Critical Reviews in Oncology / Hematology 216 (2025) 104980

Contents lists available at ScienceDirect

Critical Reviews in Oncology / Hematology

journal homepage: www.elsevier.com/locate/critrevonc

Predictive models for immune checkpoint inhibitor response in cancer: A review of current approaches and future directions

Emmanuel O. Oisakede [a,b], Oluwatosin Akinro [c], Oluwakemi Jumoke Bello [d], Claret Chinenyenwa Analikwu [e], Eghosasere Egbon [e], David B. Olawade [f,g,h,*]

[a] Department of Clinical Oncology, Leeds Teaching Hospitals Trust, Leeds, UK
[b] Department of Health Research, University of Leeds, Leeds, UK
[c] Coronary Care Unit, Kettering General Hospital NHS Foundation Trust, Kettering, UK
[d] The Clinical Research Centre, The London Clinic, 20 Devonshire Place, London W1G 6BW, United Kingdom
[e] Department of Microbiology, University Hospital Southampton NHS foundation Trust, Hampshire SO16 6YD, United Kingdom
[f] Department of Allied and Public Health, School of Health, Sport and Bioscience, University of East London, London E16 2RD, UK
[g] Department of Research and Innovation, Medway NHS Foundation Trust, Gillingham, Kent ME7 5NY, UK
[h] Department of Public Health, York St John University, London E14 2BA, UK

ARTICLE INFO

Keywords:
Checkpoint inhibitors
Predictive models
Machine learning
Biomarkers
Immunotherapy

ABSTRACT

Checkpoint inhibitors have revolutionised cancer treatment, yet only 20–30 % of patients achieve durable responses, highlighting the critical need for predictive models. This review focuses on PD-1/PD-L1 pathway inhibitors as monotherapy, examining current prediction frameworks spanning biomarker-based approaches, multi-omics integration, mathematical modelling, and artificial intelligence applications. Recent advances include SCORPIO and LORIS machine learning systems demonstrating superior statistical performance compared to traditional biomarkers, with area under curve values of 0.763. However, critical analysis reveals significant limitations in external validation across diverse healthcare settings, with many promising models failing to maintain performance outside their development institutions. Traditional pathological assessment by expert pathologists, including standardised PD-L1 scoring and tumour-infiltrating lymphocyte quantification, continues to form the foundation of clinical decision-making and provides essential validation for emerging AI approaches. Despite extensive research, established biomarkers show limited predictive accuracy, with PD-L1 demonstrating predictive value in only 28.9 % of FDA approvals. Multi-feature models incorporating genomic and clinical data show improved accuracy but face substantial validation challenges. Integration of spatial biomarkers and digital pathology has enhanced capabilities, achieving area under curve values of 0.84 in select studies. The most critical challenge is the "validation gap", many models show excellent single-institution performance but fail external validation, limiting clinical translation. Current obstacles include inadequate standardisation, interpretability concerns, and healthcare system integration difficulties. Future directions must prioritise rigorous multi-institutional validation studies, development of clinically implementable frameworks, and addressing practical deployment challenges to realise precision immunotherapy's potential.

1. Introduction

Checkpoint inhibitors targeting programmed death protein 1 (PD-1), programmed death-ligand 1 (PD-L1), and cytotoxic T-lymphocyte-associated protein 4 (CTLA-4) have fundamentally transformed cancer treatment paradigms across numerous malignancies. Since the first FDA approval in 2014, these agents have demonstrated remarkable efficacy in previously intractable cancers including melanoma, non-small cell lung cancer (NSCLC), renal cell carcinoma, and head and neck cancers (Flynn and Gerriets, 2019). However, the clinical reality reveals a significant challenge: only 20–40 % of patients achieve durable responses to these therapies, with response rates varying considerably across different tumour types and patient populations (Yang and Hu, 2019; Yarchoan et al., 2017).

* Corresponding author at: Department of Allied and Public Health, School of Health, Sport and Bioscience, University of East London, London E16 2RD, UK.
E-mail addresses: emmanuel.oisakede@gmail.com (E.O. Oisakede), oluwatosin.akinro@yahoo.com (O. Akinro), K.bello01@outlook.com (O.J. Bello), Chinenyecollette@yahoo.com (C.C. Analikwu), eghosaseregabriel@gmail.com (E. Egbon), d.olawade@uel.ac.uk (D.B. Olawade).

https://doi.org/10.1016/j.critrevonc.2025.104980
Received 19 September 2025; Received in revised form 4 October 2025; Accepted 13 October 2025
Available online 20 October 2025

While combination approaches with chemotherapy, targeted therapy, and dual checkpoint inhibition are increasingly important in clinical practice, the prediction of response to these complex regimens presents additional challenges beyond the scope of this review. The focus on monotherapy reflects both the availability of more robust predictive data and the foundational importance of understanding single-agent mechanisms before advancing to combination prediction models.

The limited response rate, combined with the substantial costs and potential for severe immune-related adverse events (irAEs), has driven intensive research into predictive models capable of identifying patients most likely to benefit from checkpoint inhibitor therapy. The global checkpoint inhibitors market, valued at USD 47.4 billion in 2023 and projected to reach USD 189.1 billion by 2032, reflects both the therapeutic potential and economic significance of optimising patient selection (Faizullabhoy and Wani, 2024). This economic imperative underscores the urgent need for reliable predictive tools that can guide treatment decisions, minimise unnecessary toxicity, and maximise therapeutic benefit.

Beyond traditional biomarkers and AI-driven prediction models, emerging research has identified metabolic reprogramming and natural bioactive compounds as potential factors influencing checkpoint inhibitor response. Tumour cells exhibit altered glucose metabolism, characterised by increased expression of glucose transporters such as GLUT1 and GLUT3, which facilitate enhanced glucose uptake to support rapid proliferation (Gökalp 2022a). These metabolic alterations not only serve as potential prognostic indicators but may also influence the tumour microenvironment and immune cell function, thereby affecting immunotherapy efficacy. Furthermore, natural compounds derived from traditional medicinal plants have demonstrated significant anticancer properties through multiple mechanisms, including inhibition of key metabolic enzymes, modulation of signalling pathways, and enhancement of immune responses (Gökalp, 2020, 2021, 2022b). For instance, thymoquinone from black cumin (*Nigella sativa*) has shown potent antitumor activity and immunomodulatory effects across various cancer types (Gökalp, 2025). Similarly, compounds such as cucurbitacins, flavonoids, and organosulfur compounds from plants like marigold (*Tagetes* species) and garlic (*Allium sativum*) have exhibited inhibitory effects on cancer cell proliferation and metastasis (Gökalp, 2023). These findings suggest that integration of metabolic biomarkers and natural compound-based approaches into predictive models could provide a more comprehensive understanding of factors influencing checkpoint inhibitor response.

Current prediction approaches span multiple domains, from traditional biomarker-based strategies to sophisticated artificial intelligence algorithms. These approaches can be categorised into two fundamental components: the underlying biological features or biomarkers (such as PD-L1 expression, tumour mutational burden, immune cell infiltration patterns, and clinical parameters), and the analytical methods used to integrate these features into predictive models (ranging from traditional pathological assessment to advanced machine learning algorithms). The landscape includes single biomarker assessments such as PD-L1 expression and tumour mutational burden (TMB), multi-feature models integrating genomic and clinical data, mathematical simulations of tumour-immune dynamics, and machine learning approaches capable of processing complex, high-dimensional datasets. Despite this diversity of approaches, no single method has achieved universal clinical adoption, reflecting the complex interplay of factors that determine treatment response.

Traditional pathological assessment by expert pathologists continues to form the cornerstone of clinical decision-making, providing standardised, reproducible evaluation of key predictive features including PD-L1 expression levels and tumour-infiltrating lymphocyte patterns. This expertise serves as both the gold standard for biomarker validation and the essential foundation upon which AI-based approaches are developed and validated.

Recent technological advances have significantly expanded predictive capabilities. The development of comprehensive genomic profiling platforms, spatial biomarker analysis, and digital pathology has provided unprecedented insights into the tumour microenvironment. Simultaneously, artificial intelligence applications have matured from experimental concepts to clinically validated tools, with models such as SCORPIO and LORIS demonstrating superior predictive performance compared to traditional biomarkers in real-world patient populations (Yoo et al. 2025; Chang et al. 2024). However, critical evaluation reveals significant limitations in the external validation and clinical implementation of these promising approaches, with many models showing reduced performance when applied outside their original development cohorts.

The rationale for this review stems from the fragmented nature of the current predictive landscape, where multiple approaches exist in isolation without systematic comparison or integration. Whilst previous reviews have focused on individual biomarker categories or specific methodological approaches, no study has comprehensively synthesised the entire spectrum of prediction models spanning traditional biomarkers, advanced artificial intelligence algorithms, mathematical simulations, and integrated multi-modal frameworks. The novelty of this review lies in its unprecedented scope, and providing the first systematic comparison of predictive performance across different methodological approaches using standardised metrics. Additionally, this review provides critical analysis of validation deficits and implementation challenges that have limited the translation of promising research findings into clinical practice.

The primary aim is to provide clinicians and researchers with a unified assessment of all available predictive tools for checkpoint inhibitor response, enabling evidence-based selection of optimal prediction strategies. The specific objectives include: systematically evaluating the predictive performance of biomarker-based, machine learning, mathematical, and integrated approaches; identifying the strengths, limitations, and clinical readiness of each methodology; comparing validation status and real-world applicability across different model types; critically analysing the challenges of external validation and clinical implementation that have hindered widespread adoption; and providing evidence-based recommendations for future research directions and clinical implementation strategies that will advance the field towards reliable, clinically-actionable precision immunotherapy tools.

2. Methods

2.1. Search strategy and data sources

A comprehensive literature search was conducted using multiple electronic databases including PubMed/MEDLINE, Embase, Web of Science, and Cochrane Library to identify relevant studies published between January 2018 and May 2024. The search strategy employed a combination of Medical Subject Headings (MeSH) terms and free-text keywords related to checkpoint inhibitors, predictive models, biomarkers, machine learning, and artificial intelligence. Key search terms included: "checkpoint inhibitor*", "PD-1", "PD-L1", "CTLA-4", "prediction model*", "biomarker*", "machine learning", "artificial intelligence", "immunotherapy response", and "cancer treatment".

Additional sources were systematically reviewed including conference abstracts from major oncology meetings (ASCO, ESMO, AACR), regulatory agency databases (FDA, EMA), and clinical trial registries (ClinicalTrials.gov, EudraCT). Reference lists of identified systematic reviews and meta-analyses were manually screened to ensure comprehensive coverage of the literature.

2.2. Study selection and inclusion criteria

Studies were included if they met the following criteria: original

research articles or systematic reviews examining predictive models for checkpoint inhibitor response in human cancer patients; studies reporting quantitative performance metrics (accuracy, sensitivity, specificity, area under the curve); publications in English language; and availability of full-text articles. Both retrospective and prospective studies were included, encompassing clinical trials, observational cohort studies, and real-world evidence analyses.

Exclusion criteria comprised: preclinical studies using animal models or cell lines only; case reports or case series with fewer than 20 patients; studies focusing solely on safety or pharmacokinetic outcomes without efficacy prediction; duplicate publications or conference abstracts without subsequent peer-reviewed publication; and studies lacking sufficient methodological detail for quality assessment.

2.3. Data extraction and quality assessment

Data extraction was performed systematically using a standardised form capturing study characteristics (design, population, sample size, follow-up duration), intervention details (checkpoint inhibitor type, treatment regimen, combination therapies), predictive model specifications (algorithm type, input variables, training methodology), and outcome measures (response rates, survival endpoints, performance metrics). Particular attention was paid to validation methodology, distinguishing between internal validation, external validation on independent datasets, and real-world clinical implementation. Quality assessment was conducted using appropriate tools including the Quality Assessment of Diagnostic Accuracy Studies (QUADAS-2) for diagnostic studies and the Newcastle-Ottawa Scale for observational studies.

Model performance was evaluated using standard metrics including area under the receiver operating characteristic curve (AUC-ROC), sensitivity, specificity, positive and negative predictive values, and overall accuracy. Where available, external validation results and clinical implementation data were extracted to assess real-world applicability. Critical evaluation focused on the distinction between development cohort performance and external validation results, as this represents a key limitation in current predictive model research.

2.4. Data synthesis and analysis

Given the heterogeneity of study designs, cancer types, and predictive approaches, a narrative synthesis was employed rather than formal meta-analysis. Studies were grouped by prediction methodology (biomarker-based, machine learning, mathematical modelling, and integrated approaches) and cancer type where appropriate. Performance metrics were compared across different approaches, with particular attention to validation status and clinical readiness. Special emphasis was placed on critically evaluating the robustness of validation studies and identifying limitations in current evidence that may impact clinical implementation.

The review followed PRISMA guidelines for systematic reviews where applicable, though the comprehensive nature of the review encompassed multiple study types and methodologies beyond those typically included in systematic reviews of diagnostic test accuracy.

3. Current biomarker landscape

3.1. Traditional single biomarkers

The development of predictive biomarkers for checkpoint inhibitor therapy has evolved significantly since the initial focus on PD-L1 expression. Traditional pathological assessment by expert pathologists forms the foundation of biomarker evaluation, providing standardised, reproducible scoring systems that have been extensively validated across multiple cancer types and clinical settings. A comprehensive analysis of 45 FDA drug approvals from 2011 to 2019 revealed that PD-L1 expression was predictive in only 28.9 % of cases, with the biomarker proving either non-predictive (53.3 %) or not tested (17.8 %) in the majority of approvals (Davis and Patel, 2019). This limited predictive accuracy stems from several factors including intratumoral heterogeneity, variable assay methodologies, and diverse scoring systems across different tumour types. The expertise of pathologists in standardising these assessments and providing quality control remains essential for ensuring reproducible clinical decision-making.

Tumour mutational burden, a measure of the number of mutations in a tumour's DNA has emerged as a complementary biomarker, with FDA approval for pembrolizumab in patients with unresectable or metastatic solid tumours harbouring high TMB (≥10 mutations per megabase) following disease progression (Marcus et al. 2021). Studies across multiple cancer types have demonstrated positive correlations between TMB and response rates, with meta-analyses showing significant associations between high mutation burden and improved clinical outcomes (Fan et al. 2020; Willis et al. 2019). However, notable exceptions exist, particularly in renal cell carcinoma and Merkel cell carcinoma, where responses exceed what TMB alone would predict, highlighting the multifactorial nature of treatment response (Knepper et al. 2019; Labriola et al. 2020).

Microsatellite instability (MSI) status represents another established biomarker, particularly valuable in colorectal cancer where MSI-high tumours demonstrate markedly improved responses to checkpoint inhibitors. MSI indicates whether a tumour's DNA is unstable enough to result in mutations during replication (Vilar and Gruber, 2010). This leads to a biological rationale on the increased neoantigen load in MSI-high tumours, resulting in enhanced immune recognition and response. Clinical studies have consistently demonstrated superior outcomes in MSI-high compared to microsatellite stable tumours across various cancer types (Van Velzen et al., 2020; Pietrantonio et al. 2021). The mechanisms through which PD-L1, TMB, and MSI function as predictive biomarkers for immune checkpoint blockade are illustrated in Fig. 1.

The exploration of metabolic biomarkers represents an expanding frontier in cancer prognostication and treatment response prediction. Glucose transporters, particularly GLUT1 and GLUT3, have emerged as significant prognostic indicators across multiple cancer types. These transporters facilitate the enhanced glucose uptake characteristic of the Warburg effect, whereby cancer cells preferentially utilise glycolysis even in the presence of oxygen (Gökalp 2022a). Overexpression of GLUT1 and GLUT3 correlates with aggressive tumour phenotypes, poor prognosis, and resistance to conventional therapies in various malignancies including lung, breast, and gastrointestinal cancers.

GLUT1, encoded by the SLC2A1 gene, is ubiquitously expressed but shows marked upregulation in hypoxic tumour regions, driven by hypoxia-inducible factor-1α (HIF-1α) signalling. Clinical studies have

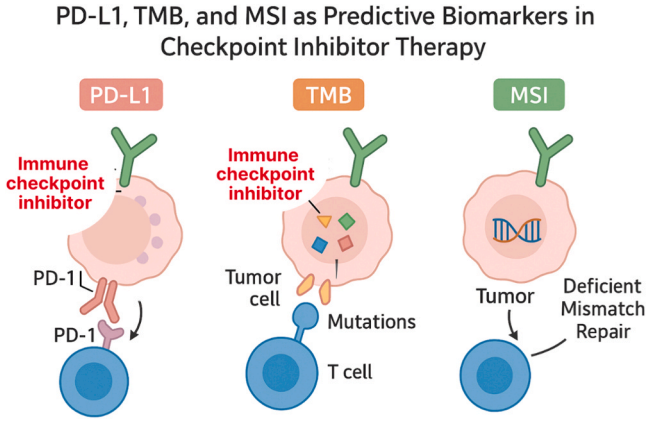

Fig. 1. Schematic illustration of the mechanisms by which PD-L1 expression, tumor mutational burden (TMB), and microsatellite instability (MSI) predict the efficacy of immune checkpoint inhibitors (ICIs).

demonstrated that elevated GLUT1 expression associates with reduced overall survival and increased metastatic potential (Gökalp 2022a). GLUT3, predominantly expressed in neurons under physiological conditions, becomes aberrantly expressed in several cancer types, particularly brain tumours and lung cancers. Its high-affinity glucose transport capacity makes it especially relevant in nutrient-deprived tumour microenvironments.

The prognostic value of glucose transporter expression extends beyond direct metabolic effects. Enhanced glycolytic metabolism in tumours can create an immunosuppressive microenvironment through lactate accumulation and acidification, potentially limiting T-cell effector function and thereby influencing checkpoint inhibitor response. Furthermore, competition for glucose between tumour cells and infiltrating immune cells may contribute to T-cell exhaustion, a key mechanism of immunotherapy resistance. Therapeutic strategies targeting glucose metabolism, including GLUT inhibitors, are under investigation as potential combination approaches with checkpoint inhibitors. The integration of metabolic biomarkers such as GLUT1 and GLUT3 expression into predictive models could enhance patient stratification by identifying those with metabolically active tumours that may require combined metabolic and immune-targeted interventions.

3.2. Emerging biomarker strategies

Recent developments in spatial biomarker analysis have provided new insights into the tumour microenvironment's role in treatment response. A comprehensive review of spatial biomarkers identified several promising approaches, including digital spatial profiling and multiplex immunofluorescence techniques that examine the spatial relationships between immune cells and cancer cells (Rossi and Radisky, 2024; Sheng et al. 2023). These technologies enable assessment of immune cell infiltration patterns, which correlate more closely with treatment response than bulk biomarker measurements.

The integration of artificial intelligence with traditional biomarker assessment has enhanced predictive accuracy. Studies utilizing AI-powered PD-L1 assessment have demonstrated improved interobserver concordance and enhanced predictive value compared to traditional pathologist scoring (Kim et al. 2024; Lee et al. 2024). These systems can process large datasets rapidly whilst maintaining consistency across different samples and institutions.

Circulating biomarkers, including circulating tumour DNA (ctDNA) and blood-based TMB assessment, offer the advantage of non-invasive monitoring and the ability to capture tumour heterogeneity more comprehensively than single tissue samples. Blood-based biomarkers extend beyond ctDNA to include metabolic markers such as circulating glucose levels and expression of glucose transporters on circulating tumour cells, which may provide real-time insights into tumour metabolic activity and treatment response (Gökalp 2022a). Studies have shown correlations between blood-based biomarkers and tissue-based assessments, though standardisation remains a challenge for widespread clinical implementation (Friedman and Postow, 2016; Raez et al. 2018). Additionally, metabolomic profiling of patient plasma samples has revealed distinct metabolic signatures associated with checkpoint inhibitor response, including alterations in glucose, amino acid, and lipid metabolism, suggesting potential for integrative biomarker panels combining genomic, immunological, and metabolic parameters. Table 1 summarises the current landscape of established and emerging biomarkers, their methodologies, regulatory status, and clinical applications.

4. Machine learning and artificial intelligence approaches

4.1. Clinical implementation of AI models

The integration of artificial intelligence into checkpoint inhibitor prediction has advanced from experimental concepts to clinically validated tools. However, critical evaluation reveals that whilst these models show statistical superiority over individual biomarkers, their clinical implementation faces significant challenges related to external validation and real-world performance. The SCORPIO machine learning system represents a significant milestone, utilising routine blood tests and clinical characteristics from 9745 patients across 21 cancer types to predict treatment response (Yoo et al. 2025). Trained on data from Memorial Sloan Kettering Cancer Center, SCORPIO achieved median area under the curve values of 0.763 and 0.759 for predicting overall survival at multiple time points, substantially outperforming traditional TMB assessment. Whilst these results are promising, the model's performance was primarily demonstrated within the Memorial Sloan Kettering system, and broader external validation across diverse healthcare settings with different patient populations and clinical practices remains limited.

The LORIS (Logistic Regression-Based Immunotherapy-Response Score) model has demonstrated comparable success using six readily available clinical and genomic features: patient age, cancer type, treatment history, blood albumin levels, neutrophil-to-lymphocyte ratio, and TMB (Chang et al. 2024). This model's strength lies in its reliance on routinely collected clinical data, making it immediately applicable in most healthcare settings without requiring sophisticated genomic testing infrastructure. However, similar to SCORPIO, comprehensive validation across different healthcare systems and patient populations is still needed to establish generalisability and clinical utility.

4.2. Advanced AI methodologies

Deep learning approaches have shown particular promise in processing complex, high-dimensional datasets. Convolutional neural networks have been successfully applied to histopathological image analysis, enabling automated assessment of tumour-infiltrating lymphocytes and spatial immune cell distributions (Millward et al. 2025; Choi et al. 2023). These AI-based approaches build upon decades of traditional pathological expertise in recognising immune infiltration patterns, requiring extensive validation by expert pathologists to ensure clinical accuracy and reproducibility. These systems can identify subtle patterns in tissue architecture that may not be apparent to human observers, potentially uncovering novel predictive features. However, the interpretability of these patterns and their biological relevance requires ongoing validation through traditional pathological assessment.

Multi-modal AI approaches that integrate diverse data types have demonstrated superior performance compared to single-modality models. Studies combining genomic data, imaging features, and clinical variables have achieved area under the curve values exceeding 0.8 in several cancer types (Li et al. 2025; Goyal et al. 2024). The challenge lies in standardising data collection and processing across different institutions to ensure model generalisability, a critical limitation that has hindered the clinical implementation of many promising AI approaches.

Natural language processing techniques have been applied to electronic health records to extract relevant clinical information that might influence treatment response (Clay et al. 2025). These approaches can identify patterns in clinical notes, laboratory results, and treatment histories that correlate with outcomes, providing additional predictive features for machine learning models. However, the variability in clinical documentation practices across institutions presents significant challenges for model generalisability.

4.3. Interpretability and clinical adoption

A critical challenge in AI implementation is the interpretability of complex algorithms. Healthcare providers require understanding of how models reach their predictions to make informed treatment decisions. Explainable AI techniques, including SHAP (Shapley Additive Explanations) values and attention mechanisms, are being developed to provide insights into model decision-making processes (Sadeghi et al. 2024;

Table 1
Current biomarkers and their clinical applications.

Biomarker	Methodology	FDA Approval Status	Cancer Types	Predictive Value	Limitations	Standardisation Issues
PD-L1 Expression (Davis and Patel, 2019)	Immunohistochemistry with expert pathologist scoring	Approved for multiple indications	NSCLC, bladder, gastric, cervical, TNBC	Limited predictive accuracy (28.9 % of approvals show predictive value)	Intratumoral heterogeneity, assay variability, different scoring systems	Multiple companion diagnostics with varying cut-offs across tumour types
Tumour Mutational Burden (Stenzinger et al., 2019)	NGS panels	Approved for solid tumours	Pan-cancer with $\geq$ 10 mutations/Mb threshold	Statistically significant but clinically modest correlation with response	Cost, long turnaround time, tissue requirements	Ongoing efforts to standardise assay platforms, bioinformatics pipelines, and reporting methods
Microsatellite Instability (Li et al., 2020; Marques et al., 2022)	PCR/NGS	Approved entity-agnostically for MSI-H/dMMR solid tumours	Pan-cancer MSI-H/dMMR tumours (not limited to specific populations)	High predictive value in MSI-H tumours (>80 % response rates in some studies)	Low prevalence in most solid tumours (2–4 % overall)	Well-established methodologies with good inter-laboratory concordance
Blood-based TMB (Sivapalan et al., 2023; Boukovala et al., 2024)	Liquid biopsy NGS	Investigational	Multiple	Emerging correlative evidence with tissue TMB	Concordance with tissue TMB varies (60–80 %), lower sensitivity for low mutation burden	Significant standardisation needs: pre-analytical variables, platform differences, cut-off determination
Circulating Tumour DNA (Dang and Park, 2022; Guigal-Stephan et al., 2025)	NGS liquid biopsy	Investigational	Multiple	Promising early data for monitoring treatment response	Technical challenges in detection sensitivity, tumour fraction variability	No standardisation issues listed as methodology still in development phase
Immune Cell Infiltration (Barua et al., 2018; Page et al., 2023)	Traditional pathologist assessment and spatial analysis platforms	Research use	Multiple	Strong association with response (TILs show consistent prognostic value)	Requires specialised platforms for quantitative analysis beyond routine pathology	Working groups developing standardised scoring systems for routine implementation

Mathew et al. 2025). The development of interpretable AI systems is essential for clinical adoption, as physicians need to understand the biological and clinical rationale underlying predictions to maintain confidence in AI-assisted decision-making.

The development of user-friendly interfaces and clinical decision support systems is essential for widespread adoption. Successful implementations require integration with existing healthcare information systems and workflows, ensuring that AI tools enhance rather than complicate clinical decision-making processes. Current evidence suggests that whilst AI models show statistical promise, the practical challenges of clinical implementation, including system integration, user training, real-world applicability benefits and ongoing model maintenance, represent significant barriers to widespread adoption (Macheka et al. 2024).

Table 2 provides a comprehensive overview of current artificial intelligence applications in checkpoint inhibitor prediction, detailing their methodologies, performance metrics, and implementation status, with clarified definitions of validation stages and technical approaches.

The integration of artificial intelligence into checkpoint inhibitor prediction has advanced from experimental concepts to clinically validated tools, as illustrated in Fig. 2. This includes clinical models such as SCORPIO and LORIS, as well as advanced deep learning and multimodal approaches, each addressing unique challenges in prediction, interpretability, and clinical integration (Fig. 2).

5. Mathematical and mechanistic models

5.1. Systems biology approaches

Mathematical models provide unique insights into the dynamic interactions between tumours and the immune system, offering mechanistic understanding that complements empirical prediction models. These approaches represent a distinct analytical methodology that focuses on underlying biological mechanisms rather than pattern recognition from large datasets. A translational mathematical model has been developed that captures checkpoint inhibitor efficacy through three key parameters: tumour growth rate, immune infiltration, and immunotherapy-mediated amplification of anti-tumour response (Butner et al. 2021). This model demonstrated 81.4 % accuracy in classifying treatment response using only tumour volume measurements within two months of treatment initiation across multiple solid tumour types. However, this validation was conducted retrospectively on a relatively small cohort (189 patients), and prospective validation across diverse clinical settings is needed to establish clinical utility.

The mathematical framework enables simulation of various treatment scenarios and can guide optimal dosing strategies and combination therapies. These models can incorporate pharmacokinetic parameters, immune cell dynamics, and tumour heterogeneity to provide personalised treatment predictions. The mechanistic basis of these models makes them particularly valuable for understanding treatment resistance mechanisms and identifying potential combination targets, though their clinical implementation requires validation of the underlying biological assumptions across different patient populations.

5.2. Computational immune monitoring

Advanced computational approaches have been developed to monitor immune system dynamics during treatment. These models integrate multiple data streams including immune cell populations, cytokine levels, and tumour markers to provide real-time assessment of treatment response (Qin et al. 2024). Such approaches can identify early indicators of treatment failure or immune-related adverse events, enabling proactive treatment modifications, though the complexity of these models presents challenges for routine clinical implementation.

The integration of mathematical models with clinical data has enabled the development of predictive frameworks that account for individual patient characteristics and tumour biology. These personalised models can simulate treatment outcomes under different scenarios, supporting clinical decision-making in complex cases where traditional biomarkers provide ambiguous results. However, the clinical validation of these mechanistic approaches remains limited compared to empirical biomarker-based models.

6. Integrated multi-modal approaches

6.1. Comprehensive biomarker frameworks

The recognition that single biomarkers have limited predictive accuracy has driven development of integrated approaches that combine multiple predictive modalities. It is crucial to distinguish between the underlying biological features (biomarkers such as PD-L1, TMB, immune infiltration patterns, and metabolic markers) and the computational

Table 2
Artificial Intelligence Applications in Checkpoint Inhibitor Prediction.

AI Approach	Input Data Types	Technical Method	Clinical Application	Performance Metrics	Implementation Status	External Validation
Deep Learning for Histopathology (Shamai et al., 2022)	H&E slides, IHC images	Convolutional neural networks	Automated TIL quantification, computer-assisted PD-L1 scoring	AUC 0.91–0.93	Proof-of-concept validation: demonstrated technical feasibility with promising performance in controlled research settings	Limited to original development institutions
Multi-modal Integration (Goyal et al., 2024; Li et al., 2025)	Genomics, imaging, clinical	Ensemble methods	Comprehensive treatment selection incorporating multiple data types	AUC > 0.85	Proof-of-concept validation: shows technical feasibility but requires extensive multi-institutional validation	External validation limited; most studies single-institution
Natural Language Processing (Clay et al., 2025)	Electronic health records	Transformer models	Automated clinical feature extraction from unstructured data	Accuracy 75–85 %	Early development: preliminary algorithms under development and testing	Not yet undertaken
Radiomics Analysis (Jiang et al., 2020)	CT, PET scans	Various machine learning algorithms	Non-invasive imaging-based response prediction	AUC 0.75–0.91	Research applications: applied in research settings with promising initial results	Limited cross-institutional validation
SCORPIO (Yoo et al., 2025)	Clinical & laboratory blood test	Machine learning ensemble	Routine clinical prediction using readily available data	AUC > 0.75	Clinical validation: demonstrated in large clinical cohort with robust methodology	Single-institution validation at Memorial Sloan Kettering
Explainable AI (Sadeghi et al., 2024; Mathew et al., 2025)	Multiple data types	SHAP, attention mechanisms	Interpretable clinical decision support providing rationale for predictions	Interpretability metrics rather than predictive accuracy	Active research: ongoing development of interpretability methods	Methodology development phase

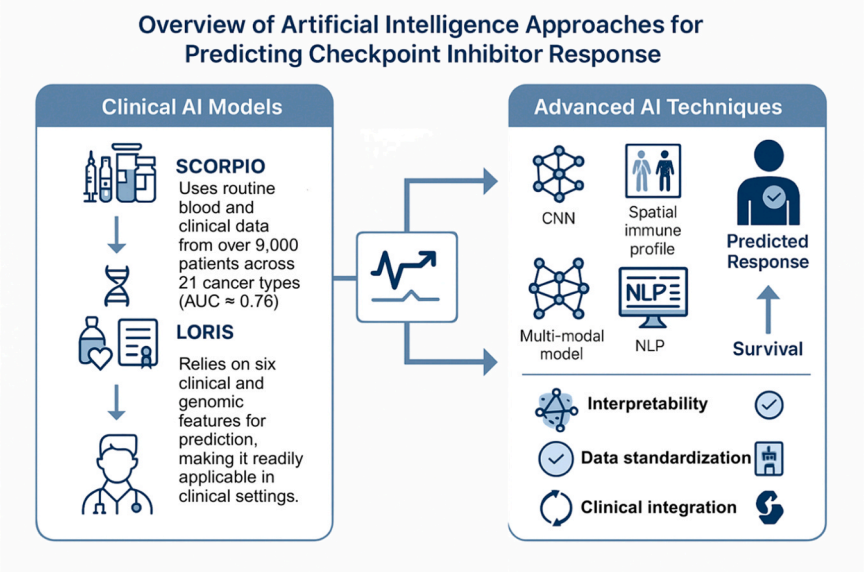

Fig. 2. Artificial intelligence approaches for predicting response to checkpoint inhibitor therapy. On the left, clinical AI models such as SCORPIO utilize routine blood and clinical data from over 9000 patients across 21 cancer types (AUC ≈ 0.76), while LORIS relies on six clinical and genomic features for prediction, making it readily applicable in clinical settings. On the right, advanced AI techniques include convolutional neural networks (CNN) for histopathological image analysis, spatial immune profiling, multi-modal models that integrate diverse data types, and natural language processing (NLP) applied to electronic health records.

methods used to integrate these features into predictive models. Recent studies have demonstrated that combining PD-L1 expression with TMB and immune cell infiltration markers can achieve area under the curve values exceeding 0.85 in certain cancer types (So et al. 2023; Yamaguchi et al. 2024). A performance comparison is detailed in Table 3, though direct comparisons are challenging due to differences in study populations, methodologies, and validation approaches. These multi-dimensional approaches provide more comprehensive assessment of the tumour immune microenvironment.

Beyond traditional genomic and immunological biomarkers, emerging multi-modal frameworks increasingly incorporate metabolic parameters. The integration of glucose transporter expression (GLUT1

Table 3
Performance comparison of major predictive models for checkpoint inhibitor response (Revised).

Model/ Approach	Dataset Size	Cancer Types	Primary Endpoint	AUC/Accuracy	Key Features	Validation Status	External Validation	Study Population
SCORPIO (Yoo et al., 2025)	9745 patients	21 cancer types	Overall survival	0.76 (median AUC)	Complete blood count, comprehensive metabolic panel, age, cancer type	Multi-institution development and internal validation	Limited to Memorial Sloan Kettering Cancer Center system	Predominantly US population, single healthcare system
LORIS (Chang et al., 2024)	Multiple cohorts (>3000 patients)	Pan-cancer	Treatment response	81.4 % accuracy	Age, cancer type, prior therapy, albumin, neutrophil-to-lymphocyte ratio, TMB	External validation across multiple published cohorts	Validated across different study populations and institutions	International cohorts from multiple published studies
Spatial Biomarker Model (Song et al., 2023)	18 patients	NSCLC	Treatment response	0.84	18-protein spatial signature analysis in tumour microenvironment	Proof-of-concept study	No external validation	Single-institution, small pilot study
PD-L1 Expression (Huang and Teng, 2020)	Meta-analysis of > 10,000 patients	Multiple	Treatment response	0.65–0.78 (varies by cancer type)	Immunohistochemistry with standardised pathologist scoring	Regulatory approval with extensive validation	Validated across multiple international clinical trials	Global multi-institutional clinical trial populations
TMB Assessment (Litchfield et al., 2021)	Multiple cohorts	Pan-cancer	Treatment response	0.60–0.68	Whole exome or targeted panel sequencing	Regulatory approval based on clinical trial data	Validated across multiple clinical trials and real-world studies	Diverse international populations
Mathematical Model (Butner et al., 2021)	189 patients	Multiple solid tumours	Clinical response	81.4 % accuracy	Tumour volume kinetics from serial imaging	Retrospective single-institution validation	No prospective or external validation	Single-institution retrospective cohort

and GLUT3) with PD-L1 status and TMB has shown promise in refining patient stratification. For instance, tumours exhibiting high glycolytic activity (indicated by elevated GLUT1/GLUT3 expression) combined with low PD-L1 expression may represent a distinct subgroup requiring alternative therapeutic strategies or combination approaches targeting both immune checkpoints and metabolic pathways (Gökalp 2022a).

Furthermore, computational modelling approaches integrating metabolic flux analysis with immune cell infiltration patterns have revealed complex interactions between tumour metabolism and immune evasion mechanisms. These multi-dimensional models suggest that metabolic reprogramming not only supports tumour growth but actively shapes the immune microenvironment, creating zones of immune privilege through nutrient depletion and accumulation of immunosuppressive metabolites.

Spatial biomarker strategies have shown particular promise, with studies demonstrating that stromal signature scores comprising 18 protein targets achieved superior predictive power (AUC 0.84) compared to bulk PD-L1 expression (AUC 0.78) and TMB (AUC 0.53) alone (Song et al. 2023). The integration of spatial relationships between different cell types provides insights into the functional organisation of the tumour microenvironment that correlate with treatment response, building upon traditional pathological assessment of immune infiltration patterns that have long been recognised as prognostically important. Recent advances in immunotherapy prediction leverage integrated multi-modal biomarker frameworks and dynamic monitoring strategies, as illustrated in Fig. 3. These approaches combine established and emerging biomarkers to improve predictive accuracy and enable early, adaptive assessment of treatment response (Fig. 3).

6.2. *Dynamic monitoring strategies*

The development of dynamic monitoring approaches extends beyond traditional imaging and blood-based biomarkers to include metabolic monitoring. Serial assessment of tumour glucose metabolism through imaging modalities such as 18F-FDG PET-CT, combined with measurement of circulating metabolites and expression of glucose transporters, enables real-time evaluation of metabolic response to checkpoint inhibitor therapy. Changes in tumour glucose uptake patterns may predict treatment response earlier than radiographic changes, as metabolic alterations often precede morphological changes (Gökalp 2022a).

Integration of metabolic monitoring with immune biomarkers provides a comprehensive view of treatment-induced changes. For example, successful checkpoint inhibitor therapy may be associated with decreased tumour glucose uptake (reflecting reduced metabolic activity) concurrent with increased infiltration of activated T-cells. Conversely, persistent high metabolic activity despite immune cell infiltration may indicate ongoing immune evasion and potential treatment failure. Such dynamic, multi-parameter monitoring strategies could enable early identification of non-responders and guide adaptive treatment modifications.

The development of dynamic monitoring approaches that track changes in biomarker expression during treatment has shown promise for early prediction of treatment response. Serial assessment of circulating biomarkers, including ctDNA and immune cell populations, can identify patients experiencing treatment benefit before radiographic response becomes apparent (Thompson et al. 2021). Liquid biopsy offers the advantage of capturing tumour heterogeneity more comprehensively than single tissue samples (Ma et al. 2023). Studies have demonstrated correlations between circulating biomarkers and treatment outcomes, though standardisation of collection and analysis methods remains a challenge for widespread clinical implementation. However, the clinical utility of dynamic monitoring approaches requires prospective validation to demonstrate that early biomarker changes translate into actionable clinical decisions that improve patient outcomes.

Table 3 presents a comprehensive comparison of major predictive models, highlighting their performance characteristics, validation status, and key features across different cancer types and patient populations, with enhanced detail on study characteristics and validation methodology.

7. Clinical validation and implementation challenges

7.1. *Regulatory and validation requirements*

The translation of predictive models from research settings to clinical practice faces significant regulatory and validation hurdles. External validation represents one of the most critical challenges in predictive

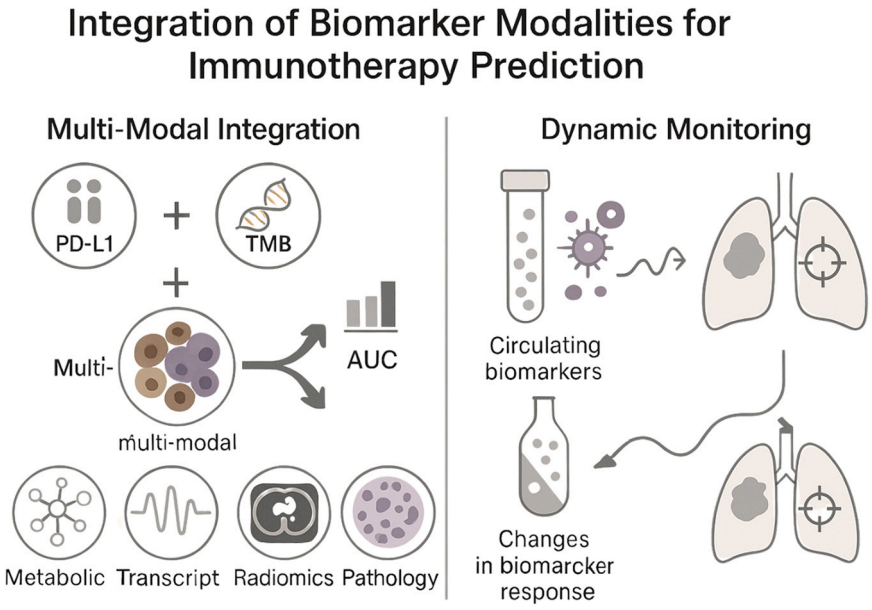

Fig. 3. Integration of biomarker modalities for immunotherapy prediction. Multi-modal integration combines biomarkers such as PD-L1, TMB, and additional molecular and imaging features to improve predictive accuracy (AUC). Dynamic monitoring uses serial circulating biomarkers to track treatment response in real time, enabling earlier and more comprehensive assessment of immunotherapy outcomes.

model development, with many promising approaches failing to maintain performance when applied outside their original development settings. The FDA has established guidelines for the validation of biomarker-based diagnostic tests, requiring demonstration of analytical validity, clinical validity, and clinical utility across diverse patient populations. Several approvals occurred after extensive validation from clinical trials. Currently, there are three FDA-approved predictive biomarkers namely PD-L1, MSI, and TMB routinely used for patient selection for immune checkpoint inhibitor response in clinical practice (Wang et al. 2021). Notably, few predictive models beyond these established biomarkers have achieved the level of validation required for regulatory approval, limiting their clinical adoption.

External validation studies are essential for demonstrating model generalisability across different institutions and patient populations. A critical example is provided by a validation study of a melanoma prediction model: whilst the original single-institution study reported high accuracy, subsequent validation in a national cohort of advanced melanoma patients treated with anti-PD-1 monotherapy could not reproduce the initial performance, highlighting fundamental generalisability concerns (van der Kooij et al., 2023). This pattern of reduced performance in external validation has been observed across multiple AI and machine learning approaches, representing a significant barrier to clinical implementation.

Many promising models have shown excellent performance in single-institution studies but have failed to maintain accuracy when applied to external datasets, highlighting the importance of robust validation frameworks. This phenomenon, known as the "validation gap," represents one of the most significant challenges facing the field and explains why few predictive models have achieved widespread clinical adoption despite promising initial results.

7.2. Data standardisation and quality

The integration of diverse data types requires standardisation of collection methods, processing protocols, and quality control measures. Variations in sample collection, storage conditions, and analytical techniques can significantly impact model performance. The development of standard operating procedures and quality metrics is essential for ensuring reproducible results across different healthcare settings, yet many promising AI models have been developed without adequate attention to these standardisation requirements.

International collaborative efforts are addressing standardisation challenges through initiatives such as the Global Alliance for Genomics and Health (GA4GH) and the International Cancer Genome Consortium (ICGC). These organisations are developing frameworks for data sharing, standardisation, and quality control that will facilitate the development and validation of predictive models (Global Alliance for Genomics and Health, 2025). However, the implementation of these standards across diverse healthcare systems remains a significant challenge.

7.3. Healthcare system integration and clinical implementability

The successful implementation of predictive models requires integration with existing healthcare information systems and clinical workflows, representing a major practical challenge that is often underestimated in research settings. Models must be accessible to clinicians at the point of care, with user-friendly interfaces that provide clear, actionable recommendations. The development of clinical decision support systems that integrate predictive models with electronic health records is essential for widespread adoption. However, the practical challenges of system integration, including software compatibility, data security requirements, and workflow disruption, have limited the clinical implementation of many technically successful models.

Training and education programmes are necessary to ensure that healthcare providers understand how to interpret and apply predictive model results. The complexity of many models requires ongoing support and education to ensure appropriate clinical use. Additionally, the cost-effectiveness of implementing sophisticated predictive models, including the required infrastructure, personnel training, and ongoing maintenance, represents a significant consideration for healthcare systems with limited resources.

The ease of clinical implementation varies significantly across different predictive approaches. Traditional biomarkers such as PD-L1 and MSI leverage existing pathology infrastructure and expertise, facilitating their clinical adoption. In contrast, sophisticated AI models requiring specialised computational resources and expertise face greater implementation barriers, regardless of their technical performance.

8. Future directions and emerging technologies

8.1. Real-time adaptive modelling

The development of real-time adaptive models that continuously update predictions based on treatment response represents an exciting frontier in personalised medicine. These systems would integrate streaming data from various sources including wearable devices, regular blood tests, and imaging studies to provide dynamic predictions of treatment response and toxicity risk. However, the clinical validation and regulatory approval of such dynamic systems present unprecedented challenges, as traditional clinical trial designs are not well-suited to evaluating continuously adaptive algorithms.

The integration of pharmacokinetic modelling with predictive algorithms could enable personalised dosing strategies that optimise efficacy whilst minimising toxicity. Such approaches could account for individual patient characteristics, drug metabolism, and treatment response patterns to provide optimal treatment regimens. However, the clinical implementation of personalised dosing based on predictive models requires extensive safety validation and regulatory oversight.

8.2. Novel data sources and technologies

Emerging technologies including single-cell sequencing, spatial transcriptomics, and advanced imaging techniques are providing unprecedented insights into tumour biology and immune system dynamics. Metabolomic and lipidomic profiling technologies have advanced significantly, enabling comprehensive characterisation of tumour metabolic states and their relationship to immune checkpoint expression and response. High-resolution mass spectrometry-based metabolomics can simultaneously quantify hundreds of metabolites, providing detailed metabolic signatures that may predict checkpoint inhibitor response (Gökalp 2022a).

Spatial metabolomics, combining mass spectrometry imaging with immunohistochemistry, enables visualisation of metabolite distributions within the tumour microenvironment and their spatial relationship to immune cell infiltrates. This technology has revealed metabolic zonation within tumours, with glucose-depleted regions showing reduced T-cell infiltration and function, highlighting the importance of metabolic-immune interactions in determining treatment response.

The integration of these technologies with predictive models has the potential to significantly enhance prediction accuracy and provide deeper mechanistic understanding, though the cost and complexity of these approaches may limit their widespread clinical implementation.

Wearable devices and mobile health technologies offer opportunities for continuous monitoring of patient status during treatment. Beyond physiological parameters, emerging biosensor technologies enable non-invasive monitoring of circulating metabolites and immune markers, providing real-time feedback on treatment-induced metabolic and immunological changes. These devices can track physiological parameters, activity levels, and patient-reported outcomes that may correlate with treatment response and toxicity risk. However, the clinical

validation of digital biomarkers and their integration into treatment decision-making algorithms remains in early development stages.

8.3. Natural bioactive compounds and complementary therapeutic strategies

An emerging area of investigation involves the potential of natural bioactive compounds to enhance checkpoint inhibitor efficacy through complementary mechanisms. These compounds, derived from traditional medicinal plants used across diverse cultures for centuries, have demonstrated multifaceted anticancer properties that may synergise with immunotherapy.

Thymoquinone, the principal bioactive compound in black cumin (*Nigella sativa*), has exhibited potent anticancer effects across multiple cancer types including lung, pancreatic, cervical, and breast cancers (Gökalp, 2020; Gökalp, 2021; Gökalp, 2025). Its mechanisms of action include inhibition of histone deacetylase 2 (HDAC2), a key epigenetic regulator overexpressed in many cancers, modulation of cell cycle progression through effects on cyclin D1 and p53 expression, and induction of apoptosis through mitochondrial pathways. Significantly, thymoquinone has also demonstrated immunomodulatory properties, including enhancement of natural killer cell activity and modulation of cytokine production, which could potentially complement checkpoint inhibitor mechanisms (Gökalp, 2025; Alhmied et al. 2021; Randhawa and Alghamdi, 2011).

Related compounds including thymol and carvacrol, monoterpenoid phenols found in various aromatic plants, have shown similar anticancer properties with distinct molecular targets. These compounds inhibit multiple cancer-associated enzymes including carbonic anhydrase isoforms, acetylcholinesterase, and α-glycosidase, while also demonstrating direct cytotoxic effects on cancer cells (Gökalp, 2020; Gökalp, 2021; Güzel et al. 2019). The inhibition of carbonic anhydrase, particularly the CA IX and CA XII isoforms upregulated in hypoxic tumours, represents a particularly relevant mechanism, as these enzymes contribute to the acidic tumour microenvironment that impairs immune cell function.

Organosulfur compounds from garlic (*Allium sativum*), including diallyl sulfide (DAS), diallyl disulfide (DADS), and diallyl trisulfide (DATS), have demonstrated significant anticancer effects through inhibition of cancer cell proliferation, induction of apoptosis, and suppression of angiogenesis (Gökalp, 2020; Gökalp, 2021). These compounds modulate multiple signalling pathways relevant to cancer progression and immune evasion, including NF-κB, MAPK, and PI3K/AKT pathways. Their ability to reduce oxidative stress while selectively inducing apoptosis in cancer cells suggests potential for combination with immunotherapy to enhance antitumour immune responses.

Cucurbitacins, triterpenoid compounds found in plants of the Cucurbitaceae family and marigold (*Tagetes* species), have shown potent cytotoxic effects against various cancer cell lines, including cervical, ovarian, and lung cancers (Gökalp, 2021; Gökalp, 2023). Cucurbitacin I, in particular, inhibits the JAK2/STAT3 signalling pathway, which is frequently hyperactivated in cancers and contributes to immune evasion through upregulation of immunosuppressive factors. By inhibiting this pathway, cucurbitacins may potentially enhance the tumour microenvironment's permissiveness to immune-mediated killing. Additionally, α-terthienyl and quercetagetin from marigold have demonstrated inhibitory effects on nematode and insect receptors through molecular mechanisms that may extend to cancer cell signalling pathways (Gökalp, 2023).

The integration of these natural compounds into checkpoint inhibitor strategies could occur through several approaches:

1. Metabolic modulation: Natural compounds targeting tumour metabolism (such as HDAC inhibitors like thymoquinone or carbonic anhydrase inhibitors like thymol and carvacrol) could reshape the tumour microenvironment to favour immune cell infiltration and function, potentially enhancing checkpoint inhibitor response.

2. Immune enhancement: Compounds with immunomodulatory properties could directly augment antitumour immune responses, complementing the mechanism of checkpoint inhibitors. For instance, thymoquinone's effects on cytokine production and natural killer cell activity may synergise with PD-1/PD-L1 blockade.

3. Epigenetic reprogramming: HDAC inhibitors from natural sources could reverse epigenetic silencing of immune-related genes in both tumour and immune cells, potentially overcoming intrinsic resistance mechanisms to checkpoint inhibitors.

4. Multi-target effects: Many natural compounds simultaneously affect multiple pathways relevant to cancer progression and immune evasion, potentially addressing the multifactorial nature of checkpoint inhibitor resistance more effectively than single-target synthetic inhibitors.

However, clinical translation of these findings faces substantial challenges. The pharmacokinetic properties of many natural compounds limit their bioavailability and tissue distribution. Chemical modification or nanoparticle-based delivery systems may be required to achieve therapeutic concentrations at tumour sites. Additionally, potential drug-drug interactions between natural compounds and checkpoint inhibitors require careful investigation, as some natural compounds may affect the metabolism or efficacy of immunotherapy agents through cytochrome P450 enzyme modulation or other mechanisms.

Rigorous preclinical studies using appropriate animal models and patient-derived xenografts are essential to establish proof-of-concept for combination strategies. Subsequent clinical trials should employ adaptive designs to identify optimal dosing, timing, and patient populations most likely to benefit. Biomarker-driven trial designs incorporating metabolic, immunological, and pharmacological parameters could accelerate identification of responsive patient subgroups and inform personalised combination approaches.

8.4. Combination therapy prediction

The increasing use of combination immunotherapy regimens presents new challenges and opportunities for predictive modelling that extend beyond the scope of this review focused on monotherapy approaches. Models must account for the complex interactions between different therapeutic agents and their combined effects on the immune system and tumour microenvironment. The development of predictive models for novel immunotherapy combinations, including checkpoint inhibitors with targeted therapies, chemotherapy, and radiation therapy, will require sophisticated approaches that account for multifactorial interactions and temporal dynamics. This represents a critical area for future research as combination approaches become increasingly standard in clinical practice.

Future research priorities should include: prospective validation studies designed specifically to test model generalisability across diverse healthcare settings; development of standardised protocols for model validation and implementation; creation of interpretable AI systems that provide clinically actionable insights with clear biological rationale; and establishment of regulatory frameworks for evaluating and approving dynamic, adaptive prediction systems.

9. Conclusions

The field of checkpoint inhibitor prediction has evolved rapidly from simple biomarker assessments to sophisticated multi-modal approaches incorporating artificial intelligence, metabolic parameters, and systems biology. However, critical analysis reveals a significant gap between promising research findings and clinical implementation, primarily due to limited external validation and practical implementation challenges. Whilst significant progress has been made, several challenges remain before these tools can achieve widespread clinical adoption. The development of robust validation frameworks, standardisation of data collection and analysis methods, and integration with healthcare

systems are essential for translating research advances into clinical practice, yet these fundamental requirements have been inadequately addressed in much of the current literature.

Traditional pathological assessment by expert pathologists continues to provide the foundation for clinical decision-making and serves as the essential validation standard for emerging AI-based approaches. The integration of metabolic biomarkers, particularly glucose transporter expression, into existing prediction frameworks represents a promising avenue for enhancing patient stratification. Furthermore, the investigation of natural bioactive compounds as complementary therapeutic agents offers potential strategies to overcome checkpoint inhibitor resistance and improve treatment outcomes through metabolic modulation, immune enhancement, and multi-target effects. The expertise and standardised protocols developed through decades of pathological practice remain irreplaceable for ensuring accurate, reproducible biomarker assessment.

The most promising approaches appear to be integrated multi-modal frameworks that combine traditional biomarkers with artificial intelligence algorithms, metabolic profiling, and dynamic monitoring strategies. However, critical evaluation reveals that while these approaches show statistical superiority in development cohorts, their clinical utility advantage over established biomarkers requires further substantiation through rigorous external validation studies. These approaches offer the potential to significantly improve patient selection for checkpoint inhibitor therapy, optimise treatment regimens, and minimise adverse events. However, the translation of this potential into clinical reality requires addressing fundamental challenges of external validation, standardisation, and practical implementation that have hindered the field's progress.

Future research priorities should encompass not only prospective validation studies and interpretable AI systems but also investigation of metabolic-immune interactions, development of combination strategies incorporating natural bioactive compounds, and integration of metabolic monitoring into dynamic response assessment protocols. The convergence of traditional biomarkers, advanced computational approaches, metabolic profiling, and novel therapeutic strategies holds promise for truly personalised immunotherapy, but realisation of this potential requires addressing fundamental challenges of validation, standardisation, and clinical implementation that have hindered progress in the field.

Future research should prioritise prospective validation studies designed to test model generalisability across diverse healthcare settings and patient populations, development of explainable AI systems that provide clinically actionable insights with clear biological rationale, and creation of adaptive models that continuously update predictions based on treatment response while maintaining clinical interpretability. The ultimate goal is the development of comprehensive decision support systems that guide personalised immunotherapy selection and management, maximising therapeutic benefit whilst minimising harm. Critical success factors include ensuring external validation across diverse populations, maintaining clinical interpretability, and addressing practical implementation challenges including cost-effectiveness and integration with existing healthcare workflows.

The successful implementation of these predictive tools will require collaboration between researchers, clinicians, regulatory agencies, and healthcare systems, with particular attention to addressing the validation deficits and implementation challenges that have limited the clinical translation of promising research findings. Only through coordinated efforts that prioritise rigorous validation and practical implementability can the promise of precision immunotherapy be realised, ensuring that the right patients receive the right treatments at the right time.

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

References

Alhmied, F., Alammar, A., Alsultan, B., Alshehri, M., Pottoo, F.H., 2021. Molecular mechanisms of thymoquinone as anticancer agent. Comb. Chem. High. Throughput Screen. 24 (10), 1644–1653. https://doi.org/10.2174/1386207323999201027225305.

Barua, S., Fang, P., Sharma, A., Fujimoto, J., Wistuba, I., Rao, A.U., Lin, S.H., 2018. Spatial interaction of tumor cells and regulatory t cells correlates with survival in non-small cell lung cancer. Lung Cancer 117, 73–79. https://doi.org/10.1016/j.lungcan.2018.01.022.

Boukovala, M., Westphalen, C.B., Probst, V., 2024. Liquid biopsy into the clinics: current evidence. J. Liq. Biopsy 4, 100146. https://doi.org/10.1016/j.jlb.2024.100146.

Butner, J.D., Martin, G.V., Wang, Z., Corradetti, B., Ferrari, M., Esnaola, N., Chung, C., Hong, D.S., Welsh, J.W., Hasegawa, N., Mittendorf, E.A., 2021. Early prediction of clinical response to checkpoint inhibitor therapy in human solid tumors through mathematical modeling. Elife 10, e70130. https://doi.org/10.7554/eLife.70130.

Chang, T.G., Cao, Y., Sfreddo, H.J., Dhruba, S.R., Lee, S.H., Valero, C., Yoo, S.K., Chowell, D., Morris, L.G., Ruppin, E., 2024. LORIS robustly predicts patient outcomes with immune checkpoint blockade therapy using common clinical, pathologic and genomic features. Nat. Cancer 5 (8), 1158–1175. https://doi.org/10.1038/s43018-024-00772-7.

Choi, S., Cho, S.I., Jung, W., Lee, T., Choi, S.J., Song, S., Park, G., Park, S., Ma, M., Pereira, S., Yoo, D., 2023. Deep learning model improves tumor-infiltrating lymphocyte evaluation and therapeutic response prediction in breast cancer. NPJ Breast Cancer 9 (1), 71. https://doi.org/10.1038/s41523-023-00577-4.

Clay, B., Bergman, H.I., Salim, S., Pergola, G., Shalhoub, J., Davies, A.H., 2025. Natural language processing techniques applied to the electronic health record in clinical research and practice-an introduction to methodologies. Comput. Biol. Med. 188, 109808. https://doi.org/10.1016/j.compbiomed.2025.109808.

Dang, D.K., Park, B.H., 2022. Circulating tumor DNA: current challenges for clinical utility. J. Clin. Investig. 132 (12). https://doi.org/10.1172/jci154941.

Davis, A.A., Patel, V.G., 2019. The role of PD-L1 expression as a predictive biomarker: an analysis of all US food and drug administration (FDA) approvals of immune checkpoint inhibitors. J. Immunother. Cancer 7 (1), 278. https://doi.org/10.1186/s40425-019-0768-9.

Faizullabhoy, M., Wani, G., 2024. Immune checkpoint inhibitors market - by type (PD-1, PD-L1, CTLA-4), application (lung cancer, breast cancer, bladder cancer, melanoma, cervical cancer), end-use (hospitals & clinics, cancer centers), global forecast 2024–2032. Global Market Insights Inc. ⟨https://www.gminsights.com/industry-analysis/immune-checkpoint-inhibitors-market⟩.

Fan, S., Gao, X., Qin, Q., Li, H., Yuan, Z., Zhao, S., 2020. Association between tumor mutation burden and immune infiltration in ovarian cancer. Int. Immunopharmacol. 89, 107126. https://doi.org/10.1016/j.intimp.2020.107126.

Flynn, J.P. and Gerriets, V., 2019. Pembrolizumab. In: StatPearls. Treasure Island (FL): StatPearls Publishing. Available at: ⟨https://www.ncbi.nlm.nih.gov/books/NBK546616/⟩ [Accessed 2 April 2025].

Friedman, C.F., Postow, M.A., 2016. Emerging tissue and blood-based biomarkers that May predict response to immune checkpoint inhibition. Curr. Oncol. Rep. 18 (4), 21. https://doi.org/10.1007/s11912-016-0509-x.

Global Alliance for Genomics and Health, 2025. Framework for responsible sharing of genomic and health-related data. Available at: ⟨https://www.ga4gh.org/framework/⟩ [Accessed 9 June 2025].

Gökalp, F., 2020. The inhibition effect of natural food supplement active ingredients on TP63 carcinoma cell. Med. Oncol. 37, 120. https://doi.org/10.1007/s12032-020-01446-x.

Gökalp, F., 2021. The effective natural compounds for inhibiting cervical cancer. Med. Oncol. 38, 12. https://doi.org/10.1007/s12032-021-01456-3.

Gökalp, F., 2022a. An investigation into the usage of monosaccharides with GLUT1 and GLUT3 as prognostic indicators for cancer. Nutr. Cancer 74 (2), 515–519. https://doi.org/10.1080/01635581.2021.1895233.

Gökalp, F., 2022b. Therapeutic effect of some natural active compounds for breast cancer. Med. Oncol. 39, 115. https://doi.org/10.1007/s12032-022-01704-0.

Gökalp, F., 2023. A study on natural control against nematodes and whiteflies with marigold, known as an antagonist plant. J. Chem. Ecol. 49, 230–234. https://doi.org/10.1007/s10886-023-01421-6.

Gökalp, F., 2025. An investigation into the usage of black cumin derivatives against cancer and COVID-19 as the nature Medicine. J. Biomol. Struct. Dyn. 43 (9), 4362–4369. https://doi.org/10.1080/07391102.2024.2302942.

Goyal, M., Marotti, J.D., Workman, A.A., Tooker, G.M., Ramin, S.K., Kuhn, E.P., Chamberlin, M.D., diFlorio-Alexander, R.M., Hassanpour, S., 2024. A multi-model approach integrating whole-slide imaging and clinicopathologic features to predict breast cancer recurrence risk. NPJ Breast Cancer 10 (1), 93. https://doi.org/10.1038/s41523-024-00700-z.

Guigal-Stephan, N., Lockhart, B., Moser, T., Heitzer, E., 2025. A perspective review on the systematic implementation of ctDNA in phase I clinical trial drug development. J. Exp. Clin. Cancer Res. 44 (1), 79. https://doi.org/10.1186/s13046-025-03328-4.

Güzel, E., Koçyiğit, Ü.M., Arslan, B.S., Ataş, M., Taslimi, P., Gökalp, F., Nebioğlu, M., Şişman, İ., Gülçin, İ., 2019. Aminopyrazole-substituted metallophthalocyanines: preparation, aggregation behavior, and investigation of metabolic enzymes inhibition properties. Arch. Der Pharm. 352 (2), e1800292. https://doi.org/10.1002/ardp.201800292.

Huang, J., Teng, X., 2020. Expression of PD-L1 for predicting response to immune checkpoint inhibitors in metastatic urothelial carcinoma: a systematic review and meta-analysis. Curr. Oncol. 27 (6), e656. https://doi.org/10.3747/co.27.6437.

Jiang, M., Sun, D., Guo, Y., Guo, Y., Xiao, J., Wang, L., Yao, X., 2020. Assessing PD-L1 expression level by radiomic features from PET/CT in nonsmall cell lung cancer patients: an initial result. Acad. Radiol. 27 (2), 171–179. https://doi.org/10.1016/j.acra.2019.04.016.

Kim, H., Kim, S., Choi, S., Park, C., Park, S., Pereira, S., Ma, M., Yoo, D., Paeng, K., Jung, W., Park, S., 2024. Clinical validation of artificial intelligence–powered PD-L1 tumor proportion score interpretation for immune checkpoint inhibitor response prediction in non–small cell lung cancer. JCO Precis. Oncol. (8), e2300556. https://doi.org/10.1200/po.23.00556.

Knepper, T.C., Montesion, M., Russell, J.S., Sokol, E.S., Frampton, G.M., Miller, V.A., Albacker, L.A., McLeod, H.L., Eroglu, Z., Khushalani, N.I., Sondak, V.K., 2019. The genomic landscape of merkel cell carcinoma and clinicogenomic biomarkers of response to immune checkpoint inhibitor therapy. Clin. Cancer Res. 25 (19), 5961–5971. https://doi.org/10.1158/1078-0432.ccr-18-4159.

Labriola, M.K., Zhu, J., Gupta, R., McCall, S., Jackson, J., Kong, E.F., White, J.R., Cerqueira, G., Gerding, K., Simmons, J.K., George, D., 2020. Characterization of tumor mutation burden, PD-L1 and DNA repair genes to assess relationship to immune checkpoint inhibitors response in metastatic renal cell carcinoma. J. Immunother. Cancer 8 (1), e000319. https://doi.org/10.1136/jitc-2019-000319.

Lee, K.S., Choi, E., Cho, S.I., Park, S., Ryu, J., Puche, A.V., Ma, M., Park, J., Jung, W., Ro, J., Kim, S., 2024. An artificial intelligence-powered PD-L1 combined positive score (CPS) analyser in urothelial carcinoma alleviating interobserver and intersite variability. Histopathology 85 (1), 81–91. https://doi.org/10.1111/his.15176.

Li, J., Guo, Q., Tan, X., 2025. Multi-modal feature integration for thyroid nodule prediction. J. Radiat. Res. Appl. Sci. 18 (1), 101217. https://doi.org/10.1016/j.jrras.2024.101217.

Li, K., Luo, H., Huang, L., Luo, H., Zhu, X., 2020. Microsatellite instability: a review of what the oncologist should know. Cancer Cell Int. 20 (1), 16. https://doi.org/10.1186/s12935-019-1091-8.

Litchfield, K., Reading, J.L., Puttick, C., Thakkar, K., Abbosh, C., Bentham, R., Watkins, T.B., Rosenthal, R., Biswas, D., Rowan, A., Lim, E., 2021. Meta-analysis of tumor-and t cell-intrinsic mechanisms of sensitization to checkpoint inhibition. Cell 184 (3), 596–614. https://doi.org/10.1016/j.cell.2021.01.002.

Ma, S., Zhou, M., Xu, Y., Gu, X., Zou, M., Abudushalamu, G., Yao, Y., Fan, X., Wu, G., 2023. Clinical application and detection techniques of liquid biopsy in gastric cancer. Mol. Cancer 22 (1), 7. https://doi.org/10.1186/s12943-023-01715-z.

Macheka, S., Ng, P.Y., Ginsburg, O., Hope, A., Sullivan, R., Aggarwal, A., 2024. Prospective evaluation of artificial intelligence (AI) applications for use in cancer pathways following diagnosis: a systematic review. BMJ Oncol. 3 (1), e000255. https://doi.org/10.1136/bmjonc-2023-000255.

Marcus, L., Fashoyin-Aje, L.A., Donoghue, M., Yuan, M., Rodriguez, L., Gallagher, P.S., Philip, R., Ghosh, S., Theoret, M.R., Beaver, J.A., Pazdur, R., 2021. FDA approval summary: pembrolizumab for the treatment of tumor mutational burden–high solid tumors. Clin. Cancer Res. 27 (17), 4685–4689. https://doi.org/10.1158/1078-0432.ccr-21-0327.

Marques, A.C., Ferraro-Peyret, C., Michaud, F., Song, L., Smith, E., Fabre, G., Willig, A., Wong, M.M., Xing, X., Chong, C., Brayer, M., 2022. Improved NGS-based detection of microsatellite instability using tumor-only data. Front. Oncol. 12, 969238. https://doi.org/10.3389/fonc.2022.969238.

Mathew, D.E., Ebem, D.U., Ikegwu, A.C., Ukeoma, P.E., Dibiaezue, N.F., 2025. Recent emerging techniques in explainable artificial intelligence to enhance the interpretable and understanding of AI models for human. Neural Process. Lett. 57 (1), 16. https://doi.org/10.1007/s11063-025-11732-2.

Millward, J., He, Z., Nibali, A., Mouradov, D., Mielke, L.A., Tran, K., Chou, A., Hawkins, N.J., Ward, R.L., Gill, A.J., Sieber, O.M., 2025. Automated deep learning-based assessment of tumour-infiltrating lymphocyte density determines prognosis in colorectal cancer. J. Transl. Med. 23 (1), 298. https://doi.org/10.1186/s12967-025-06254-3.

Page, D.B., Broeckx, G., Jahangir, C.A., et al., 2023. Spatial analyses of immune cell infiltration in cancer. J. Pathol. 260 (5), 514–532. https://doi.org/10.1002/path.6165.

Pietrantonio, F., Randon, G., Di Bartolomeo, M., Luciani, A., Chao, J., Smyth, E.C., Petrelli, F., 2020. Predictive role of microsatellite instability for PD-1 blockade in patients with advanced gastric cancer: a meta-analysis of randomized clinical trials. ESMO Open 6, 100036. https://doi.org/10.1016/j.esmoop.2020.100036.

Qin, Y., Huo, M., Liu, X., Li, S.C., 2024. Biomarkers and computational models for predicting efficacy to tumor ICI immunotherapy. Front. Immunol. 15, 1368749. https://doi.org/10.3389/fimmu.2024.1368749.

Raez, L., Usher, J., Habaue, C., Danenberg, K., Hunis, B., Jaimes, Y., Rabizadeh, S., Castrellon, A., Danenberg, P., 2018. MA19. 06 blood based biomarkers: RNA, KRAS and PD-L1 strongly matching with tissue and showing correlation with clinical responses in NSCLC Patient's. J. Thorac. Oncol. 13 (10), S423–S424. https://doi.org/10.1016/j.jtho.2018.08.478.

Randhawa, M.A., Alghamdi, M.S., 2011. Anticancer activity of nigella sativa (black seed)—a review. Am. J. Chin. Med. 39 (06), 1075–1091. https://doi.org/10.1142/S0192415X1100941X.

Rossi, M., Radisky, D.C., 2024. Multiplex digital spatial profiling in breast cancer research: state-of-the-art technologies and applications across the translational science spectrum. Cancers 16 (9), 1615. https://doi.org/10.3390/cancers16091615.

Sadeghi, Z., Alizadehsani, R., Cifci, M.A., Kausar, S., Rehman, R., Mahanta, P., Bora, P.K., Almasri, A., Alkhawaldeh, R.S., Hussain, S., Alatas, B., 2024. A review of explainable artificial intelligence in healthcare. Comput. Electr. Eng. 118, 109370. https://doi.org/10.1016/j.compeleceng.2024.109370.

Shamai, G., Livne, A., Polónia, A., Sabo, E., Cretu, A., Bar-Sela, G., Kimmel, R., 2022. Deep learning-based image analysis predicts PD-L1 status from H&E-stained histopathology images in breast cancer. Nat. Commun. 13 (1), 6753. https://doi.org/10.1038/s41467-022-34275-9.

Sheng, W., Zhang, C., Mohiuddin, T.M., Al-Rawe, M., Zeppernick, F., Falcone, F.H., Meinhold-Heerlein, I., Hussain, A.F., 2023. Multiplex immunofluorescence: a powerful tool in cancer immunotherapy. Int. J. Mol. Sci. 24 (4), 3086. https://doi.org/10.3390/ijms24043086.

Sivapalan, L., Murray, J.C., Canzoniero, J.V., Landon, B., Jackson, J., Scott, S., Lam, V., Levy, B.P., Sausen, M., Anagnostou, V., 2023. Liquid biopsy approaches to capture tumor evolution and clinical outcomes during cancer immunotherapy. J. Immunother. Cancer 11 (1), e005924. https://doi.org/10.1136/jitc-2022-005924.

So, W.V., Dejardin, D., Rossmann, E., Charo, J., 2023. Predictive biomarkers for PD-1/PD-L1 checkpoint inhibitor response in NSCLC: an analysis of clinical trial and real-world data. J. Immunother. Cancer 11 (2), e006464. https://doi.org/10.1136/jitc-2022-006464.

Song, X., Xiong, A., Wu, F., Li, X., Wang, J., Jiang, T., Chen, P., Zhang, X., Zhao, Z., Liu, H., Cheng, L., 2023. Spatial multi-omics revealed the impact of tumor ecosystem heterogeneity on immunotherapy efficacy in patients with advanced non-small cell lung cancer treated with bispecific antibody. J. Immunother. Cancer 11 (2), e006234. https://doi.org/10.1136/jitc-2022-006234.

Stenzinger, A., Allen, J.D., Maas, J., Stewart, M.D., Merino, D.M., Wempe, M.M., Dietel, M., 2019. Tumor mutational burden standardization initiatives: recommendations for consistent tumor mutational burden assessment in clinical samples to guide immunotherapy treatment decisions. Genes Chromosomes Cancer 58 (8), 578–588. https://doi.org/10.1002/gcc.22733.

Thompson, J.C., Carpenter, E.L., Silva, B.A., Rosenstein, J., Chien, A.L., Quinn, K., Espenschied, C.R., Mak, A., Kiedrowski, L.A., Lefterova, M., Nagy, R.J., 2021. Serial monitoring of circulating tumor DNA by next-generation gene sequencing as a biomarker of response and survival in patients with advanced NSCLC receiving pembrolizumab-based therapy. JCO Precis. Oncol. (5), 510–524. https://doi.org/10.1200/po.20.00321.

van der Kooij, M.K., Joosse, A., Suijkerbuijk, K.P., Aarts, M.J., van den Berkmortel, F.W., Blank, C.U., Boers-Sonderen, M.J., van den Eertwegh, A.J., de Groot, J.W.B., Haanen, J.B., Hospers, G.A., 2023. Failure to validate existing clinical prediction scale for response to PD-1 monotherapy in advanced melanoma in national cohort study. Br. J. Cancer 128 (5), 707–710. https://doi.org/10.1038/s41416-022-02088-8.

Van Velzen, M.J.M., Derks, S., Van Grieken, N.C.T., Mohammad, N.H., van Laarhoven, H.W.M., 2020. MSI as a predictive factor for treatment outcome of gastroesophageal adenocarcinoma. Cancer Treat. Rev. 86, 102024. https://doi.org/10.1016/j.ctrv.2020.102024.

Vilar, E., Gruber, S.B., 2010. Microsatellite instability in colorectal cancer—the stable evidence. Nat. Rev. Clin. Oncol. 7 (3), 153–162. https://doi.org/10.1038/nrclinonc.2009.237.

Wang, Y., Tong, Z., Zhang, W., Zhang, W., Buzdin, A., Mu, X., Yan, Q., Zhao, X., Chang, H.H., Duhon, M., Zhou, X., 2021. FDA-approved and emerging next generation predictive biomarkers for immune checkpoint inhibitors in cancer patients. Front. Oncol. 11, 683419. https://doi.org/10.3389/fonc.2021.683419.

Willis, C., Fiander, M., Tran, D., Korytowsky, B., Thomas, J.M., Calderon, F., Zyczynski, T.M., Brixner, D., Stenehjem, D.D., 2019. Tumor mutational burden in lung cancer: a systematic literature review. Oncotarget 10 (61), 6604. https://doi.org/10.18632/oncotarget.27287.

Yamaguchi, H., Hsu, J.M., Sun, L., Wang, S.C., Hung, M.C., 2024. Advances and prospects of biomarkers for immune checkpoint inhibitors. Cell Rep. Med. 5 (7). https://doi.org/10.1016/j.xcrm.2024.101621.

Yang, J., Hu, L., 2019. Immunomodulators targeting the PD-1/PD-L1 protein-protein interaction: from antibodies to small molecules. Med. Res. Rev. 39 (1), 265–301. https://doi.org/10.1002/med.21530.

Yarchoan, M., Hopkins, A., Jaffee, E.M., 2017. Tumor mutational burden and response rate to PD-1 inhibition. New Engl. J. Med. 377 (25), 2500–2501. https://doi.org/10.1056/NEJMc1713444.

Yoo, S.K., Fitzgerald, C.W., Cho, B.A., Fitzgerald, B.G., Han, C., Koh, E.S., Pandey, A., Sfreddo, H., Crowley, F., Korostin, M.R., Debnath, N., 2025. Prediction of checkpoint inhibitor immunotherapy efficacy for cancer using routine blood tests and clinical data. Nat. Med. 31 (3), 869–880. https://doi.org/10.1038/s41591-024-03398-5.