

Review Article

Exploring Advances in Early Screening of Nasopharyngeal Carcinoma Based on Regional Medical Big Data Platforms

Guowen Zhan^{*} 

Department of Otolaryngology-Head & Neck Surgery, Ningbo Hospital of Integrated Traditional Chinese and Western Medicine, Ningbo, China

Abstract

Nasopharyngeal carcinoma (NPC) has an insidious onset and lacks obvious early symptoms, often leading to diagnosis at middle or late stages and significantly worsening patient prognosis. Regional medical big data platforms integrate clinical, genomic, and imaging data sources. These platforms provide advanced technical support and novel research perspectives, such as the identification of new biomarkers and predictive modeling approaches, to facilitate early detection of NPC. This review summarizes recent advances in NPC early detection enabled by these platforms. It highlights the application and benefits of multimodal screening strategies, including Epstein-Barr virus (EBV)-related biomarker assays, DNA methylation analysis, imaging screening techniques, and machine learning algorithms. We synthesize the latest research findings and analyze the potential of multi-omics data integration and artificial intelligence technologies to improve screening accuracy. Moreover, we discuss how these approaches can reduce false-positive findings in NPC early detection. We also explore future directions for achieving precise and personalized NPC early detection through big data platforms. This review aims to provide both theoretical foundations and practical guidance for building regional medical big data platforms, thereby facilitating the optimization and wider implementation of NPC early detection systems by enhancing data integration and analytical capabilities.

Keywords

Nasopharyngeal Carcinoma, Early Screening, Regional Health Medical Big Data, Epstein-Barr Virus, DNA Methylation, Imaging, Machine Learning

1. Introduction

NPC is a distinct malignancy characterized by unique epidemiological, histopathological, and clinical features, which set it apart from other head and neck cancers [1]. The occurrence of this phenomenon shows significant variation across different regions and ethnic groups. Endemic hotspots of this condition are mainly found in East and Southeast Asia, but they also appear in parts of North Africa, including Morocco and Ethiopia [2]. Epidemiological data from Tikur Anbessa

Specialized Hospital in Ethiopia reveal a predominance of NPC among younger populations, with a median age at diagnosis of 38 years and a male-to-female ratio of approximately 2.5:1. The undifferentiated subtype, which accounts for over 85% of cases, dominates the histopathological profile. Notably, a significant proportion of patients (86%) present at advanced stages (III and IV), severely limiting therapeutic options and negatively impacting survival outcomes [3]. Similarly, studies from

^{*}Correspondence: Guowen Zhan (zgw15258258250@163.com)

Received: 8 January 2026; **Accepted:** 20 January 2026; **Published:** 31 January 2026



Copyright: © The Author(s), 2026. Published by Science Publishing Group. This is an **Open Access** article, distributed under the terms of the Creative Commons Attribution 4.0 License (<http://creativecommons.org/licenses/by/4.0/>), which permits unrestricted use, distribution and reproduction in any medium, provided the original work is properly cited.

Northern Morocco confirm the endemic nature of NPC in that region, with a median age of 50 years and a high prevalence of undifferentiated carcinoma (96.12%). The majority of patients present with advanced disease stages accompanied by cervical lymph node involvement. This underscores the challenge of late diagnosis despite the high disease burden [4]. Consequently, this late-stage presentation is a common theme across endemic areas and highlights the critical need for improved early detection strategies to enhance prognosis and survival.

The difficulty in early diagnosis of NPC is multifactorial. The anatomical location of the tumor within the nasopharynx—a region not easily accessible to routine clinical examination—combined with non-specific early symptoms, leads to delayed diagnosis. Current standard diagnostic approaches rely heavily on invasive procedures such as nasopharyngoscopy with biopsy, which are definitive but resource-intensive and not suitable for large-scale screening [5]. Imaging modalities, including Magnetic resonance imaging (MRI), computed tomography (CT), and endoscopy, are essential for staging and treatment planning but have limitations in sensitivity and specificity for early lesions. Recent advances in radio-pharmaceuticals and molecular imaging techniques, including single-photon emission computed tomography (SPECT) and positron emission tomography (PET), have shown promise in enhancing diagnostic accuracy and delineating tumor extent. However, their widespread adoption is constrained by cost and availability [6]. Moreover, traditional serological markers and EBV-related assays, such as EBV DNA quantification and antibody detection, have been incorporated into screening programs; however, when used alone, these tests suffer from limited positive predictive values and variable sensitivity. For instance, a prospective cohort study in Hunan Province, China, demonstrated that the BNL2b antibody had superior sensitivity (83.78%) and specificity (95.65%) compared to testing with EBV DNA and dual antibody assays, particularly for early-stage NPC detection. Combining BNL2b with dual antibody assays further enhanced specificity to 99.64%, making it a valuable biomarker for identifying high-risk individuals and guiding early intervention strategies [7]. Despite these advances, challenges remain in developing noninvasive, cost-effective, and highly accurate screening tools that can be deployed at the population level, especially in resource-limited settings.

In this context, the emergence of regional health and medical big data platforms offers unprecedented opportunities to revolutionize NPC early detection. These platforms integrate diverse data streams, including clinical records, imaging datasets, molecular biology findings, and epidemiological information, creating a multidimensional repository that supports comprehensive analysis and risk stratification [8]. The integration of multimodal data facilitates the identification of complex patterns and novel biomarkers that may be elusive

when single data types are analyzed in isolation. For example, in oncology and other disease domains, multimodal artificial intelligence (AI) models, combining proteomics data, imaging, and clinical data, have demonstrated significantly improved diagnostic performance over single-modality approaches. A notable example is a lung cancer screening platform that integrated multiplexed immunoassay data with computed tomography imaging and clinical parameters. This platform achieved an area under the curve (AUC) of 0.936, outperforming conventional methods [9]. Such integrative frameworks, when adapted to NPC, could harness the power of machine learning algorithms to enhance early detection accuracy, reduce false positives, and personalize screening recommendations.

Furthermore, recent developments in AI and machine learning have shown potential in transforming disease screening and management across various medical fields. AI-driven models have been successfully applied in chronic ocular diseases, stroke, epilepsy, and cancer. Furthermore, these models enable precision diagnostics, prognostic assessments, and therapeutic optimizations by leveraging multimodal data inputs, including imaging, genetic, and clinical data [10]. For instance, AI applications in gastric cancer have evolved to integrate medical imaging, digital pathology, and multiomics data, addressing challenges related to tumor heterogeneity and fragmented clinical workflows. These advances have paved the way toward integrated AI-driven diagnostic platforms capable of real-time clinical decision support [11]. Additionally, the fusion of artificial intelligence (AI) with molecular and imaging biomarkers for nasopharyngeal carcinoma (NPC) screening can overcome the limitations of traditional methods. This approach provides noninvasive, scalable, and cost-efficient solutions tailored to specific regional epidemiological characteristics, such as incidence rates and risk factors.

Given the multifaceted challenges posed by NPC—its regional endemicity, early diagnostic difficulties, and the shortcomings of existing screening modalities—there is an urgent need to develop novel, noninvasive, and efficient early detection strategies [5]. Regional health and medical big data platforms represent a promising infrastructure to support such innovations by enabling the integration and analysis of heterogeneous data sources [12]. This review focuses on the exploration of new NPC early screening strategies based on these platforms, emphasizing the application of multimodal approaches that combine EBV-related biomarkers, DNA methylation markers, advanced imaging techniques, and AI-driven algorithms. We aim to synthesize current progress and discuss the potential of these integrated strategies to improve screening accuracy, facilitate early diagnosis, and ultimately contribute to precision medicine approaches tailored to high-risk populations in endemic regions.

2. Body

2.1. Application of EBV-related Biomarkers in Early Screening of NPC

2.1.1. EBV DNA and Antibody Detection Technology Advances

EBV DNA detection in plasma has emerged as a cornerstone for NPC screening due to its high sensitivity and specificity. Meta-analytical data encompassing over 8,000 NPC patients and 15,000 controls demonstrate that plasma EBV DNA detection achieves a sensitivity of approximately 76% and specificity of 96%, outperforming traditional EBV antibody markers such as EA-IgA, VCA-IgA, and EBNA1-IgA in diagnostic accuracy [13]. Despite this, the positive predictive value (PPV) of plasma EBV DNA remains suboptimal, limiting its standalone utility for population-wide screening. This limitation is partly due to the presence of EBV DNA in asymptomatic individuals and variability in viral load dynamics. To address these challenges, recent advances have focused on novel antibody biomarkers that complement EBV DNA testing. Notably, the anti-BNLF2b antibody (termed P85-Ab) has been identified through a peptide library screening approach as a highly promising serologic biomarker. In a large prospective cohort involving nearly 25,000 participants, P85-Ab exhibited superior sensitivity (94.4%) and specificity (99.6%) compared to the conventional dual-antibody panel of EBNA1-IgA and VCA-IgA [14]. The combination of P85-Ab with the traditional antibody panel markedly improved the PPV from 4.3% to 44.6%, significantly reducing false positives and enhancing screening efficiency. This synergistic effect underscores the potential of integrating novel antibody markers with existing assays to refine NPC early detection.

Further population-based studies have validated the high diagnostic performance of EBV DNA and antibody assays. For instance, direct comparisons of plasma EBV DNA quantification using real-time PCR combined with next-generation sequencing (NGS) and EBV antibody scoring revealed that the EBV DNA algorithm achieved higher sensitivity (93.2%) and specificity (98.1%) than the antibody score alone, with comparable sensitivity for early-stage NPC detection [15]. These findings support the use of plasma EBV DNA as a primary screening tool, supplemented by antibody testing to improve risk stratification. Additionally, emerging evidence suggests that plasma EBV DNA fragmentomics, analyzing the size and fragmentation patterns of circulating EBV DNA, may further enhance specificity and prognostic value, although this requires validation in large cohorts [16].

The clinical utility of antibody biomarkers extends beyond traditional IgA responses to VCA and EBNA1. Novel antibody targets such as BGLF2 and BARF1 have been identified, with studies showing that IgA and IgG antibodies against these antigens are elevated years before NPC diagnosis, indicating their potential as early predictive markers [17, 18].

BARF1-targeted DNA immunotherapy also highlights the immunogenicity of this antigen and its relevance in NPC pathogenesis and treatment [18]. Furthermore, prospective cohort studies in endemic regions demonstrate that combining EBV DNA and antibody testing can identify early-stage NPC cases with high accuracy, facilitating timely intervention and improved prognosis [7].

Despite these advances, challenges remain in optimizing screening algorithms to balance sensitivity, specificity, and cost-effectiveness. For example, nasopharynx swab EBV DNA testing as a reflex to seropositive individuals can reduce unnecessary referrals while maintaining high sensitivity [19]. Moreover, demographic factors such as family history and lifestyle behaviors influence biomarker positivity and NPC risk, underscoring the need for personalized screening strategies [7]. The integration of these biomarkers into regional health data platforms and large-scale prospective studies will be critical to establish standardized protocols for NPC early detection.

In summary, plasma EBV DNA detection remains the core modality for NPC screening due to its high sensitivity, but its PPV is limited when used alone. The discovery and validation of novel antibody biomarkers, particularly anti-BNLF2b (P85-Ab), have significantly improved the specificity and PPV of serological screening. Combining P85-Ab with traditional EBV antibodies and EBV DNA testing in large prospective cohorts has demonstrated enhanced diagnostic accuracy and reduced false positives. These advances, supported by robust clinical evidence, pave the way for more effective, population-based NPC screening programs, especially in endemic regions, ultimately improving early detection and patient outcomes. Continued research into additional biomarkers and integration with emerging technologies such as fragmentomics and immunotherapy targets will further refine NPC screening strategies.

2.1.2. EBV-related Gene Expression and Viral Latency Mechanisms

The EBV latent gene products, particularly latent membrane protein 1 (LMP1) and EBV-encoded small RNAs (EBERs), play crucial roles in the pathogenesis of NPC. LMP1 and EBERs are closely associated with NPC development, as their expression modulates various cellular signaling pathways that promote oncogenesis. For instance, LMP1 up-regulates the expression of insulin-like growth factor 1 receptor (IGF-1R), and a significant correlation exists between IGF-1R expression and EBV markers, which impacts patient prognosis [20]. This suggests that EBV latent gene expression not only influences tumor cell behavior but also affects clinical outcomes. Moreover, EBV latent infection is characterized by a tightly regulated balance between latency and lytic reactivation. The latent-lytic switch is controlled epigenetically and transcriptionally by host and viral factors. The histone demethylase machinery, including enzymes such as lysine-specific demethylase 1 (LSD1), zinc finger protein 217 (ZNF217),

and the CoREST corepressor complex, modulates chromatin states at viral promoters, thereby regulating viral gene expression and maintaining latency [21]. These epigenetic regulators contribute to the suppression of lytic gene expression and the maintenance of a tumor microenvironment conducive to viral persistence and oncogenesis. The CoREST complex, in particular, acts as a repressor of lytic genes, ensuring the virus remains in a latent state within infected epithelial and B cells. This viral latency is further stabilized by interactions with host nuclear lamina components and chromatin modifiers, which organize the viral episome into a repressive chromatin structure [22]. The interplay between viral gene expression and host epigenetic regulation provides novel molecular targets for therapeutic intervention and early screening strategies. Targeting components of the LSD1/ZNF217/CoREST complex or modulating LMP1 and EBERs expression could disrupt the viral latency program, potentially reactivating the virus and sensitizing tumor cells to antiviral therapies or immune clearance. Additionally, the expression patterns of these viral genes and their impact on host signaling pathways, such as IGF-1R, offer biomarkers for early detection and prognosis in NPC patients. Thus, understanding the molecular mechanisms underlying EBV latency and gene expression not only illuminates the pathogenesis of NPC but also opens avenues for developing targeted therapies and improving early screening methods based on EBV-related molecular signatures.

2.2. DNA Methylation Biomarkers and Their Multi-omics Detection Technologies

2.2.1. Screening Value of Methylated Genes and Representative Biomarkers

The aberrant methylation of tumor suppressor genes has emerged as a pivotal epigenetic alteration in NPC, offering promising avenues for early diagnosis and screening. Among these, RASSF1A, SEPTIN9, DAPK, and p16INK4 α have been extensively studied and identified as representative methylation biomarkers with significantly elevated methylation frequencies in NPC tissues compared to non-cancerous controls. RASSF1A, a well-known tumor suppressor gene, shows hypermethylation in about 60% of NPC cases versus less than 3% in controls, with an odds ratio exceeding 30, highlighting its strong link to NPC tumorigenesis. Additionally, RASSF1A has potential utility as both a diagnostic and prognostic marker [23]. This hypermethylation leads to gene silencing, contributing to uncontrolled cell proliferation. Similarly, SEPTIN9 methylation has been detected in over 90% of NPC tissue samples and is also identifiable in nasopharyngeal swabs, demonstrating a minimally invasive method for early NPC detection. The methylation status of SEPTIN9 correlates with decreased mRNA expression. Detection of this methylation status in nasal swabs achieves high diagnostic accuracy, with an AUC of 0.882, indicating its robustness as a screening biomarker. [24].

In addition to these, multiple tumor suppressor genes located on chromosome 9, including DAPK, p16INK4 α , p15INK4 α , and p14ARF, show frequent promoter hypermethylation in NPC. A case-control study revealed hypermethylation rates of 75.7%, 78.6%, 52.9%, and 87.1% respectively, with a methylation index (MI) indicating that all NPC samples had methylation in at least one of these genes. This composite methylation index demonstrated 100% sensitivity and 50% specificity for NPC detection, highlighting the potential of combining multiple methylation markers to enhance diagnostic accuracy [25]. The methylation of these genes likely contributes to NPC pathogenesis by disrupting cell cycle regulation and apoptotic pathways, underscoring their biological relevance as early biomarkers.

Furthermore, non-invasive detection methods utilizing nasopharyngeal swabs have been developed to assess methylation of RASSF1A, SEPTIN9, and other relevant genes. Among these, RASSF1A methylation demonstrates the highest classification accuracy (AUC = 0.956) for distinguishing NPC patients from healthy controls. The detection rates in swabs are significantly higher than in plasma samples, emphasizing the advantage of local sampling for early NPC diagnosis [26]. This approach aligns with the clinical need for sensitive, specific, and minimally invasive screening tools.

Beyond individual genes, recent studies employing genome-wide methylation analyses have identified additional candidate methylation biomarkers such as RERG and ZNF671 in circulating cell-free DNA (cfDNA), which also demonstrate significant diagnostic potential. Combining methylation rates of RERG and ZNF671 in cfDNA improves diagnostic accuracy, suggesting that multi-gene panels may further refine early NPC detection strategies [27].

Advances in detection technologies, including novel quantitative PCR assays and CRISPR/Cas12a-based methods, further support the clinical utility of these methylation biomarkers. These technologies offer rapid, sensitive, and specific methylation detection suitable for clinical applications. For instance, the development of methylation-sensitive restriction endonuclease-assisted assays enables detection of methylation levels as low as 0.1%, facilitating discrimination between NPC and normal cells [28].

In summary, the elevated methylation frequencies of tumor suppressor genes such as RASSF1A, SEPTIN9, DAPK, and p16INK4 α represent a robust molecular signature of NPC, with strong evidence supporting their roles as early diagnostic and screening biomarkers. The integration of these methylation markers into minimally invasive sampling platforms and advanced detection methods holds significant promise for improving early NPC detection, ultimately enhancing patient prognosis and guiding timely therapeutic interventions.

2.2.2. Multi-marker Combined Detection Strategy

The integration of multiple biomarkers, particularly combining DNA methylation profiles with EBV antibody detection, represents a promising multi-marker strategy for early

screening and risk assessment of NPC. DNA methylation changes have been consistently associated with NPC pathogenesis. In particular, hypermethylation of tumor suppressor genes such as RASSF1A, SEPTIN9, and other tumor suppressor genes located on chromosome 9 can serve as sensitive early diagnostic markers. For instance, promoter hypermethylation of genes like DAPK, p16INK4a, p15INK4a, and p14ARF showed high prevalence in NPC tissues and demonstrated diagnostic potential with high sensitivity when combined into a methylation index [25]. Similarly, methylation of RASSF1A has been validated as a robust biomarker for NPC diagnosis and prognosis through meta-analyses and clinical investigations [23, 26]. Notably, methylation detection in non-invasive samples such as nasopharyngeal swabs or plasma cell-free DNA has shown promising sensitivity and specificity, offering a practical approach for population screening [26, 29].

On the other hand, EBV serological markers, particularly IgA antibodies against EBV nuclear antigen 1 (EBNA1-IgA), viral capsid antigen (VCA-IgA), and novel antibodies such as anti-BNLF2b, have been extensively studied and widely applied in NPC screening programs. These antibodies exhibit high sensitivity and specificity, with anti-BNLF2b antibody demonstrating superior diagnostic performance compared to traditional two-antibody methods [7, 14, 30]. Moreover, the combination of EBV antibody profiles with plasma EBV DNA quantification enhances early detection accuracy and risk stratification [16, 31].

Recent studies have emphasized that combining methylation markers with EBV antibody detection can significantly improve the positive predictive value (PPV) of NPC screening, addressing the limitations of individual biomarkers. For example, integrating cfDNA methylation analysis with EBV serological testing has been proposed as a key approach to overcome the low PPV of current screening strategies. This integration can reduce false positives, thereby optimizing referral rates for invasive diagnostic procedures [32]. Additionally, multiplex assays that simultaneously evaluate EBV antibody profiles and methylation status of host genes have demonstrated enhanced sensitivity and specificity in NPC detection [33, 34].

The construction of multidimensional risk assessment models that incorporate both methylation data and EBV antibody levels allows for a comprehensive evaluation of NPC risk. Such models leverage the complementary biological information: methylation changes reflect epigenetic alterations in host tumor suppressor genes, while EBV antibodies indicate viral infection status and immune response. For instance, the combination of serum EphB2 protein levels with EBV seropositivity improved diagnostic accuracy, suggesting that integrating multiple biomarkers can capture different facets of NPC pathogenesis [35]. Similarly, machine learning approaches integrating transcriptomic data and immune profiling have identified key biomarkers and immune-related pathways, further supporting the feasibility of multi-marker models for early NPC detection [36].

The diversity of NPC subtypes affects biomarker screening, with non-keratinizing NPC needing different strategies than keratinizing NPC. Recent studies indicate that using serum biomarkers, imaging, and genetic profiling can improve screening accuracy, with ctDNA and miRNA panels helping identify high-risk patients and customize surveillance. This approach enhances early detection and monitors treatment response and disease progression in various NPC subtypes [37]. The relationship between biomarker levels and disease progression in NPC highlights the dynamic nature of tumor biology. Research indicates that as NPC advances from early to late stages, specific biomarkers, such as serum levels of EBV DNA and various cytokines, exhibit significant changes that correlate with disease severity [38]. For instance, increased levels of inflammatory markers have been observed in patients with advanced disease, suggesting their potential role in tumor progression and immune response modulation [39]. Additionally, longitudinal studies have demonstrated that the monitoring of these biomarkers can provide critical insights into disease trajectory, enabling clinicians to predict outcomes and tailor interventions accordingly [40]. Understanding the dynamic changes in biomarker levels throughout the disease continuum is essential for enhancing diagnostic precision and optimizing therapeutic strategies in NPC management.

In summary, multi-marker combined detection strategies that integrate DNA methylation signatures with EBV antibody assays provide a multidimensional risk assessment framework for NPC. These approaches enhance diagnostic sensitivity and specificity, improve PPV, and facilitate early detection, which is critical for improving patient outcomes. The current landscape reveals a gap in the comprehensive and balanced validation of these biomarkers across various NPC subtypes and disease stages, especially in the critical early phases of the disease. Future research should focus on validating these combined models in large-scale prospective cohorts, optimizing biomarker panels for cost-effectiveness, and developing standardized protocols for sample collection and analysis to enable widespread clinical application.

2.3. Imaging Techniques and AI-assisted Diagnosis

2.3.1. MRI in Early NPC Screening

MRI has emerged as a pivotal tool in the early screening of NPC, particularly due to its superior soft tissue resolution and non-invasive nature. A key advancement is the use of rapid, non-contrast-enhanced MRI sequences as screening tools. Such protocols enable the detection of early NPC lesions that may be occult to traditional endoscopic examination and biopsy, thereby identifying patients with early-stage disease who would otherwise remain undiagnosed. For instance, a prospective study involving 354 patients demonstrated that a short, contrast-free MRI screening complemented endoscopy by detecting additional NPC cases that were endoscopically

negative or biopsy-negative, with MRI achieving a sensitivity of 88.9% and specificity of 91.1% for NPC detection [41]. This indicates that rapid MRI can uncover early mucosal or submucosal changes before they become visible or accessible to biopsy, thus addressing a critical gap in early NPC screening.

Further enhancing MRI's diagnostic utility is the development of improved MRI grading systems that refine the differentiation between early NPC and benign nasopharyngeal hyperplasia. A modified MRI grading system that incorporates refined tumor criteria and nodal assessment has been shown to significantly improve diagnostic performance. In a large cohort study comparing current, modified, and non-contrast plain scan MRI grading systems, the modified system achieved a sensitivity of 97.9% and a specificity of 89.6%, outperforming both the current contrast-enhanced and plain scan systems [42]. Notably, no additional cancers were detected by contrast administration in cases graded low-risk by the plain scan, suggesting that rapid non-contrast MRI may suffice for initial screening, reserving contrast-enhanced imaging for higher-risk patients.

AI and machine learning approaches have further augmented MRI's role in early NPC detection. A convolutional neural network (CNN) applied to T2-weighted fat-suppressed MRI sequences demonstrated high accuracy (91.5%) and an area under the receiver operating characteristic curve of 0.96 in discriminating early-stage T1 NPC from benign hyperplasia [43]. This automated method offers a promising avenue for large-scale screening programs by reducing observer variability and expediting image interpretation. Similarly, radiomics models extracting quantitative imaging features have been developed to distinguish early NPC from benign lesions, with an AUC of 0.85 in training cohorts and stable feature selection techniques enhancing model reliability [44].

Advanced MRI sequences such as diffusion-weighted imaging (DWI) have also contributed to early NPC detection. Reduced field-of-view DWI techniques like FOCUS provide superior image quality and better delineation of early-stage tumors compared to conventional single-shot echo-planar imaging, with significantly lower apparent diffusion coefficient values observed in early T and N stages versus advanced disease [45]. These functional imaging parameters reflect tumor cellularity and microstructural changes, aiding in early diagnosis. Furthermore, arterial spin labeling (ASL) perfusion MRI and histogram analyses have been shown to differentiate early-stage NPC from lymphoid hyperplasia based on blood flow metrics, with variance in ASL blood flow identified as an independent predictor [46].

Collectively, these advances underscore MRI's expanding role in early NPC screening. Rapid, non-contrast MRI protocols facilitate detection of endoscopically occult lesions; refined MRI grading systems improve specificity and reduce unnecessary invasive procedures; AI-driven image analysis enhances diagnostic accuracy; and advanced functional sequences provide valuable physiological insights. Integration

of these approaches into regional health data platforms can optimize screening strategies, enabling earlier diagnosis and improved patient outcomes.

2.3.2. Advances in Endoscopic Technology and Combined Diagnostic Approaches

The evolution of endoscopic techniques has significantly enhanced the detection and characterization of nasopharyngeal lesions, particularly in the early diagnosis of NPC. Among these advancements, Narrow Band Imaging (NBI) endoscopy and Lugol's iodine staining have emerged as pivotal modalities that improve visualization of mucosal and vascular changes associated with malignant and premalignant conditions. NBI utilizes specific light wavelengths to accentuate the microvascular architecture and mucosal patterns, thereby facilitating the identification of abnormal vascular morphology and subtle mucosal alterations in the nasopharyngeal region. This enhanced visualization aids in distinguishing neoplastic lesions from benign inflammatory changes, which is critical for accurate diagnosis and timely intervention.

Lugol's iodine staining complements NBI by selectively staining glycogen-rich normal squamous epithelium, leaving dysplastic or malignant areas unstained or lightly stained, thus highlighting suspicious lesions. The combined use of NBI and Lugol's iodine staining has demonstrated improved sensitivity and specificity in differentiating nasopharyngeal carcinoma from chronic hyperplastic nasopharyngitis and other benign conditions. For instance, a study involving 159 patients showed that the diagnostic performance of combined conventional white light endoscopy, NBI, and Lugol's iodine staining was significantly superior to any single modality alone, with enhanced accuracy in detecting nasopharyngeal carcinoma [47].

In addition to these optical enhancements, AI integration into nasopharyngeal endoscopy has shown promising results in improving diagnostic accuracy and efficiency. The team led by Professor Hongmeng Yu at Fudan University Eye and ENT Hospital has pioneered the application of AI for nasopharyngeal endoscopic diagnosis of NPC. Utilizing deep learning algorithms, such as CNNs and You Only Look Once networks, their AI models have achieved high precision, recall, and mean average precision in real-time detection of nasopharyngeal carcinoma during endoscopy. For example, their AI system demonstrated precision and recall rates exceeding 90% in internal test sets, with inference speeds surpassing the typical frame rates of endoscopic video, enabling feasible real-time application [48].

Moreover, multimodal AI models that integrate both white light imaging and NBI have further enhanced diagnostic performance. The Siamese deep convolutional neural network developed by this group simultaneously processes WLE and NBI images, achieving significantly higher accuracy and AUC compared to models using either modality alone [49]. These AI-assisted diagnostic tools not only aid expert endos-

copists but also substantially improve the diagnostic capabilities of less experienced clinicians in primary care settings, thereby addressing disparities in NPC detection.

The synergistic use of advanced optical imaging and AI-based analysis represents a new frontier in nasopharyngeal endoscopy. It enables more precise identification of vascular and mucosal abnormalities, guides targeted biopsies, and supports early diagnosis, which is essential for improving prognosis in NPC patients. Future directions include further validation of AI algorithms in larger, multicenter cohorts and integration with molecular imaging techniques to provide comprehensive, non-invasive diagnostic platforms. Collectively, these technological advancements hold great promise for transforming NPC screening and diagnosis, particularly in resource-limited settings where expert endoscopic evaluation may not be readily available.

2.3.3. Deep Learning-based Automated Imaging Analysis

Deep learning, particularly CNNs, has revolutionized the automated analysis of medical images, offering promising advancements in the early detection and classification of NPC. CNN models have demonstrated high efficacy in classifying T2-weighted MRI scans to distinguish NPC from benign lesions, achieving an accuracy as high as 91.5%. This performance reflects the ability of CNNs to extract hierarchical and discriminative features from complex imaging data, surpassing traditional manual interpretation challenges posed by the nasopharynx's intricate anatomy and subtle tumor characteristics [50]. The development of optimized CNN architectures has further enhanced image quality and diagnostic accuracy by improving parameters like Dice coefficient, peak signal-to-noise ratio, and structural similarity index, leading to diagnostic accuracies exceeding 90% when combining T2-weighted imaging with diffusion-weighted and dynamic contrast-enhanced MRI sequences [51]. Moreover, weakly supervised deep learning approaches have facilitated automated T staging of NPC on MR images without requiring extensive slice-level annotations, achieving accuracies around 75.6% and AUC values of 0.943, comparable to conventional TNM staging in prognostic performance [52]. Beyond MRI, deep learning models have been applied to endoscopic imaging, where CNNs and transformer-based networks have successfully differentiated NPC from benign hyperplasia and normal nasopharynx with accuracy rates exceeding 90%, thereby aiding early detection and potentially reducing missed diagnoses in primary healthcare settings [49, 53]. The integration of multimodal imaging data, such as white light and narrow-band imaging, into deep learning frameworks like Siamese DCNNs has further improved classification performance, with accuracies approaching 95% and AUCs near 0.99, underscoring the value of combining complementary imaging modalities [49]. Additionally, advanced segmentation models utilizing deep learning, including U-Net variants and transformer-integrated

architectures, have addressed the challenges of NPC tumor delineation on MRI and PET images, achieving Dice similarity coefficients up to 0.87 and enhancing the precision of radiotherapy planning [54, 55]. These automated segmentation tools are critical given the tumor's small size and variable morphology, which complicate manual contouring. Furthermore, deep learning models have been employed to predict clinical outcomes such as distant metastasis and treatment response by analyzing intratumoral and peritumoral MRI features, with certain CNN architectures achieving AUCs up to 0.88 for metastasis prediction and providing individualized survival estimations [56, 57]. The application of explainable AI techniques, including Grad-CAM visualization, has enhanced the interpretability of these models, facilitating clinical trust and adoption. Collectively, these studies highlight that CNN-based deep learning models applied to T2-weighted MRI and complementary imaging modalities offer a highly accurate, automated approach for NPC detection, classification, staging, and prognosis. The continued refinement of these algorithms, alongside the creation of large-scale, well-annotated datasets, is essential to translate these promising technologies into routine clinical practice, ultimately improving early NPC screening and patient outcomes [50-52].

2.4. Construction of Regional Health Medical Big Data Platform and Multimodal Data Fusion

2.4.1. Advantages of Big Data Platforms Integrating Multi-source Data

The integration of multi-source data through big data platforms presents significant advantages in the context of NPC early screening, particularly by unifying diverse data types such as electronic health records (EHRs), imaging data, molecular testing results, and epidemiological information. A well-constructed big data platform enables the consolidation of heterogeneous data from multiple systems, facilitating standardized data management and comprehensive analysis. For example, the nasopharyngeal carcinoma-specific big data platform developed at Southern Hospital integrates data from 15 EHR systems, encompassing 14 modules and 640 fields, including demographic details, laboratory tests, vital signs, and imaging reports. The use of advanced information technologies such as machine learning and natural language processing ensures data cleaning, structuring, and standardization, which are critical for maintaining high data quality and completeness [58]. This unified data environment allows for efficient data retrieval and supports real-world studies on NPC risk factors, early diagnosis, treatment efficacy, and toxicity prevention, thereby enhancing research productivity and reducing costs.

Moreover, big data platforms facilitate precise stratification of screening populations by leveraging integrated multi-di-

mensional data. The ability to analyze combined clinical, molecular, and epidemiological data enables identification of high-risk subgroups, optimizing resource allocation and improving screening efficiency. For instance, data-driven stratification can guide targeted screening efforts towards individuals with specific genetic markers, imaging abnormalities, or epidemiological risk profiles, thereby increasing the yield of early NPC detection and minimizing unnecessary procedures in low-risk groups. This precision approach aligns with the concept of people-centered and integrated health services, where big data analytics support personalized health plans and risk prediction models [59]. Additionally, the integration of patient-generated health data into EHRs, although still in early stages, holds promise for continuous monitoring and dynamic risk assessment, further refining population stratification [60].

From a technical perspective, big data platforms employ scalable architectures capable of handling massive, heterogeneous datasets, ensuring real-time or near-real-time data integration and analysis. The West China Hospital of Sichuan University big data platform exemplifies this by supporting over 12 million patients and 75 million visits with 8,475 data variables, integrating data across all hospital departments via a master-slave mode and business-based metadata models for quality control [61]. Such platforms enable seamless data flow from clinical systems, imaging archives, molecular diagnostic laboratories, and public health databases, ensuring comprehensive data availability for NPC screening programs.

Furthermore, big data platforms enhance data security, privacy, and governance through encryption, de-identification, and adherence to regulatory standards, which are essential when handling sensitive health information from multiple sources [48]. They also support interoperability by adopting common data standards and semantic integration techniques, which address heterogeneity and facilitate reproducibility of analyses [62]. The integration of blockchain technology and distributed file systems has been proposed to further improve data resilience and trustworthiness in health data ecosystems [63].

In summary, big data platforms integrating multi-source data offer a unified, high-quality, and secure data environment that supports comprehensive analysis and research in NPC early screening. They enable precise population stratification, optimize resource allocation, and improve screening effectiveness. The continuous advancement in data integration technologies, standardization, and governance will further enhance the utility of these platforms, driving progress in NPC early detection and personalized medicine.

2.4.2. Multimodal Data Fusion and Machine Learning Model Construction

The integration of multimodal data sources—such as EBV biomarkers, DNA methylation profiles, and imaging features—through advanced machine learning algorithms has emerged as a promising strategy for constructing multidimensional risk prediction models for early NPC detection. This

approach leverages the complementary strengths of diverse data types to capture the complex biological and clinical heterogeneity of NPC, which traditional single-modality models often fail to address. For instance, EBV-related serological markers, including viral capsid antigen immunoglobulin A and nuclear antigen 1 immunoglobulin A, have been consistently identified as key predictors of NPC risk, with ML models such as XGBoost achieving excellent discrimination performance ($AUC > 0.95$) in both internal and external validation cohorts [64]. Simultaneously, DNA methylation biomarkers and differentially expressed genes identified via bioinformatics analyses provide molecular-level insights that can enhance the specificity of NPC risk stratification [65]. Complementing these molecular markers, radiomic features extracted from pretreatment MRI, including morphology, texture, and intensity statistics, have been successfully integrated with deep learning-derived features (e.g., stacked denoising autoencoder outputs) and clinical parameters to build robust recurrence prediction models using support vector machines [66]. These multimodal models have demonstrated superior performance in early NPC detection, with reported AUCs ranging from 0.83 to 0.99, high accuracy, and improved sensitivity and specificity compared to unimodal approaches.

The fusion of heterogeneous data modalities is typically achieved through feature-level integration, where selected features from each modality are concatenated into a unified input vector for ML classifiers, or through model-level fusion strategies that combine outputs from modality-specific models. Studies in related oncologic and clinical domains have shown that early fusion methods often yield better performance and require fewer labeled samples compared to late fusion [67, 68]. Moreover, advanced ensemble learning techniques that integrate multiple fusion strategies can further enhance predictive accuracy and robustness [69]. The choice of ML algorithms is critical; gradient boosting decision trees (GBDT), XGBoost, and support vector machines have been favored for their ability to handle high-dimensional, heterogeneous data and mitigate overfitting, especially in relatively small sample sizes common in NPC studies [70, 71]. Additionally, interpretability tools such as SHapley Additive exPlanations facilitate the identification of the most influential features, aiding clinical translation [70].

The clinical utility of these multimodal ML models is underscored by their capacity to improve early NPC detection rates and reduce false positives, addressing key limitations of current screening methods that rely heavily on invasive biopsies and singular biomarker assays. For example, models integrating EBV serology, hematological parameters, and imaging features have demonstrated high specificity (~95%) and sensitivity (~88%), enabling effective risk stratification and timely intervention [64, 66]. Furthermore, noninvasive approaches such as saliva-based DNA detection combined with surface-enhanced Raman spectroscopy and ML classification have shown promising sensitivity and specificity (~74% and 77%, respectively), highlighting the potential of multimodal

fusion in population-wide screening [72]. Despite these advances, challenges remain, including the need for multicenter prospective validation, incorporation of functional imaging and molecular data, and development of models robust to missing modalities [66, 73]. Emerging transformer-based architectures and contrastive learning frameworks offer promising avenues to enhance cross-modal interactions and handle incomplete data scenarios [73].

In summary, the construction of multimodal risk prediction models for NPC through the fusion of EBV biomarkers, DNA methylation data, and imaging features using sophisticated ML algorithms represents a significant advancement in early screening strategies. These models not only improve diagnostic accuracy but also provide a framework for personalized risk assessment and clinical decision support, ultimately contributing to better patient outcomes in NPC management.

2.4.3. Challenges and Future Directions in Platform Applications

The construction and application of regional health medical big data platforms for NPC early screening face several pivotal challenges, primarily revolving around data standardization, privacy protection, and cross-institutional data sharing. Data standardization is foundational to ensure interoperability and meaningful integration of heterogeneous data sources, including electronic health records, imaging, laboratory tests, and molecular biomarkers. For instance, a comprehensive NPC-specific big data platform integrating 15 EHR systems demonstrated the feasibility of harmonizing diverse data elements, achieving high completeness and accuracy for key clinical variables [58]. However, standardizing data formats, terminologies, and coding systems across different healthcare institutions remains a complex task that requires consensus on data element definitions and adoption of industry-wide standards. Without such standardization, the aggregation and comparative analysis of data for real-world studies and AI model training are significantly hindered.

Privacy protection is another critical concern, especially given the sensitive nature of health data and the increasing regulatory requirements worldwide. The platform must implement robust data encryption, de-identification, and access control measures to safeguard patient confidentiality while enabling legitimate research use. Balancing data utility and privacy remains a delicate issue, necessitating advanced techniques such as federated learning or secure multi-party computation to allow collaborative analysis without direct data sharing. This is particularly relevant in NPC research, where multicenter data pooling is essential to gather sufficient sample sizes for early detection biomarker validation and AI model development [58, 74].

Cross-institutional data sharing is indispensable for enhancing the statistical power and generalizability of NPC early screening strategies. Current platforms are mostly limited to single institutions or regions, which restricts the diversity of

patient populations and clinical scenarios represented. Multicenter collaboration can facilitate the development of robust diagnostic models, such as deep learning algorithms for endoscopic image analysis that have demonstrated high accuracy and generalizability across multiple hospitals [5, 53]. However, differences in data governance policies, technical infrastructures, and resource availability pose significant barriers to seamless data sharing. Establishing common data governance frameworks and interoperable technical solutions is crucial to overcome these obstacles.

Looking forward, future directions should focus on strengthening multicenter collaborations to build larger, more diverse NPC datasets that reflect real-world clinical heterogeneity. Moreover, this approach will improve the robustness and clinical applicability of screening tools, including AI-assisted diagnostic systems and molecular biomarker panels [15, 74]. Furthermore, the integration of real-time dynamic monitoring capabilities into platforms can enable continuous patient surveillance, early detection of recurrence, and timely intervention. For example, liquid biopsy approaches using circulating cell-free DNA methylation patterns or EBV DNA signatures have shown promise for non-invasive, dynamic NPC detection and monitoring, which could be incorporated into big data platforms for longitudinal tracking [29, 75].

Additionally, the development of personalized screening programs tailored to individual risk profiles represents a promising future direction. Prognostic nomograms and risk stratification models derived from big data analyses can guide risk-adapted screening intervals and modalities, optimizing resource allocation and patient outcomes [76, 77]. The incorporation of AI-driven predictive analytics can further refine personalized screening by continuously learning from accumulating data and adjusting recommendations accordingly.

In summary, addressing the challenges of data standardization, privacy protection, and cross-institutional data sharing is essential for the successful construction and application of NPC early screening platforms. Future efforts should prioritize multicenter collaboration, real-time dynamic monitoring, and personalized screening program development to fully leverage big data and AI technologies for improving early detection and clinical management of nasopharyngeal carcinoma.

3. Conclusion

In conclusion, the integration of regional health big data platforms with multimodal diagnostic approaches marks a significant advancement in the early screening of NPC. From an expert perspective, this review highlights how the convergence of EBV-related assays, DNA methylation biomarkers, advanced imaging technologies, and AI-assisted diagnostics has transformed NPC detection strategies. The combined use of multi-omics data and machine learning algorithms has notably enhanced screening performance metrics—particularly sensitivity, specificity, and positive predictive value—thereby enabling more accurate identification of early-stage NPC

cases.

This multimodal fusion approach addresses critical limitations of traditional screening methods, offering a non-invasive, cost-effective, and convenient solution suitable for large-scale population screening. These innovations are especially impactful for high-risk groups, facilitating precise risk stratification and timely clinical intervention that can substantially improve patient outcomes. The ability to integrate multiple data types—such as molecular, imaging, and clinical information—and apply advanced analytical techniques exemplifies the power of interdisciplinary collaboration in advancing oncological diagnostics.

Nevertheless, implementing these sophisticated screening frameworks presents challenges. Data integration complexities, concerns over patient privacy and data security, and the lack of standardized protocols remain significant barriers. Overcoming these obstacles will require concerted efforts across multiple disciplines, including bioinformatics, clinical oncology, data governance, and health policy. Additionally, ongoing technological innovation is essential to refine analyt-

ical models, enhance interpretability, and ensure the scalability of screening programs.

Looking ahead, the development of precision screening strategies grounded in comprehensive big data platforms is poised to become a cornerstone in early NPC screening and prevention. These strategies promise to improve individual patient prognoses and elevate regional public health standards by enabling proactive disease management. Continued research and collaboration will be vital to fully realize the potential of these integrated approaches, ultimately transforming NPC screening from a reactive to a predictive paradigm.

In summary, this review underscores the transformative impact of multimodal data integration and AI-driven analytics in early NPC detection. Balancing promising research findings with pragmatic considerations of implementation challenges provides a roadmap for future advancements. By fostering interdisciplinary partnerships and prioritizing innovation, the medical community can enhance early NPC screening efficacy, reduce disease burden, and improve survival outcomes at the population level (Figure 1).

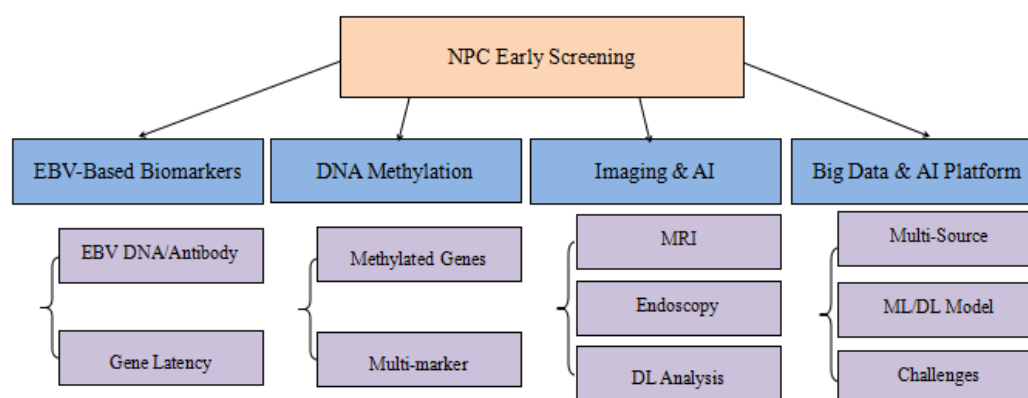


Figure 1. Different techniques for early screening of nasopharyngeal carcinoma.

Abbreviations

NPC	Nasopharyngeal Carcinoma
EBV	Epstein-Barr Virus
MRI	Magnetic Resonance Imaging
CT	Computed Tomography
SPECT	Single-photon Emission Computed Tomography
PET	Positron Emission Tomography
AI	Artificial Intelligence
AUC	Area Under the Curve
PPV	Positive Predictive Value
NGS	Next-generation Sequencing
LMP1	Latent Membrane Protein 1
EBERs	EBV-encoded Small RNAs
IGF-1R	Insulin-like Growth Factor 1 Receptor
LSD1	Lysine-specific Demethylase 1

ZNF217	Zinc Finger Protein 217
MI	Methylation Index
cfDNA	Circulating Cell-free DNA
EBNA1-IgA	IgA Antibodies Against EBV Nuclear Antigen 1
VCA-IgA	Viral Capsid Antigen
PPV	Positive Predictive Value
CNN	Convolutional Neural Network
DWI	Diffusion-weighted Imaging
ASL	Arterial Spin Labeling
NBI	Narrow Band Imaging
EHR	Selectronic Health Records

Acknowledgments

The author thanks Prof. Peng Shen, Yinzhou District Center for Disease Control and Prevention, for providing big data support.

Author Contributions

Guowen Zhan is the sole author. The author read and approved the final manuscript.

Funding

This work is supported by Yinzhou District Health Science and Technology (Grant No.2024Y15).

Conflicts of Interest

The author declares no conflicts of interest.

References

- [1] Yildirim HC, Kupik GE, Mustafayev TZ, et al. A multicenter retrospective analysis of patients with nasopharyngeal carcinoma treated in IMRT era from a nonendemic population: Turkish Society for Radiation Oncology Head and Neck Cancer Group Study (TROD 01-001). *Head Neck*. 2023; 45(5): 1194-1205. <https://doi.org/10.1002/hed.27333>
- [2] Chan SM, Paterson IC, Yap LF. Nasopharyngeal Carcinoma in Southeast Asia: Current Landscape and Future Priorities. *Br J Biomed Sci*. 82: 15902. Published 2025 None. <https://doi.org/10.3389/bjbs.2025.15902>
- [3] Beyene ET, Ketema SG, Alebachew AN, et al. Descriptive epidemiology of nasopharyngeal carcinoma at Tikur Anbessa Hospital, Ethiopia. *BMC Cancer*. 2021; 21(1): 540. <https://doi.org/10.1186/s12885-021-08311-8>
- [4] Reffai A, Mesmoudi M, Derkaoui T, et al. Epidemiological Profile and Clinicopathological, Therapeutic, and Prognostic Characteristics of Nasopharyngeal Carcinoma in Northern Morocco. *Cancer Control*. 28: 10732748211050587. <https://doi.org/10.1177/10732748211050587>
- [5] Shi Y, Li Z, Wang L, et al. Artificial intelligence-assisted detection of nasopharyngeal carcinoma on endoscopic images: a national, multicentre, model development and validation study. *Lancet Digit Health*. 2025; 7(6): 100869. <https://doi.org/10.1016/j.landig.2025.03.001>
- [6] Xu X, Tang X, Wu W, et al. Radiopharmaceuticals in Nasopharyngeal Cancer. *Bioorg Chem*. 157: 108281. <https://doi.org/10.1016/j.bioorg.2025.108281>
- [7] Qu B, Sun L, Deng H, Liu Q, et al. Diagnostic value of BNLF2b antibody, dual-antibody testing and Epstein-Barr virus DNA in nasopharyngeal carcinoma: a prospective cohort study in Hunan Province, China. *BMJ Open*. 2025; 15(5): e100538. <https://doi.org/10.1136/bmjopen-2025-100538>
- [8] Lam WKJ, King AD, Miller JA, et al. Recommendations for Epstein-Barr virus-based screening for nasopharyngeal cancer in high- and intermediate-risk regions. *J Natl Cancer Inst*. 2023; 115(4): 355-364. <https://doi.org/10.1093/jnci/djad012>
- [9] Cai Y, Ke L, Du A, et al. Machine Learning-Assisted Multimodal Early Screening of Lung Cancer Based on a Multiplexed Laser-Induced Graphene Immunosensor. *ACS Nano*. 2025; 19(28): 25697-25709. <https://doi.org/10.1021/acsnano.5c02822>
- [10] Sun K, Wang Y, Qu R, et al. Comprehensive application of artificial intelligence in colorectal cancer: A review. *iScience*. 2025; 28(7): 112980. Published 2025 Jul 18. <https://doi.org/10.1016/j.isci.2025.112980>
- [11] Fu J, Fang M, Wu L, et al. Development, advancement, and clinical integration of artificial intelligence technology in gastric cancer. *Chin Med J (Engl)*. Published online Nov 28, 2025. <https://doi.org/10.1097/CM9.0000000000003922>
- [12] Yao L, Liu Y, Wang T, et al. Global trends of big data analytics in health research: a bibliometric study. *Front Med (Lausanne)*. 12: 1456286. Published 2025 None. <https://doi.org/10.3389/fmed.2025.1456286>
- [13] Liu W, Chen G, Gong X, et al. The diagnostic value of EBV-DNA and EBV-related antibodies detection for nasopharyngeal carcinoma: a meta-analysis. *Cancer Cell Int*. 2021; 21(1): 164. <https://doi.org/10.1186/s12935-021-01862-7>
- [14] Li T, Li F, Guo X, et al. Anti-Epstein-Barr Virus BNLF2b for Mass Screening for Nasopharyngeal Cancer. *N Engl J Med*. 2023; 389(9): 808-819. <https://doi.org/10.1056/NEJMoa2301496>
- [15] Lou PJ, Jacky Lam WK, Hsu WL, et al. Performance and Operational Feasibility of Epstein-Barr Virus-Based Screening for Detection of Nasopharyngeal Carcinoma: Direct Comparison of Two Alternative Approaches. *J Clin Oncol*. 2023; 41(26): 4257-4266. <https://doi.org/10.1200/JCO.22.01979>
- [16] Lam WKJ, Ma BBY, King AD, et al. Achieving control of nasopharyngeal carcinoma: the role of Epstein-Barr virus-based screening and vaccines. *Nat Rev Clin Oncol*. 2026; 23(1): 7-21. <https://doi.org/10.1038/s41571-025-01079-x>
- [17] Simon J, Brenner N, Reich S, et al. Nasopharyngeal carcinoma patients from Norway show elevated Epstein-Barr virus IgA and IgG antibodies prior to diagnosis. *Cancer Epidemiol*. 77: 102117. <https://doi.org/10.1016/j.canep.2022.102117>
- [18] Zhu X, Perales-Puchalt A, Wojtak K, et al. DNA immunotherapy targeting BARF1 induces potent anti-tumor responses against Epstein-Barr-virus-associated carcinomas. *Mol Ther Oncolytics*. 24: 218-229. <https://doi.org/10.1016/j.omto.2021.12.017>
- [19] Chen GH, Liu Z, Yu KJ, et al. Utility of Epstein-Barr Virus DNA in Nasopharynx Swabs as a Reflex Test to Triage Seropositive Individuals in Nasopharyngeal Carcinoma Screening Programs. *Clin Chem*. 2022; 68(7): 953-962. <https://doi.org/10.1093/clinchem/hvac032>
- [20] Lv M, Shi D, Zhao X, et al. IGFBP2 up-regulation by EBV via TGF- β signaling: a key mechanism in nasopharyngeal carcinoma progression. *Virus Genes*. 2025; 61(5): 562-573. <https://doi.org/10.1007/s11262-025-02178-8>

- [21] Shareena G, Kumar D. Epigenetics of Epstein Barr virus - A review. *Biochim Biophys Acta Mol Basis Dis*. 2023; 1869(8): 166838. <https://doi.org/10.1016/j.bbdis.2023.166838>
- [22] Caruso LB, Guo R, Keith K, et al. The nuclear lamina binds the EBV genome during latency and regulates viral gene expression. *PLoS Pathog*. 2022; 18(4): e1010400. <https://doi.org/10.1371/journal.ppat.1010400>
- [23] Lao TD, Thieu HH, Nguyen DH, et al. Hypermethylation of the RASSF1A gene promoter as the tumor DNA marker for nasopharyngeal carcinoma. *Int J Biol Markers*. 2022; 37(1): 31-39. <https://doi.org/10.1177/17246008211065472>
- [24] Lyu JY, Chen JY, Zhang XJ, et al. Septin 9 Methylation in Nasopharyngeal Swabs: A Potential Minimally Invasive Biomarker for the Early Detection of Nasopharyngeal Carcinoma. *Dis Markers*. 2020: 7253531. <https://doi.org/10.1155/2020/7253531>
- [25] Thieu HH, Le TAH, Lao TD. Methylation of Tumor Suppressor Genes on Chromosome 9: Diagnostic Insights from a Nasopharyngeal Carcinoma in Vietnam. *Asian Pac J Cancer Prev*. 2025; 26(6): 2239-2245. <https://doi.org/10.31557/APJCP.2025.26.6.2239>
- [26] Qin ZH, Chen SY, Zhou S, et al. Nasopharyngeal carcinoma detected noninvasively in the real world using three gene methylation analyses from automatically processed bilateral nasal swab samples. *BMC Cancer*. 2025; 25(1): 1147. <https://doi.org/10.1186/s12885-025-14508-y>
- [27] Xu Y, Zhao W, Mo Y, et al. Combination of RERG and ZNF671 methylation rates in circulating cell-free DNA: A novel biomarker for screening of nasopharyngeal carcinoma. *Cancer Sci*. 2020; 111(7): 2536-2545. <https://doi.org/10.1111/cas.14431>
- [28] Lu Z, Ye Z, Li P, et al. An MSRE-Assisted Glycerol-Enhanced RPA-CRISPR/Cas12a Method for Methylation Detection. *Biosensors (Basel)*. 2024; 14(12). <https://doi.org/10.3390/bios14120608>
- [29] Fu XY, Zhou ZY, Yang TY, et al. Plasma cfDNA VILL gene methylation as a diagnostic marker for nasopharyngeal carcinoma. *Clin Epigenetics*. 2025; 17(1): 38. <https://doi.org/10.1186/s13148-025-01847-7>
- [30] Ma L, Wang TM, He YQ, et al. Multiplex assays reveal anti-EBV antibody profile and its implication in detection and diagnosis of nasopharyngeal carcinoma. *Int J Cancer*. 2024; 155(10): 1874-1885. <https://doi.org/10.1002/ijc.35061>
- [31] Chen GH, Liu Z, Ji MF, et al. Prospective assessment of a nasopharyngeal carcinoma risk score in a population undergoing screening. *Int J Cancer*. 2021; 148(10): 2398-2406. <https://doi.org/10.1002/ijc.33424>
- [32] Yuan Y, Ye F, Wu JH, et al. Early screening of nasopharyngeal carcinoma. *Head Neck*. 2023; 45(10): 2700-2709. <https://doi.org/10.1002/hed.27466>
- [33] Tang C, Li X, Zhang Y, et al. Blind Brush Biopsy: Quantification of Epstein-Barr Virus and Its Host DNA Methylation in the Detection of Nasopharyngeal Carcinoma. *Research (Wash D C)*. 7: 0475. <https://doi.org/10.34133/research.0475>
- [34] Zheng XH, Li XZ, Tang CL, et al. Detection of Epstein-Barr virus DNA methylation as tumor markers of nasopharyngeal carcinoma patients in saliva, oropharyngeal swab, oral swab, and mouthwash. *MedComm* (2020). 2024; 5(9): e673. <https://doi.org/10.1002/mco2.673>
- [35] Huang Q, Hou T, Ren Y, et al. Diagnostic and prognostic value of EphB2 in nasopharyngeal carcinoma. *Transl Oncol*. 64: 102641. <https://doi.org/10.1016/j.tranon.2025.102641>
- [36] Wang H, Zhang J, Cheng P, et al. Integrating transcriptomics and hybrid machine learning enables high-accuracy diagnostic modeling for nasopharyngeal carcinoma. *Discov Oncol*. 2025; 16(1): 1067. <https://doi.org/10.1007/s12672-025-02932-2>
- [37] Li Y, Pan Y, Huang Z, et al. Comorbidity profiling identifies potential subtype of elderly patients with nasopharyngeal carcinoma. *Oncologist*. 2024; 29(8): e1020-e1030. <https://doi.org/10.1093/oncolo/oyae063>
- [38] Ghose S, Roy S, Ghosh V, et al. The plasma EBV DNA load with IL-6 and VEGF levels as predictive and prognostic biomarker in nasopharyngeal carcinoma. *Viro J*. 2024; 21(1): 224. Published 2024 Sep 20. <https://doi.org/10.1186/s12985-024-02473-0>
- [39] Peng L, Wan L, Liu M, et al. Diagnostic and prognostic performance of plasma neurofilament light chain in multiple system atrophy: a cross-sectional and longitudinal study. *J Neurol*. 2023; 270(9): 4248-4261. <https://doi.org/10.1007/s00415-023-11741-y>
- [40] Trieu C, van Harten AC, van Leeuwenstijn MSSA, et al. Longitudinal Blood-Based Biomarkers and Clinical Progression in Subjective Cognitive Decline. *JAMA Netw Open*. 2025; 8(12): e2545862. Published 2025 Dec 1. <https://doi.org/10.1001/jamanetworkopen.2025.45862>
- [41] King AD, Ai QYH, Lam WKJ, et al. Early detection of nasopharyngeal carcinoma: performance of a short contrast-free screening magnetic resonance imaging. *J Natl Cancer Inst*. 2024; 116(5): 665-672. <https://doi.org/10.1093/jnci/djad260>
- [42] King AD, Woo JKS, Ai QY, et al. Early Detection of Cancer: Evaluation of MR Imaging Grading Systems in Patients with Suspected Nasopharyngeal Carcinoma. *AJNR Am J Neuroradiol*. 2020; 41(3): 515-521. <https://doi.org/10.3174/ajnr.A6444>
- [43] Wong LM, King AD, Ai QYH, et al. Convolutional neural network for discriminating nasopharyngeal carcinoma and benign hyperplasia on MRI. *Eur Radiol*. 2021; 31(6): 3856-3863. <https://doi.org/10.1007/s00330-020-07451-y>
- [44] Wong LM, Ai QYH, Zhang R, et al. Radiomics for Discrimination between Early-Stage Nasopharyngeal Carcinoma and Benign Hyperplasia with Stable Feature Selection on MRI. *Cancers (Basel)*. 2022; 14(14). <https://doi.org/10.3390/cancers14143433>
- [45] Meng T, Liu H, Liu J, et al. The investigation of reduced field-of-view diffusion-weighted imaging (DWI) in patients with nasopharyngeal carcinoma: comparison with conventional DWI. *Acta Radiol*. 2023; 64(6): 2118-2125. <https://doi.org/10.1177/02841851231159389>

- [46] Xiao B, Wang P, Zhao Y, et al. Using arterial spin labeling blood flow and its histogram analysis to distinguish early-stage nasopharyngeal carcinoma from lymphoid hyperplasia. *Medicine (Baltimore)*. 2021; 100(8): e24955. <https://doi.org/10.1097/MD.00000000000024955>
- [47] Yang F, Huang N, Chen X, et al. Application of narrow band imaging and Lugol's iodine staining in screening for nasopharyngeal carcinoma. *World J Surg Oncol*. 2023; 21(1): 376. <https://doi.org/10.1186/s12957-023-03258-5>
- [48] He Z, Zhang K, Zhao N, et al. Deep learning for real-time detection of nasopharyngeal carcinoma during nasopharyngeal endoscopy. *iScience*. 2023; 26(10): 107463. <https://doi.org/10.1016/j.isci.2023.107463>
- [49] Xu J, Wang J, Bian X, et al. Deep Learning for nasopharyngeal Carcinoma Identification Using Both White Light and Narrow-Band Imaging Endoscopy. *Laryngoscope*. 2022; 132(5): 999-1007. <https://doi.org/10.1002/lary.29894>
- [50] Li S, Deng YQ, Zhu ZL, et al. A Comprehensive Review on Radiomics and Deep Learning for Nasopharyngeal Carcinoma Imaging. *Diagnostics (Basel)*. 2021; 11(9). <https://doi.org/10.3390/diagnostics11091523>
- [51] Huang R, Zhou Z, Wang X, et al. Magnetic Resonance Imaging Features on Deep Learning Algorithm for the Diagnosis of Nasopharyngeal Carcinoma. *Contrast Media Mol Imaging*. 2022; 3790269. <https://doi.org/10.1155/2022/3790269>
- [52] Yang Q, Guo Y, Ou X, et al. Automatic T Staging Using Weakly Supervised Deep Learning for Nasopharyngeal Carcinoma on MR Images. *J Magn Reson Imaging*. 2020; 52(4): 1074-1082. <https://doi.org/10.1002/jmri.27202>
- [53] Yue Y, Zeng X, Lin H, et al. A deep learning based smartphone application for early detection of nasopharyngeal carcinoma using endoscopic images. *NPJ Digit Med*. 2024; 7(1): 384. <https://doi.org/10.1038/s41746-024-01403-2>
- [54] Wang CK, Wang TW, Yang YX, et al. Deep Learning for Nasopharyngeal Carcinoma Segmentation in Magnetic Resonance Imaging: A Systematic Review and Meta-Analysis. *Bioengineering (Basel)*. 2024; 11(5). <https://doi.org/10.3390/bioengineering11050504>
- [55] Zeng Y, Li J, Zhao Z, et al. WET-UNet: Wavelet integrated efficient transformer networks for nasopharyngeal carcinoma tumor segmentation. *Sci Prog*. 2024 Apr-Jun; 107(2): 368504241232537. <https://doi.org/10.1177/00368504241232537>
- [56] Hua HL, Deng YQ, Li S, et al. Deep Learning for Predicting Distant Metastasis in Patients with Nasopharyngeal Carcinoma Based on Pre-Radiotherapy Magnetic Resonance Imaging. *Comb Chem High Throughput Screen*. 2023; 26(7): 1351-1363. <https://doi.org/10.2174/1386207325666220919091210>
- [57] Cao X, Chen X, Lin ZC, et al. Add-on individualizing prediction of nasopharyngeal carcinoma using deep-learning based on MRI: A multicentre, validation study. *iScience*. 2022; 25(9): 104841. <https://doi.org/10.1016/j.isci.2022.104841>
- [58] Chen ZK, Wang XQ, Xiao LL, et al. Construction and application of nasopharyngeal carcinoma-specific big data platform based on electronic health records. *Am J Otolaryngol*. 2024 May-Jun; 45(3): 104204. <https://doi.org/10.1016/j.amjoto.2023.104204>
- [59] Schulte T, Bohnet-Joschko S. How can Big Data Analytics Support People-Centred and Integrated Health Services: A Scoping Review. *Int J Integr Care*. 2022 Apr-Jun; 22(2): 23. <https://doi.org/10.5334/ijic.5543>
- [60] Tiase VL, Hull W, McFarland MM, et al. Patient-generated health data and electronic health record integration: a scoping review. *JAMIA Open*. 2020; 3(4): 619-627. <https://doi.org/10.1093/jamiaopen/ooaa052>
- [61] Wang M, Li S, Zheng T, et al. Big Data Health Care Platform With Multisource Heterogeneous Data Integration and Massive High-Dimensional Data Governance for Large Hospitals: Design, Development, and Application. *JMIR Med Inform*. 2022; 10(4): e36481. <https://doi.org/10.2196/36481>
- [62] Zhang H, Lyu T, Yin P, et al. A scoping review of semantic integration of health data and information. *Int J Med Inform*. 165: 104834. <https://doi.org/10.1016/j.ijmedinf.2022.104834>
- [63] Shukla M, Lin J, Seneviratne O. BlockIoT: Blockchain-based Health Data Integration using IoT Devices. *AMIA Annu Symp Proc*. 2021: 1119-1128.
- [64] Yang W, Zhou C, Tang M, et al. Machine learning-based prediction of nasopharyngeal carcinoma risk: a clinical approach. *Front Immunol*. 16: 1648648. <https://doi.org/10.3389/fimmu.2025.1648648>
- [65] Fu C, Sun L, Zhang L, et al. Identification of key biomarkers and potential therapeutic drugs in nasopharyngeal carcinoma based on comprehensive bioinformatics analysis. *Discov Oncol*. 2025; 16(1): 1189. <https://doi.org/10.1007/s12672-025-03047-4>
- [66] Liu Y, Wang X, Li J, et al. Prediction of Recurrence using a Stacked Denoising Autoencoder and Multifaceted Feature Analysis of Pretreatment MRI in Patients with Nasopharyngeal Carcinoma. *Curr Radiopharm*. 2025; 18(3): 224-243. <https://doi.org/10.2174/0118744710384129250327060846>
- [67] Lopez K, Fodeh SJ, Allam A, et al. Reducing Annotation Burden Through Multimodal Learning. *Front Big Data*. 3: 19. <https://doi.org/10.3389/fdata.2020.00019>
- [68] Mohsen F, Ali H, El Hajj N, et al. Artificial intelligence-based methods for fusion of electronic health records and imaging data. *Sci Rep*. 2022; 12(1): 17981. <https://doi.org/10.1038/s41598-022-22514-4>
- [69] Imrie F, Denner S, Brunschwig LS, et al. Automated Ensemble Multimodal Machine Learning for Healthcare. *IEEE J Biomed Health Inform*. 2025; 29(6): 4213-4226. <https://doi.org/10.1109/JBHI.2025.3530156>
- [70] Li X, Wang Z, Chen W, et al. Construction and validation of a machine learning model to predict the risk of nasopharyngeal carcinoma using multimodal clinical a single-center, retrospective study. *Clin Transl Oncol*. Published online Jul 15, 2025. <https://doi.org/10.1007/s12094-025-03992-0>

- [71] Li Y, Zhang D, Wang Y, et al. Construction of an oligometastatic prediction model for nasopharyngeal carcinoma patients based on pathomics features and dynamic multi-swarm particle swarm optimization support vector machine. *Front Oncol*. 15: 1589919. <https://doi.org/10.3389/fonc.2025.1589919>
- [72] Hou S, Wu Q, Qiu S, et al. Noninvasive nasopharyngeal cancer screening via salivary DNA detection using label-free surface-enhanced Raman spectroscopy. *Talanta*. 299: 129113. <https://doi.org/10.1016/j.talanta.2025.129113>
- [73] Liu J, Capurro D, Nguyen A, et al. Attention-based multimodal fusion with contrast for robust clinical prediction in the face of missing modalities. *J Biomed Inform*. 145: 104466. <https://doi.org/10.1016/j.jbi.2023.104466>
- [74] Yang X, Wu J, Chen X. Application of Artificial Intelligence to the Diagnosis and Therapy of Nasopharyngeal Carcinoma. *J Clin Med*. 2023; 12(9). <https://doi.org/10.3390/jcm12093077>
- [75] Qiu SF, Zhang QZ, Wu ZY, et al. Establishment and validation of circulating cell-free DNA signatures for nasopharyngeal carcinoma detection. *EBioMedicine*. 108: 105321. <https://doi.org/10.1016/j.ebiom.2024.105321>
- [76] Zhang LL, Xu F, He WT, et al. Development and validation of a prognostic nomogram for the pre-treatment prediction of early metachronous metastasis in endemic nasopharyngeal carcinoma: a big-data intelligence platform-based analysis. *Ther Adv Med Oncol*. 12: 1758835920978132. <https://doi.org/10.1177/1758835920978132>
- [77] Wu CF, Lv JW, Lin L, et al. Development and validation of a web-based calculator to predict individualized conditional risk of site-specific recurrence in nasopharyngeal carcinoma: Analysis of 10,058 endemic cases. *Cancer Commun (Lond)*. 2021; 41(1): 37-50. <https://doi.org/10.1002/cac2.12113>