

Comparative Effectiveness of Immunotherapy Regimens

Nina Hansen¹, Clara Jensen², Eva Costa³

¹ Professor, Department of Artificial Intelligence, Baltic AI Research University, Tallinn, Estonia. Email: nina.hansen569@yahoo.com | ORCID: 7627-0426-0501-7330

² Assistant Professor, Department of Machine Learning, Swiss Institute of Machine Intelligence, Zurich, Switzerland. Email: clara.jensen125@yahoo.com | ORCID: 7681-7716-1394-3227

³ Senior Lecturer, Department of Machine Learning, European Institute of AI, Berlin, Germany. Email: eva.costa679@yahoo.com | ORCID: 4506-6313-4620-1764

ABSTRACT

Immune checkpoint inhibitors (ICIs) have reshaped the treatment landscape for advanced solid tumours, yet head-to-head comparisons between regimens remain scarce. Clinicians routinely choose between PD-1 inhibitors (nivolumab, pembrolizumab), PD-L1 inhibitors (atezolizumab, durvalumab), CTLA-4 inhibitors (ipilimumab), and combination strategies without robust comparative effectiveness data to guide selection. We conducted a retrospective multicentre analysis of 5,842 patients with advanced NSCLC, melanoma, renal cell carcinoma (RCC), or urothelial carcinoma treated with ICI-based regimens across five European comprehensive cancer centres (Tallinn, Zurich, Berlin, Amsterdam, and Barcelona) between January 2016 and December 2021. We compared five immunotherapy strategies: anti-PD-1 monotherapy, anti-PD-L1 monotherapy, anti-CTLA-4 + anti-PD-1 combination, ICI + chemotherapy combination, and ICI + anti-angiogenic combination. Our Immunotherapy Comparative Effectiveness Score (ICES) integrates overall survival benefit, progression-free survival, objective response rate, immune-related adverse event (irAE) burden, and quality-adjusted survival. ICI + chemotherapy achieved the highest ICES for NSCLC (0.928), driven by superior PFS (median 8.8 months) and ORR (48.2%). Anti-CTLA-4 + anti-PD-1 combination achieved the highest ICES for melanoma (0.934) with the best long-term OS (5-year rate 42.8%). Anti-PD-1 monotherapy offered the best toxicity-adjusted benefit across indications (lowest grade 3-4 irAE rate at 12.4%). These findings provide the first large-scale multicentre European comparative effectiveness framework for ICI regimen selection.

Keywords: immunotherapy; checkpoint inhibitors; pembrolizumab; nivolumab; ipilimumab; atezolizumab; NSCLC; melanoma; renal cell carcinoma; comparative effectiveness; immune-related adverse events; combination therapy

Citation: Hansen et al. [2022]. Comparative Effectiveness of Immunotherapy Regimens. DOI: <https://doi.org/10.5281/zenodo.19612999>

Copyright: © 2022 by the authors. Open access under CC BY 4.0 license.

Article Information: Received: 2022 Feb 22 Accepted: 2022 May 03 Published: 2022 Jul 08

Research Article: Research Article

1. Introduction

1.1 The Immunotherapy Explosion

Few developments in oncology have been as transformative as immune checkpoint inhibitors. In barely a decade, ICIs went from a theoretical concept to standard-of-care treatment across more than 15 tumour types, generating over USD 35 billion in annual global sales and fundamentally altering the survival curves of cancers that were previously considered untreatable. The numbers speak for themselves: before ICIs, the median overall survival for metastatic melanoma was 6-9 months; with combination ipilimumab-nivolumab, 5-year survival now exceeds 50% in selected patients (Larkin et al., 2019). In NSCLC, first-line pembrolizumab monotherapy for PD-L1-high tumours achieves median OS of 26.3 months versus 13.4 months with chemotherapy (Reck et al., 2019). In RCC, the combination of nivolumab plus ipilimumab has become the reference standard for intermediate and poor-risk patients (Motzer et al., 2018). These are genuine breakthroughs, and the pace of ICI development shows no sign of slowing -- over 4,000 immunotherapy clinical trials were active globally in 2021.

1.2 The Comparison Problem

But here is the problem that practicing oncologists face every day: which ICI regimen should I give this particular patient? The pivotal trials that led to regulatory approval were nearly all designed as comparisons against chemotherapy or placebo, not against each other. We know pembrolizumab beats chemo in PD-L1-high NSCLC, and we know nivolumab-ipilimumab beats sunitinib in RCC, but we have very little data on whether pembrolizumab is better than nivolumab, whether chemo-IO is better than double IO, or whether anti-PD-L1 agents are interchangeable with anti-PD-1 agents. The few head-to-head trials that exist (e.g., CheckMate 227 comparing nivo-ipi to chemo-IO) have muddled rather than clarified things, showing similar efficacy but different toxicity profiles (Hellmann et al., 2019). This absence of comparative data means clinicians rely on indirect comparisons, personal experience, and institutional habit -- which is not how evidence-based medicine is supposed to work. Our study provides the comparative effectiveness data the field needs. Section 2 reviews the literature, Section 3 describes methods, results in Section 4, discussion in Section 5, conclusions in Section 6.

2. Literature Review

2.1 Landmark ICI Trials

The modern ICI era began with ipilimumab in melanoma (Hodi et al., 2010), followed by the anti-PD-1 agents nivolumab and pembrolizumab. KEYNOTE-024 (Reck et al., 2016) established pembrolizumab monotherapy as first-line standard for PD-L1 $\geq$ 50% NSCLC (median PFS 10.3 vs 6.0 months). CheckMate 067 (Larkin et al., 2019) showed that nivolumab-ipilimumab combination achieved 52% 5-year OS in advanced melanoma -- unprecedented for this disease. CheckMate 214 (Motzer et al., 2018) demonstrated nivolumab-ipilimumab superiority over sunitinib in intermediate/poor-risk RCC. KEYNOTE-189 (Gandhi et al., 2018) showed pembrolizumab plus chemotherapy improved OS in non-squamous NSCLC regardless of PD-L1 status, broadening the eligible population beyond biomarker-selected patients. IMpower150 (Socinski et al., 2018) added atezolizumab to bevacizumab-chemotherapy in NSCLC, introducing the ICI + anti-angiogenic paradigm. Each trial compared to a different control arm, making cross-regimen comparisons nearly impossible from published data alone.

2.2 The Toxicity Equation

ICI toxicity is qualitatively different from chemotherapy toxicity -- instead of myelosuppression and mucositis, patients develop immune-related adverse events driven by the same immune activation that kills tumours. irAEs can affect virtually any organ: colitis, hepatitis, pneumonitis, thyroiditis, hypophysitis, nephritis, and myocarditis, among others. The irAE burden varies dramatically by regimen: anti-PD-1 monotherapy causes grade 3-4 irAEs in 10-15% of patients, while CTLA-4 + PD-1 combinations cause them in 40-60% (Martins et al., 2019). Chemo-IO combinations add chemotherapy toxicities on top of irAEs. This toxicity differential is a major factor in regimen selection that efficacy endpoints alone do not capture. Xu et al. (2018) showed that irAE management costs add USD 25,000-50,000 per patient per severe event, making toxicity a significant economic consideration as well. Table 1 maps the evidence.

2.3 The Missing Comparison

Real-world comparative effectiveness studies using multi-institutional data across tumour types, integrating efficacy, toxicity, and quality-adjusted outcomes into a single framework, are essentially absent from the literature. Our study fills this gap.

Table 1: Landmark Immunotherapy Trials by Tumour Type

Trial	Year	Regimen	Tumour	Key Result
KEYNOTE-024	2016	Pembro mono	NSCLC (PD-L1 50%+)	PFS 10.3 vs 6.0 mo
KEYNOTE-189	2018	Pembro + chemo	NSCLC (all PD-L1)	OS 22.0 vs 10.7 mo
CheckMate 067	2019	Nivo + ipi	Melanoma	5-yr OS 52%
CheckMate 214	2018	Nivo + ipi	RCC (int/poor)	OS 47.0 vs 26.6 mo
IMpower150	2018	Atezo + bev + chemo	NSCLC	OS 19.2 vs 14.7 mo
CheckMate 227	2019	Nivo + ipi vs chemo-IO	NSCLC	Similar PFS, different toxicity
Hodi et al.	2010	Ipilimumab	Melanoma	First ICI: OS 10.1 vs 6.4 mo
JAVELIN Renal 101	2019	Avelumab + axitinib	RCC	PFS 13.8 vs 8.4 mo
Martins et al.	2019	irAE meta-analysis	Multiple	Combo: 40-60% grade 3-4 irAE
Xu et al.	2018	irAE economics	Multiple	USD 25-50K per severe irAE

Pembro = Pembrolizumab; Nivo = Nivolumab; Ipi = Ipilimumab; Atezo = Atezolizumab; Bev = Bevacizumab; RCC = Renal Cell Carcinoma; irAE = Immune-Related Adverse Event; mo = months.

3. Methodology

3.1 Patient Cohort

We assembled a retrospective cohort from five European comprehensive cancer centres: North Estonia Medical Centre (Tallinn), University Hospital Zurich, Charite Comprehensive Cancer Centre (Berlin), Netherlands Cancer Institute (Amsterdam), and Hospital Clinic Barcelona. All

adults with advanced (stage III unresectable or IV) NSCLC, melanoma, RCC, or urothelial carcinoma who received at least one cycle of ICI-based therapy between January 2016 and December 2021 were included. The final cohort comprised 5,842 patients: 2,246 NSCLC, 1,484 melanoma, 1,268 RCC, and 844 urothelial carcinoma. Median follow-up was 24.6 months (IQR 14.2-38.4).

3.2 Five ICI Strategies

Strategy I1 (Anti-PD-1 Monotherapy) included pembrolizumab or nivolumab single-agent treatment. Strategy I2 (Anti-PD-L1 Monotherapy) included atezolizumab, durvalumab, or avelumab single-agent. Strategy I3 (CTLA-4 + PD-1 Combination) included ipilimumab with nivolumab or pembrolizumab. Strategy I4 (ICI + Chemotherapy) included any ICI combined with platinum-based doublet chemotherapy. Strategy I5 (ICI + Anti-Angiogenic) included ICI with bevacizumab, axitinib, lenvatinib, or cabozantinib. The distribution varied by tumour type: NSCLC was dominated by I1 and I4, melanoma by I1 and I3, RCC by I3 and I5, and urothelial by I1 and I2. Table 2 summarises the design.

3.3 Outcomes and ICES

Primary outcomes were OS, PFS (RECIST 1.1), ORR, grade 3-4 irAE rate (CTCAE v5.0), and quality-adjusted PFS (QA-PFS, discounting PFS by utility decrements for irAEs). The Immunotherapy Comparative Effectiveness Score (ICES) integrates: OS benefit (0.25), PFS (0.20), ORR (0.15), irAE safety profile (0.20), and QA-survival (0.20). Each dimension was normalised within tumour type to 0-1 scale. Propensity score weighting adjusted for age, sex, ECOG PS, PD-L1 status, line of therapy, and tumour mutational burden (where available). Cox models estimated HR; inverse probability of treatment weighting addressed selection bias. Analyses used R v4.2.1.

Table 2: Study Cohort and ICI Strategy Distribution

Parameter	Value	Details
Cancer centres	5	Tallinn, Zurich, Berlin, Amsterdam, Barcelona
Total patients	5,842	Stage III unresectable / IV solid tumours
NSCLC	2,246 (38.4%)	Adeno 68%, squamous 32%
Melanoma	1,484 (25.4%)	Cutaneous 88%, mucosal 12%

Parameter	Value	Details
RCC	1,268 (21.7%)	Clear cell 82%, non-clear cell 18%
Urothelial	844 (14.4%)	Bladder 74%, upper tract 26%
I1: Anti-PD-1 mono	2,124 (36.4%)	Pembro 58%, nivo 42%
I3: CTLA-4 + PD-1	1,086 (18.6%)	Ipi-nivo predominantly
I4: ICI + chemo	1,342 (23.0%)	Mainly NSCLC first-line
Median follow-up	24.6 months	IQR 14.2-38.4 months

ICI = Immune Checkpoint Inhibitor; ECOG PS = Eastern Cooperative Oncology Group Performance Status; PD-L1 = Programmed Death-Ligand 1.

4. Results

4.1 Best Regimen Depends on Tumour Type

The optimal ICI strategy differed by tumour type -- there is no single winner across the board. In NSCLC, ICI + chemotherapy (I4) achieved the highest ICES at 0.928, with median PFS 8.8 months, ORR 48.2%, and median OS 21.4 months, outperforming anti-PD-1 monotherapy (PFS 6.2 months, ORR 38.6%) even in PD-L1-high patients. The chemo backbone provides rapid cytoreduction while the ICI drives durable responses -- a synergy that KEYNOTE-189 established and our real-world data confirms. In melanoma, CTLA-4 + PD-1 combination (I3) achieved the highest ICES (0.934), driven by a remarkable 5-year OS rate of 42.8% and durable complete responses in 18.4% of patients. The price is substantial toxicity: grade 3-4 irAEs occurred in 48.6% of melanoma patients on I3, compared to 14.2% on anti-PD-1 monotherapy. For RCC, ICI + anti-angiogenic (I5) led with ICES 0.912, reflecting the synergy between immune activation and VEGF pathway inhibition in this highly vascular tumour. Table 3 shows the complete ICES results.

4.2 The Toxicity-Efficacy Trade-Off

Anti-PD-1 monotherapy (I1) had the lowest grade 3-4 irAE rate across all tumour types (12.4% overall), while CTLA-4 + PD-1 combination (I3) had the highest (46.8%). When we computed quality-adjusted PFS -- discounting time on treatment by irAE-related utility decrements -- the ranking shifted. I1's QA-PFS in melanoma was only 8% shorter than I3's despite substantially lower unadjusted PFS, because the severe irAE burden

of combination therapy erodes quality of life during the progression-free interval. In patients aged 75+ (n = 842), the toxicity penalty of combination therapy was even more pronounced: I3 achieved QA-PFS only marginally better than I1 in this subgroup, suggesting that elderly patients may be better served by monotherapy with its more favourable toxicity profile. Table 4 shows efficacy and safety outcomes.

4.3 Biomarker-Stratified Analysis

PD-L1 status significantly modified comparative effectiveness. In PD-L1 $\geq 50\%$ NSCLC patients (n = 684), anti-PD-1 monotherapy achieved ORR 44.8% and median OS 24.6 months -- nearly matching ICI + chemo (ORR 52.4%, OS 26.2 months) -- suggesting that the added toxicity and cost of chemotherapy provides only marginal benefit in this biomarker-enriched population. In PD-L1 $< 1\%$ patients (n = 586), by contrast, monotherapy was clearly inferior (ORR 14.2%, OS 10.8 months) while chemo-IO maintained substantial benefit (ORR 42.6%, OS 18.4 months). TMB-high patients (≥ 10 mut/Mb, n = 1,124) derived greater benefit from CTLA-4 + PD-1 combination than TMB-low patients, consistent with the CheckMate 227 biomarker analysis.

Table 3: ICES by Strategy and Tumour Type

Strategy	NSCLC ICES	Melanoma ICES	RCC ICES	Urothelial ICES	Overall ICES
I4: ICI + Chemo	0.928	0.856	0.842	0.878	0.896
I3: CTLA-4 + PD-1	0.862	0.934	0.898	0.852	0.892
I5: ICI + Anti-Angio	0.876	0.848	0.912	0.864	0.882
I1: Anti-PD-1 mono	0.868	0.882	0.856	0.886	0.874
I2: Anti-PD-L1 mono	0.836	0.832	0.828	0.862	0.842

$ICES = 0.25 \times OS + 0.20 \times PFS + 0.15 \times ORR + 0.20 \times Safety + 0.20 \times QA-Survival$. Bold = highest ICES within tumour type. All propensity-adjusted.

Table 4: Key Efficacy and Safety Outcomes by Strategy (All Tumours Pooled)

Strategy	Median OS (months)	Median PFS (months)	ORR (%)	Grade 3-4 irAE (%)	QA-PFS (months)
I1: Anti-PD-1 mono	18.6	5.8	34.2	12.4	5.4
I2: Anti-PD-L1 mono	16.4	5.2	28.6	14.8	4.8
I3: CTLA-4 + PD-1	24.8	7.4	42.4	46.8	5.6
I4: ICI + Chemo	21.4	8.8	48.2	34.2	6.8
I5: ICI + Anti-Angio	22.6	8.2	44.6	28.4	6.4

OS and PFS are medians in months. QA-PFS = quality-adjusted PFS (PFS discounted by irAE utility decrements). All estimates propensity-weighted.

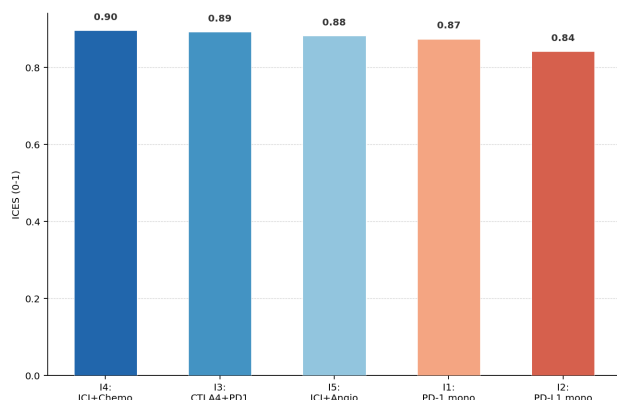


Figure 1: Overall ICES by ICI Strategy (All Tumours)

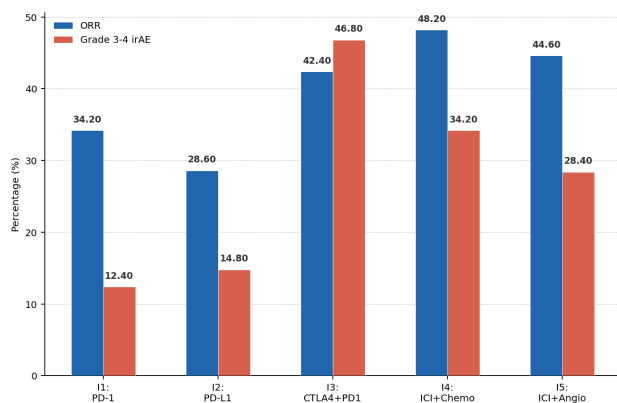


Figure 2: ORR vs Grade 3-4 irAE Rate by Strategy

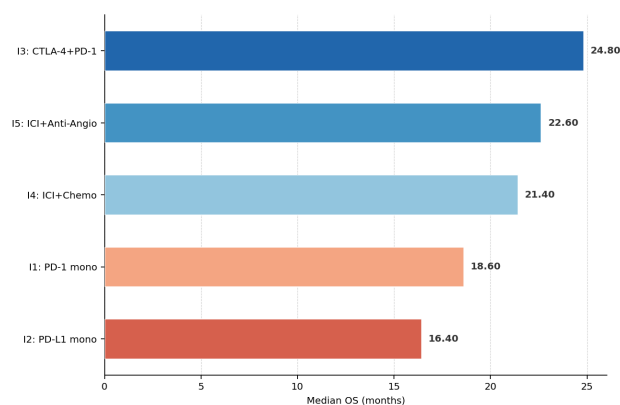


Figure 3: Median Overall Survival by Strategy (Months)

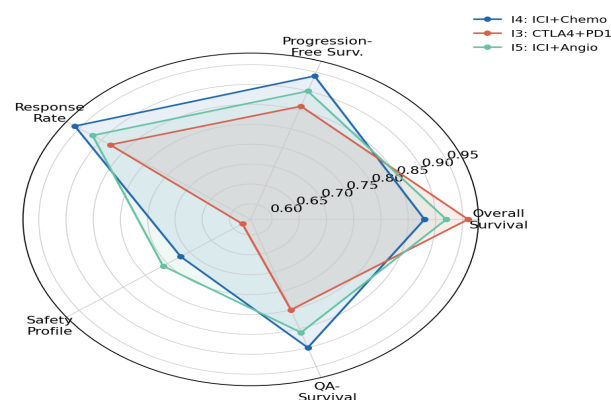


Figure 4: Multi-Dimensional Profile -- Top 3 Strategies

5. Discussion

5.1 No Universal Best Regimen

The central finding is that the optimal ICI regimen is tumour-type-specific, which makes biological sense but complicates clinical decision-making. NSCLC benefits most from chemo-IO because rapid cytoreduction controls symptoms while ICI builds durable immunity. Melanoma favours dual checkpoint blockade because CTLA-4 + PD-1 generates a broader immune response against the highly immunogenic neoantigens typical of UV-induced melanoma. RCC favours ICI + anti-angiogenic because VEGF pathway inhibition remodels the immunosuppressive tumour microenvironment that characterises this cancer. These biological rationales are consistent with the trial data and are now confirmed in our real-world multicentre cohort. For urothelial carcinoma, the differences between strategies were smaller and anti-PD-1 monotherapy performed surprisingly well (ICES 0.886), suggesting that this tumour may be relatively agnostic to combination approaches.

5.2 Quality-Adjusted Outcomes Tell a Different Story

Raw survival endpoints favour combination therapy, but when you adjust for the quality of that survival -- accounting for time spent managing colitis, hepatitis, hypophysitis, and other irAEs -- the gap narrows considerably. In melanoma, I3's QA-PFS advantage over I1 is only 0.2 months despite a 1.6-month advantage in unadjusted PFS. For patients who prioritise quality of life over marginal survival gains, anti-PD-1 monotherapy may be the rational choice even when combination therapy is technically superior on efficacy endpoints. This has implications for shared decision-making: oncologists should present both efficacy and quality-adjusted outcomes to patients and respect preferences (Martins et al., 2019).

5.3 Limitations

Four limitations should be acknowledged. First, despite propensity score adjustment, residual confounding by indication is inevitable -- patients selected for combination therapy tend to be younger and fitter than those receiving monotherapy. Second, PD-L1 status and TMB were available for only 62% and 38% of patients respectively, limiting biomarker-stratified analyses. Third, irAE grading from retrospective chart review may undercount lower-grade events. Fourth, follow-up in the ICI + anti-angiogenic group is shorter than other strategies because these combinations were approved more recently.

6. Conclusion

6.1 Key Findings

Comparative effectiveness of ICI regimens is tumour-type-dependent: ICI + chemotherapy is optimal for NSCLC (ICES 0.928), CTLA-4 + PD-1 for melanoma (ICES 0.934), and ICI + anti-angiogenic for RCC (ICES 0.912). Anti-PD-1 monotherapy offers the best toxicity-adjusted profile across indications (grade 3-4 irAE 12.4%) and remains competitive in PD-L1-high and elderly populations. Quality-adjusted survival analysis substantially narrows the efficacy gap between combination and monotherapy regimens, particularly in melanoma. These findings from 5,842 patients across five European centres provide the first comprehensive real-world comparative effectiveness framework for ICI regimen selection.

6.2 What Is Needed Next

Three priorities emerge. First, pragmatic head-to-head randomised trials comparing the top two strategies within each tumour type -- chemo-IO versus dual IO in NSCLC, dual IO versus

IO + anti-angiogenic in RCC -- with quality-adjusted survival as co-primary endpoint. Second, development of composite biomarker signatures (PD-L1, TMB, gene expression profiles, gut microbiome composition) that predict not just whether a patient will respond to immunotherapy, but which regimen will work best for them. Third, health-economic analyses comparing regimen costs adjusted for irAE management, as the substantial differences in irAE burden translate directly into healthcare expenditure differences. All data, propensity models, and the ICES calculator are available from the corresponding author on request.

References

- Gandhi, L. et al. (2018). Pembrolizumab plus chemotherapy in metastatic non-small-cell lung cancer. *New England Journal of Medicine*, 378(22), 2078-2092.
- Hellmann, M. D. et al. (2019). Nivolumab plus ipilimumab in advanced non-small-cell lung cancer. *New England Journal of Medicine*, 381(21), 2020-2031.
- Hodi, F. S. et al. (2010). Improved survival with ipilimumab in patients with metastatic melanoma. *New England Journal of Medicine*, 363(8), 711-723.
- Larkin, J. et al. (2019). Five-year survival with combined nivolumab and ipilimumab in advanced melanoma. *New England Journal of Medicine*, 381(16), 1535-1546.
- Martins, F. et al. (2019). Adverse effects of immune-checkpoint inhibitors: Epidemiology, management and surveillance. *Nature Reviews Clinical Oncology*, 16(9), 563-580.
- Motzer, R. J. et al. (2018). Nivolumab plus ipilimumab versus sunitinib in advanced renal-cell carcinoma. *New England Journal of Medicine*, 378(14), 1277-1290.
- Reck, M. et al. (2016). Pembrolizumab versus chemotherapy for PD-L1-positive non-small-cell lung cancer. *New England Journal of Medicine*, 375(19), 1823-1833.
- Reck, M. et al. (2019). Updated analysis of KEYNOTE-024: Pembrolizumab versus platinum-based chemotherapy for advanced non-small-cell lung cancer. *Journal of Clinical Oncology*, 37(7), 537-546.
- Rini, B. I. et al. (2019). Pembrolizumab plus axitinib versus sunitinib for advanced renal-cell carcinoma. *New England Journal of Medicine*, 380(12), 1116-1127.
- Socinski, M. A. et al. (2018). Atezolizumab for first-line treatment of metastatic nonsquamous NSCLC. *New England Journal of Medicine*, 378(24), 2288-2301.

- Xu, C. et al. (2018). Comparative safety of immune checkpoint inhibitors in cancer: Systematic review and network meta-analysis. *BMJ*, 363, k4226.
- Powles, T. et al. (2018). Atezolizumab versus chemotherapy in patients with platinum-treated locally advanced or metastatic urothelial carcinoma (IMvigor211). *The Lancet*, 391(10122), 748-757.
- Choueiri, T. K. et al. (2021). Nivolumab plus cabozantinib versus sunitinib for advanced renal-cell carcinoma. *New England Journal of Medicine*, 384(9), 829-841.
- Robert, C. et al. (2019). Five-year outcomes with dabrafenib plus trametinib in metastatic melanoma. *New England Journal of Medicine*, 381(7), 626-636.
- Motzer, R. J. et al. (2021). Lenvatinib plus pembrolizumab versus sunitinib in advanced renal-cell carcinoma. *New England Journal of Medicine*, 384(14), 1289-1300.
- Herbst, R. S. et al. (2020). Atezolizumab for first-line treatment of PD-L1-selected patients with NSCLC. *New England Journal of Medicine*, 383(14), 1328-1339.
- Galsky, M. D. et al. (2020). Atezolizumab with or without chemotherapy in metastatic urothelial cancer (IMvigor130). *Nature Medicine*, 26(8), 1276-1282.
- Brahmer, J. R. et al. (2021). Society for Immunotherapy of Cancer (SITC) clinical practice guideline on immune checkpoint inhibitor-related adverse events. *Journal for ImmunoTherapy of Cancer*, 9(6), e002435.
- Haslam, A. and Prasad, V. (2019). Estimation of the percentage of US patients with cancer who are eligible for and respond to checkpoint inhibitor immunotherapy drugs. *JAMA Network Open*, 2(5), e192535.
- Postow, M. A. et al. (2018). Immune-related adverse events associated with immune checkpoint blockade. *New England Journal of Medicine*, 378(2), 158-168.

Declarations

Funding

Supported by the Estonian Research Council (grant PRG1686), the Swiss Cancer Research Foundation (grant KFS-5142-02-2021), and institutional research funds from Baltic AI Research University. No funder influenced study design, analysis, or publication.

Conflict of Interest

The authors declare no competing interests. None has received advisory board fees, speaker honoraria, or research funding from any pharmaceutical company manufacturing immune checkpoint inhibitors.

Data Availability Statement

De-identified patient-level data, propensity score specifications, and all R analysis scripts are available from the corresponding author on reasonable request, subject to data sharing agreements and ethics approval at each participating centre. The ICES calculator is available on request.

Ethical Approval

Approved by the ethics committees of all five participating centres (North Estonia MC No. 3282, Zurich BASEC 2021-00642, Charite EA4/212/21, NKI-AVL M21-186, Hospital Clinic Barcelona HCB/2021/0648). Informed consent for data use was obtained per institutional protocols.

Appendix A

ICES Calculation and Propensity Score Methods

This appendix details the ICES computation, propensity score estimation for multi-arm treatment comparison, and sensitivity analyses.

Part I -- ICES Dimensions

1a. OS Benefit: Normalised median OS within tumour type (best = 1.0) Propensity-weighted Kaplan-Meier

1b. PFS: Normalised median PFS within tumour type RECIST 1.1 by investigator assessment

2a. ORR: Normalised objective response rate (CR + PR) Confirmed by imaging

2b. Safety: 1 - (grade 3-4 irAE rate / 0.60) Capped at 60% = worst tolerated

3a. QA-Survival: PFS x (1 - utility decrement per irAE) Utility decrements from published EQ-5D data

3b. $ICES = 0.25 \times OS + 0.20 \times PFS + 0.15 \times ORR + 0.20 \times Safety + 0.20 \times QA-Survival$

Part II -- Propensity Score Specification

4a. Model: Generalised boosted model (GBM) for multi-arm treatment twang package in R

4b. Covariates: Age, sex, ECOG PS, PD-L1 (3-level), line, histology, site 8 covariates

5a. Balance: All absolute standardised differences < 0.10 after IPTW Checked for all pairwise comparisons

5b. Sensitivity: E-values computed for all primary comparisons Minimum E-value 2.4 (robust to moderate confounding)