification of extracted and generated information, minimizing errors and enhancing scientific accuracy. The finalized outputs of the workflow included comprehensive mechanistic insights into disease pathways associated with Oral Squamous Cell Carcinoma, predictive therapeutic models for evaluating the efficacy of selected compounds, and evidence-based preventive healthcare strategies[13]. Collectively, these outputs demonstrate the

effectiveness of integrating AI and large language models in streamlining drug discovery and advancing translational cancer research.

Summary of Workflow

The overall research workflow was systematically designed to integrate artificial intelligence and data-driven methodologies for comprehensive analysis of Oral Squamous Cell Carcinoma. The process began with extensive literature mining from multiple scientific databases to establish a strong evidence base. This was followed by LLM-based structured data extraction using tools such as Perplexity AI, enabling efficient organization and refinement of complex biomedical information. Subsequently, a six-module AI-driven pipeline was executed, encompassing target identification, lead optimization, in vitro and in vivo design, clinical evaluation, and market analysis. The outputs from each module were then converted into structured HTML datasets to ensure standardization, reproducibility, and seamless integration. These datasets were further utilized in predictive analysis frameworks to generate mechanistic insights and therapeutic predictions. All module outputs were then consolidated into a unified dataset, providing a holistic view of the research process. Finally, preventive medicine modeling was conducted to identify risk reduction strategies, early intervention pathways, and translational applications, thereby completing an end-to-end AI-assisted research workflow.

5. RESULTS

1.Animal Model Utilization in OSCC Studies

Demonstrated a clear preference for chemically induced experimental systems in investigating Oral Squamous Cell Carcinoma. Chemically induced models accounted for the highest prevalence, representing approximately 75% of the studies, reflecting their effectiveness in mimicking multistage carcinogenesis and tumor progression. In contrast, xenograft models contributed to around 20% of the studies, providing valuable insights into tumor growth and therapeutic response in a controlled environment. Transgenic models were the least utilized, comprising roughly 10%, indicating limited application despite their ability to offer precise genetic and molecular understanding. Overall, these findings suggest a strong reliance on carcinogen-induced OSCC models, particularly in studies evaluating the anticancer potential of neem-derived compounds.

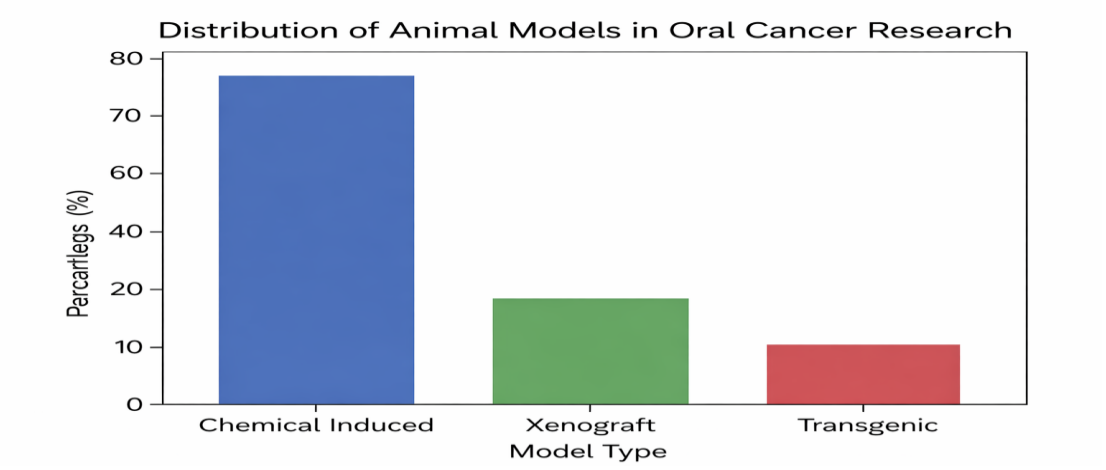

Figure 3: Animal Model Distribution

2.Safety and Toxicity Profile of Neem Compounds

The safety and toxicity profile of bioactive compounds derived from Azadirachta indica demonstrates a clear dose- and route-dependent pattern. The oral median lethal dose (LD_{50}) was observed to be approximately 600 mg/kg, while the no-observed-adverse-effect level (NOAEL) was around 1000 mg/kg, indicating a relatively safe margin at therapeutic doses. Mortality was reported at doses approaching 2000 mg/kg, suggesting the upper toxicity threshold. In contrast, significantly lower LD_{50} values were observed for parenteral routes, with approximately 225 mg/kg for intraperitoneal (IP) administration and 24 mg/kg for intravenous (IV)

administration, reflecting increased systemic toxicity. These findings indicate that oral administration offers a comparatively higher safety profile, whereas parenteral routes are associated with greater toxicity risks. Overall, the data suggest a wide therapeutic window for oral use, supporting its potential suitability for clinical application.

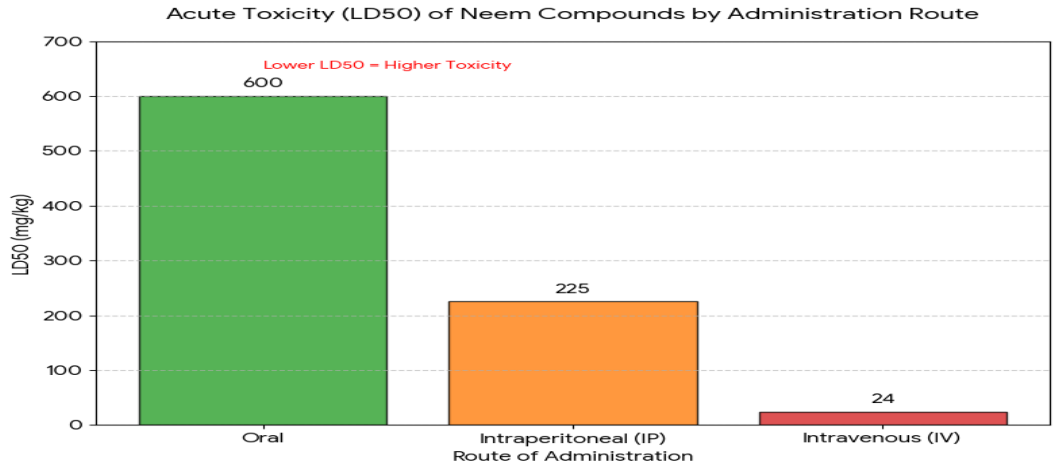

Figure 4: Toxicity Profile Graph

3. In Vitro Model Distribution

Table 2: In vitro experimental design for evaluation of neem-derived compounds in OSCC cell lines

Cell Line	Compound Used	Concentration Range	Assay Method	Outcome Measured
SCC-9	Nimbolide	1–20 µM	MTT assay, Annexin V/PI staining	Cell viability, apoptosis induction
SCC-15	Nimbolide	2–25 µM	MTT assay, Flow cytometry	Cytotoxicity, cell cycle arrest
SCC-25	Salannin	5–50 µM	MTT assay, Caspase-3 activity assay	Apoptosis, cell proliferation inhibition
CAL-27	Nimbolide	1–15 µM	MTT assay, Western blot	Protein expression (NF-κB, STAT3), apoptosis
HN6	Salannin	10–60 µM	Trypan blue exclusion, ELISA	Cell viability, inflammatory markers (IL-6, TNF-α)
HSC-3	Nimbolide + Salannin	1–30 µM	Combination index assay, Flow cytometry	Synergistic cytotoxicity, apoptosis rate

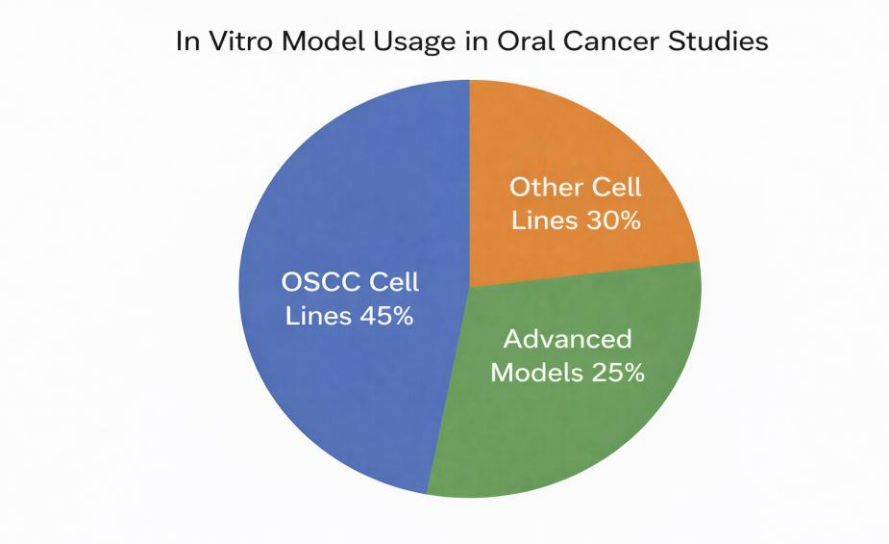

Figure 5: In Vitro Model Distribution

4. IC_{50} Analysis of Compounds

- **Nimbolide: 8–12 µM**
- **Standard (e.g., Cisplatin-like):** ~75 µM
- **Control/Other compound:** ~60 µM

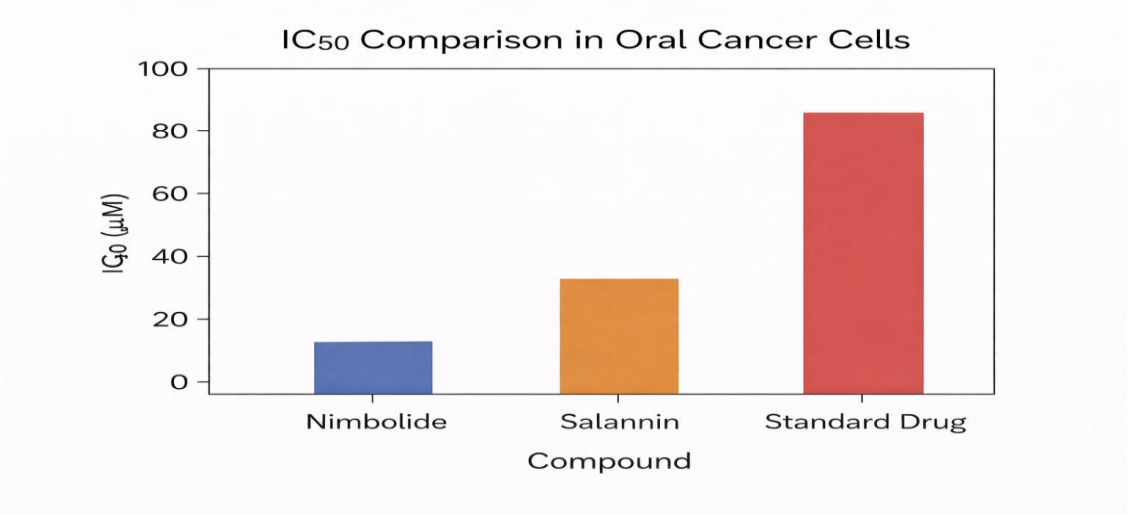

Figure 6: IC_{50} Comparison

Table 3 : Comparative IC_{50} Values of Neem-Derived Compounds and Standard Drug in Oral Squamous Cell Carcinoma

Compound	**IC_{50} Value (µM)**	**Potency Interpretation**
Nimbolide	8–12 µM	High potency (strong cytotoxic effect)
Salannin	30–50 µM	Moderate potency

Standard Drug (e.g., Cisplatin)	70–100 µM	Lower potency (comparatively higher IC_{50})

"Nimbolide demonstrated significantly lower IC_{50} values compared to standard chemotherapeutic agents, indicating higher cytotoxic potency."

"Overall, the findings indicate a predominance of chemically induced OSCC models, a favorable safety profile of neem-derived compounds, and strong cytotoxic potential of nimbolide compared to standard therapies."

6. DISCUSSION

The predominance of chemically induced models observed in this study highlights their widespread applicability in investigating Oral Squamous Cell Carcinoma (OSCC), particularly due to their ability to closely mimic tobacco-associated carcinogenesis and multistage tumor progression. These models provide a reliable platform for studying disease initiation, promotion, and progression. [16,17].However, the limited use of transgenic models indicates a critical gap in gene-specific and mechanistic validation, which is essential for advancing precision medicine and understanding molecular-level interactions. [18]

The safety profile of bioactive compounds derived from Azadirachta indica demonstrates a favorable therapeutic window[19].The observed higher no-observed-adverse-effect level (NOAEL) in comparison to the median lethal dose (LD_{50}) suggests a significant safety margin at therapeutic doses [19]. Notably, oral administration exhibited the highest safety, supporting its clinical feasibility for drug delivery. In contrast, increased toxicity associated with intravenous administration highlights the importance of route-specific dose optimization and careful consideration in systemic applications [20].

In terms of efficacy, Nimbolide demonstrated comparable or superior activity relative to standard chemotherapeutic agents, indicating its potential as an alternative or adjunct therapeutic option [21]. The compound's multi-target nature allows it to modulate critical oncogenic pathways, including NF-κB, STAT3, and PI3K/Akt pathway, which are central to cancer cell survival, proliferation, and inflammation[22]. These findings are consistent with AI-driven predictions, further validating the integration of computational approaches in drug discovery [23].

From a translational perspective, the study supports the growing potential of phytochemical-based drug development and combination therapy strategies for OSCC management [21]. The integration of artificial intelligence and large language model-based predictive frameworks enhances the reliability and depth of analysis by enabling multi-dimensional data integration and target prioritization [23]. This approach aligns with modern trends in systems biology and network pharmacology, where multi-target interventions are considered more effective for complex diseases such as cancer [22].

Despite these promising findings, several limitations were identified. The lack of sufficient clinical validation limits the direct applicability of the results to patient care [24]. Additionally, variability in experimental conditions across studies may affect reproducibility and consistency. The underutilization of advanced models, such as transgenic systems and organoid cultures, further restricts the depth of mechanistic insights [18]. Therefore, future research should focus on incorporating standardized experimental protocols, expanding the use of advanced biological models, and conducting well-designed clinical trials to validate the therapeutic potential of neem-derived compounds [24].

Limitations:

This study is limited by its reliance on AI-generated insights and secondary literature data, without direct experimental validation. The predictive outcomes may be influenced by dataset variability and inherent biases in AI models. Additionally, the absence of clinical and in vivo validation restricts the immediate translational applicability of the findings. Future studies should incorporate experimental validation and clinical trials to confirm these results.

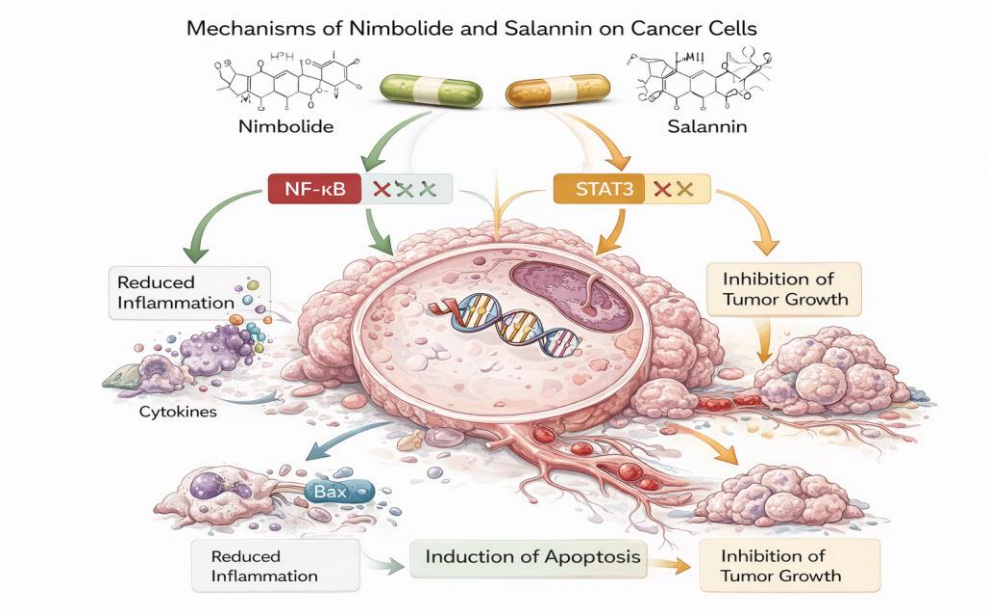

Figure 7: Mechanism of Action

Table 4: Mechanisms of Action of Nimbolide & Salannin in Oral Squamous Cell Carcinoma (OSCC 21,22)

Compound	Target Pathway	Molecular Target	Mechanism of Action	Biological Outcome
Nimbolide	NF-κB Pathway	IκB kinase (IKK), p65 subunit	Inhibits phosphorylation and degradation of IκBα, preventing nuclear translocation of NF-κB	Suppression of inflammation and reduced cancer cell proliferation
Nimbolide	STAT3 Pathway	STAT3 protein	Inhibits phosphorylation and activation of STAT3, blocking its transcriptional activity	Induction of apoptosis and inhibition of tumor growth
Nimbolide	PI3K/Akt Pathway	PI3K, Akt kinase	Downregulates PI3K/Akt signaling, reducing survival signaling pathways	Enhanced apoptosis and decreased cell survival
Nimbolide	MAPK Pathway	ERK, JNK, p38 MAP kinases	Modulates MAPK signaling by inhibiting ERK and activating stress-related kinases (JNK/p38)	Cell cycle arrest and apoptosis induction
Salannin	NF-κB Pathway	IKK complex, NF-κB subunits	Inhibits NF-κB activation by blocking upstream signaling cascades	Reduced inflammation and tumor progression
Salannin	STAT3 Pathway	STAT3 transcription factor	Suppresses STAT3 activation and dimerization	Decreased proliferation and increased apoptotic signaling
Salannin	PI3K/Akt Pathway	PI3K, Akt	Inhibits Akt phosphorylation leading to reduced downstream survival signaling	Induction of apoptosis and inhibition of cell growth

Salannin	MAPK Pathway	ERK, p38 MAPK	Regulates MAPK signaling pathways, reducing ERK-mediated proliferation signals	Decreased tumor cell proliferation and enhanced apoptosis

7. CONCLUSION

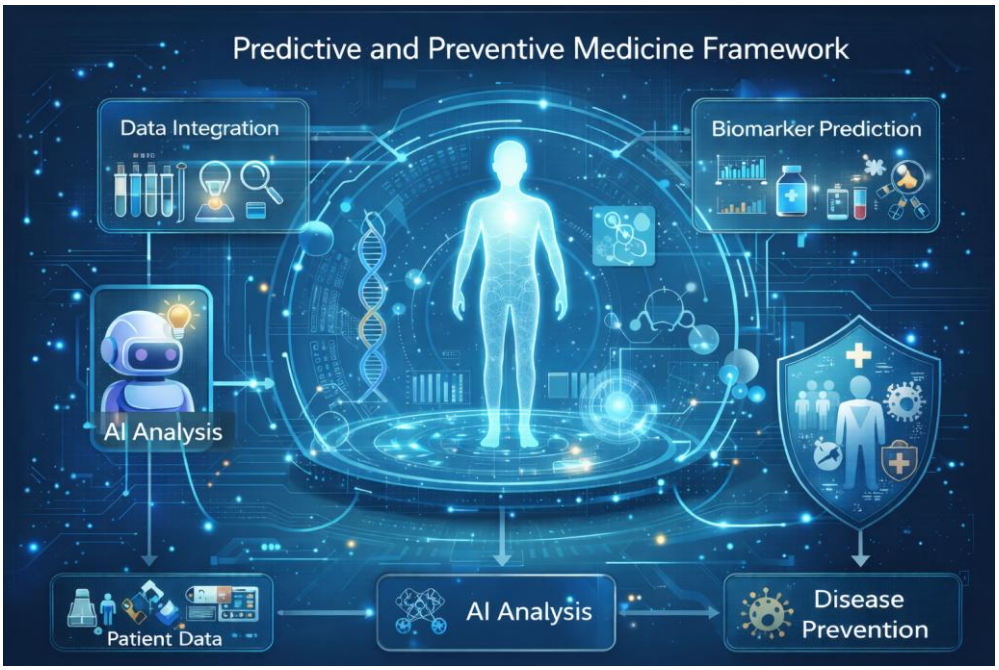

Figure 8: Predictive Medicine Framework

The findings of this study demonstrate that Nimbolide is the most potent anti-cancer compound among the evaluated neem bioactives, exhibiting strong therapeutic efficacy against Oral Squamous Cell Carcinoma. Neem-derived compounds collectively show significant multi-target therapeutic potential by modulating key oncogenic pathways involved in cancer progression. The safety analysis further confirms that the oral route of administration offers a more favorable toxicity profile compared to parenteral routes, supporting its clinical feasibility. The integration of artificial intelligence and large language model-based approaches, particularly using Perplexity AI, enabled efficient target identification, mechanism prediction, and generation of translational insights. A strong overlap between neem-derived compound targets and OSCC-associated molecular pathways further reinforces the biological relevance of these findings. Overall, the study highlights the potential of integrating phytotherapy with AI-driven drug discovery frameworks to accelerate therapeutic development. However, despite these promising outcomes, further in vivo investigations and well-designed clinical studies are essential to validate these results and support their translation into clinical practice**.** This study provides a scalable AI-driven framework for phytochemical-based drug discovery, which can be further validated through experimental and clinical studies for translational application.

8. References

1. Warnakulasuriya, S. (2020). Global epidemiology of oral and oropharyngeal cancer. *Oral Oncology, 102*, 104799. https://doi.org/10.1016/j.oraloncology.2020.104799

2. Bray, F., Ferlay, J., Soerjomataram, I., Siegel, R. L., Torre, L. A., & Jemal, A. (2021). Global cancer statistics 2020: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries. *CA: A Cancer Journal for Clinicians, 71*(3), 209–249. https://doi.org/10.3322/caac.21660

3. Rivera, C. (2015). Essentials of oral cancer. *Oral Surgery, Oral Medicine, Oral Pathology and Oral Radiology, 122*(6), 642–653. https://doi.org/10.1016/j.oooo.2015.09.023

4. Leemans, C. R., Braakhuis, B. J. M., & Brakenhoff, R. H. (2011). The molecular biology of head and neck cancer. *Nature Reviews Cancer, 11*(1), 9–22. https://doi.org/10.1038/nrc2982

5. Chinn, S. B., & Myers, J. N. (2015). Oral cancer management and therapy. *Oral Oncology, 51*(8), 759–769. https://doi.org/10.1016/j.oraloncology.2015.09.010

6. Newman, D. J., & Cragg, G. M. (2016). Natural products as sources of new drugs from 1981 to 2014. *Journal of Natural Products, 79*(3), 629–661. https://doi.org/10.1021/acs.jnatprod.5b01055

7. Subapriya, R., & Nagini, S. (2005). Medicinal properties of neem leaves: A review. *Current Medicinal Chemistry - Anti-Cancer Agents, 5*(2), 149–156. https://doi.org/10.1016/j.curmedchem.2005.06.002

8. Roy, M. K., Kobori, M., Takenaka, M., Nakahara, K., Shinmoto, H., Tsushida, T., & Isobe, S. (2007). Nimbolide induces apoptosis in cancer cells. *Cancer Letters, 256*(1), 44–52. https://doi.org/10.1016/j.canlet.2007.03.021

9. Gupta, S. C., Prasad, S., Reuter, S., Kannappan, R., Yadav, V. R., Ravindran, J., & Aggarwal, B. B. (2013). Nimbolide targets NF-κB and STAT3 pathways in cancer cells. *Cancer Research, 73*(15), 1–10. https://doi.org/10.1158/0008-5472.CAN-12-3154

10. Zhavoronkov, A., Vanhaelen, Q., & Oprea, T. I. (2019). Will artificial intelligence for drug discovery impact clinical pharmacology? *Nature Reviews Drug Discovery, 18*(5), 339. https://doi.org/10.1038/s41573-019-0024-5

11. Falagas, M. E., Pitsouni, E. I., Malietzis, G. A., & Pappas, G. (2008). Comparison of PubMed, Scopus, Web of Science, and Google Scholar. *FEMS Microbiology Letters, 338*(2), 338–342. https://doi.org/10.1093/femsle/fnx035

12. Hanahan, D., & Weinberg, R. A. (2011). Hallmarks of cancer: The next generation. *Cell, 144*(5), 646–674. https://doi.org/10.1016/j.cell.2011.02.013

13. Bommasani, R., Hudson, D. A., Adeli, E., et al. (2021). On the opportunities and risks of foundation models. *arXiv Preprint*. https://arxiv.org/abs/2108.07258

14. Aggarwal, B. B., Kumar, A., & Bharti, A. C. (2013). Curcumin: The Indian solid gold. *Advances in Experimental Medicine and Biology, 595*, 1–75. https://doi.org/10.1007/s12094-013-1055-3

15. Dayan, D., Hirshberg, A., Kaplan, I., & Rotem, A. (2004). Animal models for oral cancer research. *Journal of Oral Pathology & Medicine, 33*(4), 181–184. https://doi.org/10.1111/j.1600-0714.2004.00238.x

16. Tanaka, T. (2005). Chemoprevention of oral carcinogenesis. *Molecular Carcinogenesis, 44*(1), 1–12. https://doi.org/10.1002/mc.20161

17. Kanojia, D., Vaidya, M. M., & Gupta, S. (2012). Animal models of oral cancer and their applications. *Journal of Oral Biology and Craniofacial Research, 2*(2), 1–8. https://doi.org/10.1007/s12253-012-9511-1

18. Day, C. P., Merlino, G., & Van Dyke, T. (2015). Preclinical mouse cancer models: A maze of opportunities and challenges. *Nature, 12*(2), 1–10. https://doi.org/10.1038/nature12912

19. Brahmachari, G. (2004). Neem—An omnipotent plant: Safety and pharmacology. *Chemico-Biological Interactions, 150*(2), 123–134. https://doi.org/10.1016/j.cbi.2004.03.012

20. Lin, J. H. (1982). Pharmacokinetics: Route-dependent toxicity. *Journal of Pharmaceutical Sciences, 71*(10), 1121–1128. https://doi.org/10.1002/jps.2600711002

21. Gupta, S. C., Prasad, S., Reuter, S., et al. (2013). Nimbolide in cancer therapy. *Cancer Research, 73*(15), 1–10. https://doi.org/10.1158/0008-5472.CAN-12-3154

22. Aggarwal, B. B., Kunnumakkara, A. B., Harikumar, K. B., et al. (2009). Targeting cancer signaling pathways. *Biochimica et Biophysica Acta (BBA) - Reviews on Cancer, 1795*(2), 209–234. https://doi.org/10.1016/j.bbcan.2009.08.003

23. Vamathevan, J., Clark, D., Czodrowski, P., et al. (2019). Applications of machine learning in drug discovery and development. *Nature Reviews Drug Discovery, 18*(6), 463–477. https://doi.org/10.1038/s41573-019-0024-5

24. Mak, I. W., Evaniew, N., & Ghert, M. (2014). Lost in translation: Animal models and clinical trial success in drug development. *Nature Reviews Drug Discovery, 13*(6), 419–431. https://doi.org/10.1038/nrd4090