**Introduction**

Immune checkpoint blockade (ICB) has significantly altered the treatment landscape for advanced melanoma, offering prolonged survival in a subset of patients who achieve durable responses. Single-agent antibodies targeting the programmed cell-death protein 1 (PD-1) pathway, such as nivolumab and pembrolizumab, can lead to lasting remissions in some individuals. Still, many patients either fail to respond or eventually develop resistance. Combination therapy with anti–PD–1 and anti–CTLA–4 agents (e.g., ipilimumab plus nivolumab) can further improve response rates and progression-free survival, but often at the cost of substantial immune-related toxicity. This trade-off highlights the need for better biomarkers to guide treatment intensity, ensuring that combination therapy is reserved for those most likely to benefit while sparing others from avoidable toxicity and expense.

Recent work has suggested that features such as high genomic heterogeneity and low ploidy may robustly identify patients intrinsically resistant to PD-1 blockade. In a series of studies, Tarantino et al. used whole-exome sequencing of pretreatment tumors to develop a model with high positive predictive value for resistance to single-agent PD-1 therapy. Their findings also suggested that patients predicted to be intrinsically resistant might benefit disproportionately from combination therapy, further underscoring the potential for a biomarker-guided approach.

Despite these promising insights, there has been limited investigation into the clinical and economic implications of using such biomarkers to personalize first-line immunotherapy decisions. It remains unclear how accurate a predictive model must be to justify the additional toxicity and cost of combination therapy for patients labeled as resistant to PD-1 blockade. This question is especially pertinent given the potential for performance drift or generalizability issues when biomarkers are translated from tightly controlled research settings to broader clinical practice.

In this study, we address these gaps by conducting a cost-effectiveness analysis incorporating the latest survival data from the CheckMate 067 trial and the biomarker model proposed by Tarantino et al. Using a partitioned survival approach, we simulate long-term outcomes and costs over a 10-year horizon for three strategies: (1) single-agent PD-1 blockade for all patients, (2) combination anti–PD–1 and anti–CTLA-4 therapy for all patients, and (3) a biomarker-driven strategy that escalates patients to combination therapy only if they are predicted to be intrinsically resistant. In addition, we introduce a framework for evaluating how variations in model performance (e.g., sensitivity and specificity) affect the overall value of a biomarker-driven approach. By quantifying how good a model needs to be before providing clinical and economic benefit, our work highlights the critical interplay between predictive accuracy and resource utilization in pursuing personalized immunotherapy for advanced melanoma.

**Methods**

**Model Structure and Treatment Strategies**

This analysis adheres to established guidelines for cost-effectiveness research (Sanders et al., 2016; CHEERS, 2022). We developed a partitioned survival model to evaluate three first-line treatment strategies for advanced melanoma over a 10-year horizon with monthly cycles. An overview of the model structure is shown in Figure 1. The strategies compared were:

1. **Nivolumab Monotherapy (NIVO):** Patients receive single-agent PD-1 blockade following the regimen used in CheckMate 067, where nivolumab was administered at 3 mg/kg every two weeks until disease progression or unacceptable toxicity.
2. **Combination Therapy (NIVO+IPI):** Patients receive combination treatment with nivolumab plus ipilimumab as per the CheckMate 067 protocol. In this arm, patients received four doses of ipilimumab (3 mg/kg every three weeks) concurrently with nivolumab (1 mg/kg), followed by nivolumab maintenance therapy. Although combination therapy has demonstrated improved response rates and progression-free survival (PFS), it is associated with higher toxicity and costs.
3. **Biomarker-Guided Strategy:** All patients undergo whole-exome sequencing (WES) to assess genomic heterogeneity and ploidy, following the methodology of Tarantino et al. The Tarantino model classifies patients based on their predicted response to PD-1 blockade. Patients predicted to be intrinsically resistant to PD-1 monotherapy (base-case sensitivity of 60% and specificity of 88%) are triaged to receive NIVO+IPI, while those predicted to be responders receive NIVO. This strategy is designed to mitigate unnecessary toxicity by reserving combination therapy for those most likely to benefit.

Figure : Partitioned Survival Analysis Model for the Study Scenarios

A diagram of a disease

Description automatically generated

**Survival Modeling and Calibration**

Transition probabilities for overall survival (OS) and PFS were derived via partitioned survival analysis using the CheckMate 067 trial data. We digitized published Kaplan–Meier curves using GetData Graph Digitizer and generated pseudo–individual patient data with the IPDfromKM package in R. Competing parametric models—including exponential, Weibull, and log-logistic functions—as well as natural cubic spline models were fitted to the reconstructed data. The best-fitting models were selected based on the Akaike information criterion, ensuring robust survival extrapolation beyond the trial period (Figure 2 displays the overlay of our fitted curves with the observed Kaplan–Meier data).

Figure : Model Validation

**A graph of different types of results

Description automatically generated with medium confidence**

**Counterfactual Survival Estimates**

Recognizing that the NIVO arm in CheckMate 067 represents a composite of patients with varying responses to PD-1 blockade, we derived counterfactual survival functions for two subpopulations: true responders and intrinsically resistant patients. We hypothesized that the survival outcomes of true responders are upper-bounded by those observed in the NIVO+IPI arm, while intrinsically resistant patients have markedly poorer outcomes. The survival function for the NIVO arm was modeled as a weighted average:

,

where  (assumed to be 0.6 in the base case) represents the proportion of true responders. A similar approach was applied to derive PFS curves. Figure 3 presents these counterfactual survival estimates, with the solid line depicting the observed NIVO curve and the dashed lines representing the derived curves for responders and resistant patients. These functions were incorporated into our sensitivity analyses to evaluate the impact of uncertainty in subgroup classification on cost-effectiveness outcomes.

Figure : Base-case Counterfactual OS and PFS for Anti-PD-1 Responders and Resistant

A graph of different colored lines

Description automatically generated

**Costs, Utilities, and Adverse Events**

Direct medical costs in the model encompassed drug acquisition, administration, and treatment-related adverse events (AEs) management. Drug costs were based on national pricing data and adjusted to current U.S. dollars, while AE management costs were derived from published literature and clinical guidelines. Health-state utilities for stable disease and progressive disease were obtained from melanoma-specific studies, with additional disutility applied for AEs. Detailed input parameters, including cost, utility values, and their associated probability distributions for probabilistic sensitivity analysis (PSA), are summarized in Table 1.

**Sensitivity Analyses**

We conducted comprehensive sensitivity analyses to assess the robustness of our findings. One-way sensitivity analyses varied key parameters (e.g., drug costs, AE management costs, utility values, and the performance characteristics of the Tarantino model) over plausible ranges. A two-way sensitivity analysis specifically examined the impact of varying sensitivity and specificity of the Tarantino model on overall quality-adjusted life-years (QALYs). Using Monte Carlo simulation with 10,000 iterations, PSA was performed, assigning gamma distributions to cost parameters and beta distributions to probability and utility inputs. The influence of ML model accuracy on overall outcomes is further illustrated by a heat map of QALYs (Figure 4), which identifies the threshold performance necessary for the biomarker-guided strategy to be cost-effective.

| **Parameter** | **Value (95% CI)** | **Distribution** | **Source** |
| --- | --- | --- | --- |
| **Drug Costs per Cycle** |  |  |  |
| Nivolumab | $14,975 ($9,703–$21,417) | Gamma | AWP |
| Combination (NIVO+IPI) | $26,425 ($17,089–$37,662) | Gamma | AWP |
| Second-line Treatment (NIVO+IPI arm) a | $8,908 ($5,764–$12,735) | Gamma | AWP |
| Second-line Treatment (NIVO arm) a | $12,093 ($7,824–$17,255) | Gamma | AWP |
| **Drug Toxic Effects Decrement** b |  |  |  |
| Disutility for NIVO+IPI Toxicity c | 0.012 | Beta |  |
| Disutility for NIVO Toxicity c | 0.0036 | Beta |  |
| **Disease Costs per Cycle** |  |  |  |
| Stable Disease | $2,166 ($1,397–$3,098) | Gamma | Insinga et al., 2019 |
| Progressed Disease | $4,000 ($2,575–$5,712) | Gamma | Insinga et al., 2019 |
| Palliative Care & Death (1-time cost) | $15,957 ($10,335–$22,818) | Gamma | Insinga et al., 2019 |
| **Health Utilities** |  |  |  |
| Stable Disease Utility (per year) | 0.754 (0.407–0.970) | Beta | Nafees et al., 2017 |
| Disease Progression (Decrement) | 0.180 (0.115–0.367) | Beta | Nafees et al., 2008 |
| Death Utility | 0 | NA | NA |
| **Tarantino Model Accuracies** |  |  |  |
| Positive Predictive Value (PPV) | 90% | Fixed | Tarantino et al. |
| Specificity | 97% | Fixed | Tarantino et al. |
| Sensitivity | 33% | Fixed | Tarantino et al. |

a Calculated as the average cost of treatment using weighted frequencies of individual second-line therapeutic agents received by each treatment arm in the CheckMate 067 clinical trial.

b Calculated as the average cost of toxic effects using weighted frequencies of grade 3 to 4 treatment related adverse events for each treatment arm in the CheckMate 067 clinical trial. Costs of individual toxic effects were derived from the literature and include all care required to manage each toxic effect. References for individual toxic effect costs are summarized in eTable 2 in the Supplement.

c Calculated as the average disutility of toxic effects using weighted frequencies of grade 3 to 4 treatment-related adverse events for each treatment arm in the CheckMate 067 clinical trial. Disutility from experiencing toxic effects occurred over a 1-month period. Disutilities of individual toxic effects were derived from the literature. References for individual toxic effect disutilities are summarized in eTable 3 in the Supplement.

A chart of a model

Description automatically generated

Appendix

Table A1. Disutility from Grade 3–5 Treatment-Related Adverse Events in CheckMate 067

| **Adverse Event Category** | **G3–4 no. (%)** | **Median Time to Resolution (months)** | **Disutility per Month** | **Total Disutility** | **Source** |
| --- | --- | --- | --- | --- | --- |
| *NIVO+IPI (N=314)* |  |  |  |  |  |
| **Skin** (e.g., rash, pruritus, maculopapular rash, vitiligo) | 15 (4.8%) | 0.71 | 0.013 | 0.0092 |  |
| **Gastrointestinal** (e.g., diarrhea, colitis) | 45 (14.3%) | 0.71 | 0.018 | 0.0128 |  |
| **Endocrine** (e.g., hypothyroidism, hyperthyroidism, hypophysitis) | 14 (4.5%) | 4.28 | 0.02 | 0.0856 |  |
| **Hepatic** (e.g., elevated ALT, AST) | 40 (12.7%) | 0.87 | 0.05 | 0.0435 |  |
|  |  |  |  |  |  |
| *Weighted average* | — | — | — | 0.012 | — |
| **Any AE Leading to Discontinuation** | 105 (33.5%) | — | — | — | — |
| *NIVO (N=316)* |  |  |  |  |  |
| **Skin** (e.g., rash, pruritus, maculopapular rash, vitiligo) | 4 (1.3%) | 0.94 | 0.013 | 0.0122 |  |
| **Gastrointestinal** (e.g., diarrhea, colitis) | 8 (2.5%) | 1.24 | 0.018 | 0.0223 |  |
| **Endocrine** (e.g., hypothyroidism, hyperthyroidism, hypophysitis) | 3 (0.9%) | NR\* | 0.02 | 0.0856 |  |
| **Hepatic** (e.g., elevated ALT, AST) | 7 (2.2%) | 1.84 | 0.05 | 0.092 |  |
|  |  |  |  |  |  |
| *Weighted average* | — | — | — | 0.0036 | — |
| **Any AE Leading to Discontinuation** | 29 (9.3%) | — | — | — | — |

**eMethods - Partitioned Survival Analysis for Advanced Melanoma Treatment**

**1. Model Structure and Design**

Our cost-effectiveness analysis employed a partitioned survival model with three health states: progression-free disease, progressed disease, and death. This structure reflects the natural history of advanced melanoma and aligns with established practice in oncology modeling. State occupancy was derived directly from parametric survival curves fitted to the CheckMate 067 trial data with 10-year follow-up.

The model used a 10-year (120-month) time horizon corresponding to the maximum follow-up period in the CheckMate 067 trial. Monthly cycles were implemented to capture treatment dynamics with sufficient granularity, and a standard 3% annual discount rate was applied to both costs and outcomes following health economic guidelines.

**2. Survival Analysis and Data Extraction**

**Digitization of Survival Curves**

To obtain patient-level data for our analysis, we digitized the published Kaplan-Meier curves from the CheckMate 067 trial using the IPDfromKM package in R. This algorithm reconstructs individual patient data by iteratively estimating the underlying individual event and censoring times that would generate the published survival curves. This approach allowed us to work with reconstructed individual-level data rather than relying solely on published summary statistics.

**Parametric Modeling Approach**

We fitted natural cubic spline models to the reconstructed patient-level data for both overall survival (OS) and progression-free survival (PFS). Natural splines were selected after comparison with standard parametric distributions (Weibull, exponential, log-normal, log-logistic, Gompertz, and generalized gamma), as they provided superior fit to the observed data, particularly for capturing the plateau in the survival curves characteristic of immunotherapy.

The spline models were fitted independently to each treatment arm's data due to clear violations of the proportional hazards assumption, as evidenced by log-log plots and statistical testing. This approach ensures that treatment-specific survival patterns are accurately captured.

**Time Horizon Considerations**

Importantly, our model strictly uses the 10-year observed data from CheckMate 067 and does not extrapolate survival beyond this period. This approach eliminates uncertainty associated with extrapolation and ensures that all conclusions are based on actual observed clinical outcomes rather than modeling assumptions about long-term survival.

Our model demonstrated strong validation against the trial data, with predicted 10-year OS rates of 43.8% for nivolumab plus ipilimumab and 37.3% for nivolumab, closely aligning with the reported trial rates of 43% and 37%, respectively.

**3. Treatment Protocol Implementation**

Treatment duration in the model was limited to a maximum of 24 months (2 years), reflecting established clinical practice for immunotherapy. This approach is supported by treatment guidelines indicating that patients who achieve response and discontinue after 2 years typically maintain that response. The trial data further supports this assumption, as at 5-year follow-up, only 8% of nivolumab plus ipilimumab and 18% of nivolumab patients remained on treatment.

Early treatment discontinuation due to adverse events was explicitly modeled based on trial data. CheckMate 067 reported that 44.4% of patients receiving combination therapy and 15.7% receiving nivolumab monotherapy discontinued treatment due to toxicity. We implemented a gradual discontinuation approach over the first three months, reflecting the timing of adverse events observed in the trial. After this period, only 55.6% of combination patients and 84.3% of nivolumab patients continued receiving treatment until progression or until reaching the 24-month cap.

For combination therapy, we separately modeled the induction phase (nivolumab plus ipilimumab) and maintenance phase (nivolumab monotherapy), accounting for the 43% of patients who discontinued during induction based on trial observations.

**4. Cost Parameters**

Drug acquisition costs were estimated at $26,425 per month for combination therapy and $14,975 per month for nivolumab monotherapy. These costs were applied only to patients in the progression-free state who remained on treatment, accounting for early discontinuation patterns and the maximum treatment duration.

Disease management costs of $2,166 per month were applied to patients in the progression-free state, covering routine monitoring, physician visits, and supportive care.

Subsequent treatment costs were modeled according to utilization patterns observed in CheckMate 067. The trial reported that 36% of combination therapy patients and 50% of nivolumab patients received subsequent systemic therapy after progression. Rather than assuming all progressed patients received second-line therapy, we adjusted the costs proportionally to reflect real-world utilization. The base monthly costs ($8,908 for post-combination and $12,093 for post-nivolumab) were multiplied by these utilization rates, resulting in adjusted costs of $3,207 and $6,047 per month, respectively. This approach more accurately captures the economic benefit of reduced subsequent therapy needs with combination treatment.

For local therapy costs (radiation or surgery), we incorporated a one-time cost of $5,000 at the time of progression, adjusted by the proportion of patients receiving such interventions in the trial (46% for combination therapy and 56% for nivolumab). Additionally, a one-time end-of-life cost of $15,957 was applied at death for all patients to account for palliative care.

**5. Health Utility Parameters**

Health state utilities were set at 0.754 for progression-free disease and 0.574 for progressed disease (applying a decrement of 0.180), based on published quality-of-life data for advanced melanoma.

Adverse event disutilities were derived from a detailed analysis of grade 3-4 treatment-related adverse events reported in CheckMate 067. From the trial data, we calculated weighted average disutility values by combining:

1. The incidence of each adverse event type (skin, gastrointestinal, endocrine, hepatic)
2. The median time to resolution for each event type
3. The established disutility weight per event type

This analysis yielded weighted average disutilities of 0.012 for combination therapy and 0.0036 for nivolumab monotherapy. The calculation method is illustrated below:

For nivolumab plus ipilimumab:

* Skin events: 4.8% incidence, 0.71 months duration, 0.013 disutility = 0.0092 total impact
* Gastrointestinal events: 14.3% incidence, 0.71 months duration, 0.018 disutility = 0.0128 total impact
* Endocrine events: 4.5% incidence, 4.28 months duration, 0.02 disutility = 0.0856 total impact
* Hepatic events: 12.7% incidence, 0.87 months duration, 0.05 disutility = 0.0435 total impact

The weighted average of these impacts (0.012) was then applied in a time-dependent manner to reflect the changing burden of adverse events over time:

* Full impact (100%) in months 1-3, when most adverse events occur and are most severe
* Moderate impact (60%) in months 4-6, as many events begin resolving
* Minimal impact (20%) in months 7-24, primarily reflecting persistent endocrine events
* No impact after 24 months

This time-dependent approach is supported by the trial's documentation of adverse event timing and resolution patterns. Table S13 in the CheckMate 067 appendix shows that most adverse events resolved within weeks to months, with median resolution times of 3-9 weeks for most events, though endocrine events often persisted longer.

**6. Model Validation and Sensitivity Analyses**

We validated our model through multiple approaches. Internal validation confirmed that predicted survival outcomes closely matched observed trial data. Face validity was established by ensuring that survival patterns showed the expected plateau characteristic of immunotherapy and that economic outcomes aligned with clinical expectations.

Sensitivity analyses were conducted to test the robustness of our findings. A probabilistic sensitivity analysis with 1,000 Monte Carlo simulations varied all key parameters simultaneously, including survival model parameters, utilities, costs, and adverse event incidences. Deterministic sensitivity analyses examined the impact of individual parameters, such as discount rates, utility values, cost inputs, and treatment duration assumptions.

Additional scenario analyses explored alternative parametric survival distributions, different treatment durations, and variations in early discontinuation rates to further test the stability of our conclusions.

**7. Limitations**

Our model has several limitations worth noting. While our approach of using natural splines with the complete 10-year trial data eliminates extrapolation uncertainty, it is bound by the follow-up period of the trial. Our approach to modeling treatment discontinuation simplifies complex real-world patterns. The costs of subsequent therapies are based on aggregate estimates rather than patient-level data on specific regimens. Additionally, our model does not explicitly account for heterogeneity in treatment effects across different patient subgroups, which may be clinically relevant.

Despite these limitations, our partitioned survival analysis provides a robust framework for evaluating the cost-effectiveness of nivolumab plus ipilimumab compared to nivolumab monotherapy in advanced melanoma, based on the most mature clinical data available without requiring extrapolation beyond the observed trial period.