

Electronic Supplementary Material

Article title

Antiangiogenic Class and Trial Population Are Confounded in First-Line Immunotherapy for Hepatocellular Carcinoma: A Bayesian Network Meta-Analysis of Nine Randomised Trials

Journal

BMC Gastroenterology

Authors

Burkhoniddin Tillashaykhov, Iroda Ikromova, Dilnavoz Berdiyeva

Affiliation

Central Asian University, School of Medicine, 264 Milliy Bog St, Tashkent 111221, Uzbekistan

Corresponding author

Dilnavoz Berdiyeva, PhD — d.berdiyeva@centralasian.uz

Registration

PROSPERO CRD420261353528

Supplementary Material

Comparative Efficacy and Toxicity of Antiangiogenic Classes in First-Line Immunotherapy Combinations for Advanced Hepatocellular Carcinoma: A Bayesian Network Meta-Analysis

Contents

Supplementary Methods — heterogeneity priors and model

Supplementary Table S1 — Baseline characteristics

Supplementary Table S2 — PFS and ORR estimates

Supplementary Table S3 — Sensitivity analyses (OS)

Supplementary Table S4 — GRADE certainty of evidence (OS)

Supplementary Table S5 — League table, all pairwise OS comparisons

Supplementary Table S6 — Risk of bias (RoB 2), domain-level

Supplementary Table S7—Effect-modifier correlations and meta-regression

Supplementary Table S8 — GRADE certainty for the class-level comparison (anti-VEGF antibody vs VEGFR-TKI)

Supplementary Figure S1 — PRISMA-NMA flow diagram

Supplementary Figure S2 — Grade ≥ 3 treatment-related adverse events, odds ratios versus sorafenib

Supplementary Methods. Heterogeneity priors and model specification

Because the network is a star centred on sorafenib with a single trial per treatment, the between-study heterogeneity variance (τ^2) is not identifiable from the data. We specified τ^2 using the empirical predictive distributions of Turner et al. (2015), which summarise heterogeneity across large collections of meta-analyses by outcome type, using the pharmacological-versus-pharmacological column matched to each endpoint: overall survival, log-normal τ^2 with location -4.18 and scale 1.41^2 ; progression-free survival, -3.95 and 1.79^2 ; objective response, -2.29 and 1.58^2 ; grade ≥ 3 adverse events, -2.10 and 1.58^2 . The posterior median τ for overall survival was approximately 0.12 . Models were fitted in PyMC (Python) using the No-U-Turn Sampler, four chains of 8000 draws after 6000 tuning iterations; convergence was confirmed by $R\text{-hat} < 1.01$ and the absence of divergent transitions. Treatment effects are expressed relative to sorafenib, with lenvatinib bridging pembrolizumab-plus-lenvatinib through REFLECT and LEAP-002.

Supplementary Table S1. Baseline characteristics of included trials

Trial	Regimen	n ICI	n ctrl	Design	Phase	Region	HBV %	HCV %	Non-viral%	CP-A %	MVI/EHS%	AFP≥400%	f/u mo
ORIENT-32	Sintilimab + bevacizumab*	380	191	OL	II/III	China	94.2	2.5	3.3	95.8	79.3	43.1	10
Finotonlimab + bev	Finotonlimab + bevacizumab*	230	116	OL	II/III	China	89	4.9	6.1	93	79.5	47.4	19.7
HEPATORCH	Toripalimab + bevacizumab	162	164	OL	III	China/TW/SG	89.9	4.6	5.5	100	79	47	16.4
APOLLO	Anlotinib + penpulimab	433	216	OL	III	China	84.1	3.5	12.3	92.4	80.3	49.2	15.3
CARES-310	Camrelizumab + rivoceranib	272	271	OL	III	Global	74.6	9.4	16	100	74	36	22.1
REFLECT	Lenvatinib (bridge)	478	476	OL	III	Global	50.2	22.7	27.1	99	70	43	27.5
LEAP-002	Pembrolizumab + lenvatinib†	395	399	DB	III	Global	49	23	38.2	99	67	31.6	32.1
IMbrave150	Atezolizumab + bevacizumab	336	165	OL	III	Global	47.9	21.6	30.5	99	75	37	15.6
COSMIC-312	Atezolizumab + cabozantinib	432	217	OL	III	Global	29.4	31	39	100	69	35.1	22.1

OL, open-label; DB, double-blind; TW, Taiwan; SG, Singapore; f/u, follow-up; CP-A, Child–Pugh A; MVI/EHS, macrovascular invasion or extrahepatic spread; AFP, alpha-fetoprotein. *Bevacizumab biosimilar. †Lenvatinib comparator (LEAP-002); lenvatinib-vs-sorafenib bridge (REFLECT).

Supplementary Table S2. Progression-free survival and objective response rate versus sorafenib

Treatment	Class	PFS — HR (95% CrI)	ORR — OR (95% CrI)
Finotonlimab + bevacizumab	ICI + anti-VEGF	0.50 (0.28–0.89) *	9.51 (2.47–36.74) *
Anlotinib + penpulimab	ICI + TKI	0.52 (0.29–0.94) *	3.44 (1.03–11.25) *
Camrelizumab + rivoceranib	ICI + TKI	0.54 (0.30–0.97) *	5.72 (1.74–18.56) *
Pembrolizumab + lenvatinib	ICI + TKI	0.55 (0.24–1.25)	5.05 (1.02–25.29) *
Sintilimab + bevacizumab	ICI + anti-VEGF	0.56 (0.31–1.01)	5.61 (1.52–20.50) *
Lenvatinib	TKI monotherapy	0.64 (0.36–1.13)	3.07 (0.97–9.65)
Atezolizumab + bevacizumab	ICI + anti-VEGF	0.65 (0.36–1.15)	3.25 (0.96–10.53)
Toripalimab + bevacizumab	ICI + anti-VEGF	0.69 (0.38–1.25)	4.99 (1.44–17.86) *
Atezolizumab + cabozantinib	ICI + TKI	0.74 (0.40–1.36)	3.02 (0.82–11.13)

Posterior median hazard ratios (PFS) and odds ratios (ORR) with 95% credible intervals. PFS values below 1 favour the experimental regimen; ORR values above 1 favour it. Asterisks denote credible intervals excluding unity. $\tau = 0.14$ (PFS), 0.31 (ORR).

Supplementary Table S3. Sensitivity analyses of the overall-survival class comparison

Analysis	P(anti-VEGF >TKI)	Interpretation
Primary (Turner prior, all-cause mortality)	75.0%	Reference result
Fixed-effect model	85.5%	Higher when heterogeneity is assumed absent
Alternative prior: pharmacological vs placebo	74.0%	Insensitive to choice of prior cell
Alternative prior: cause-specific mortality/composite	73.9%	Insensitive to outcome categorisation
High-HBV subset ($\geq 80\%$)	88.4%	Direction preserved; four trials only
High-HBV subset ($\geq 70\%$)	75.2%	Unchanged from primary
Leave-one-trial-out (range across nine)	59.6% – 85.2%	Above 50% for every omission
Seven-trial star, bridged evidence removed	78.2%	Independent of the lenvatinib bridge

P(anti-VEGF > TKI) is the posterior probability that the best ICI-plus-anti-VEGF-antibody regimen is superior to the best ICI-plus-TKI regimen for OS. Values above 50% favour the anti-VEGF-antibody class. Removing REFLECT also removes LEAP-002.

Supplementary Table S4. GRADE certainty of evidence for overall survival

Comparison vs sorafenib	OS HR (95% CrI)	Risk of bias	Indirectness	Imprecision	Publication bias	Certainty
Sintilimab + bevacizumab	0.57 (0.35–0.92)	Not serious	Serious‡	Serious§	Undetected	Low
Finotonlimab + bevacizumab	0.60 (0.37–0.99)	Not serious	Serious‡	Serious§	Undetected	Low
Camrelizumab + rivoceranib	0.64 (0.41–0.99)	Not serious	Serious‡	Serious§	Undetected	Low
Atezolizumab + bevacizumab	0.66 (0.42–1.04)	Not serious	Serious‡	Serious§	Undetected	Low
Anlotinib + penpulimab	0.69 (0.44–1.08)	Not serious	Serious‡	Serious§	Undetected	Low
Toripalimab + bevacizumab	0.76 (0.47–1.20)	Not serious	Serious‡	Serious§	Undetected	Low
Atezolizumab + cabozantinib	0.98 (0.62–1.54)	Not serious	Serious‡	Serious§	Undetected	Low
Pembrolizumab + lenvatinib	0.77 (0.42–1.40)	Not serious	Very serious¶	Very serious#	Undetected	Very low

■ Low certainty ■ Very low certainty

Certainty rated with GRADE for network meta-analysis. All comparisons derive from randomised trials and begin at high certainty. †Inconsistency (incoherence) cannot be assessed statistically because the network is a star with no closed loops; direct–indirect agreement is therefore not testable, and transitivity was assessed instead through the distribution of effect modifiers (Table 1). ‡Indirectness: each estimate rests on a single trial, and several are confined to China-only populations, limiting applicability. §Imprecision: wide 95% credible intervals, several crossing unity, reflecting single-trial evidence with empirically specified between-study heterogeneity. ¶Very serious indirectness: pembrolizumab-plus-levatinib is estimated only indirectly, through the REFLECT–LEAP-002 lenvatinib bridge, with no trial comparing it to sorafenib. #Very serious imprecision: the doubly indirect estimate carries the widest interval. All comparisons between active regimens are wholly indirect and were rated low or very low. CrI, credible interval; NA, not assessable.

Supplementary Table S5. League table of all pairwise overall-survival comparisons

Sint+Bev	0.95 (0.48–1.89)	0.89 (0.47–1.69)	0.86 (0.45–1.69)	0.83 (0.43–1.62)	0.75 (0.38–1.47)	0.74 (0.35–1.61)	0.62 (0.33–1.19)	0.58 (0.30–1.14)	0.57 (0.35–0.92)
	Fino+Bev	0.94 (0.48–1.85)	0.91 (0.47–1.79)	0.87 (0.45–1.70)	0.79 (0.40–1.56)	0.78 (0.36–1.69)	0.65 (0.34–1.25)	0.61 (0.31–1.20)	0.60 (0.37–0.99)
		Cam+Riv	0.97 (0.51–1.85)	0.93 (0.50–1.74)	0.84 (0.44–1.61)	0.83 (0.39–1.77)	0.69 (0.38–1.32)	0.65 (0.35–1.23)	0.64 (0.41–0.99)
			Atezo+Bev	0.96 (0.51–1.81)	0.87 (0.45–1.67)	0.85 (0.40–1.83)	0.72 (0.39–1.36)	0.67 (0.36–1.28)	0.66 (0.42–1.04)
				Anlo+Penp	0.91 (0.47–1.75)	0.89 (0.42–1.89)	0.75 (0.40–1.41)	0.70 (0.37–1.35)	0.69 (0.44–1.08)
					Tori+Bev	0.99 (0.46–2.13)	0.83 (0.43–1.57)	0.78 (0.41–1.49)	0.76 (0.47–1.20)
						Pembro+Lenv	0.84 (0.55–1.29)	0.79 (0.37–1.64)	0.77 (0.42–1.40)
							Lenv	0.94 (0.50–1.72)	0.92 (0.59–1.40)
								Atezo+Cab	0.98 (0.62–1.54)
									Soraf

Posterior median hazard ratios (95% credible intervals) from the Bayesian random-effects network meta-analysis for the row treatment versus the column treatment; values below 1 favour the row treatment. The diagonal gives regimen abbreviations, ordered by overall-survival SUCRA (most to least favourable). The upper triangle is shown; the lower triangle is the reciprocal. Bold indicates a credible interval excluding unity. Blue shading favours the row regimen, red the column regimen. Estimates versus sorafenib (final column) match Figure 2 of the main text.

Supplementary Table S6. Risk-of-bias assessment (Cochrane RoB 2)

Trial	D1	D2	D3	D4	D5	Overall
IMbrave150	Low	Low	Low	Low	Low	Low
CARES-310	Low	Low	Low	Low	Low	Low
COSMIC-312	Low	Low	Low	Low	Low	Low
LEAP-002	Low	Low	Low	Low	Low	Low
REFLECT	Low	Low	Low	Low	Low	Low
ORIENT-32	Low	Some concerns	Low	Some concerns	Low	Some concerns
HEPATORCH	Low	Some concerns	Low	Some concerns	Some concerns	Some concerns
Finotonlimab + bevacizumab	Low	Some concerns	Low	Some concerns	Low	Some concerns
APOLLO	Low	Some concerns	Low	Some concerns	Low	Some concerns

■ Low risk ■ Some concerns

RoB 2 domains: D1, randomisation process; D2, deviations from intended interventions; D3, missing outcome data; D4, measurement of the outcome; D5, selection of the reported result. Judgements apply to overall survival unless a progression-based concern is noted. Some concerns in the four China-only, open-label trials reflect unconfirmed blinded independent central review for progression endpoints (D2/D4) and, for HEPATORCH, reporting after a pre-planned interim analysis (D5). Overall survival, an objective endpoint, was at low risk in all trials.

Supplementary Table S7. Effect-modifier correlations and meta-regression

A. Pairwise correlations (Pearson r)

	HBV	MVI/EHS	AFP ≥ 400	China
HBV prevalence	1.00	0.90	0.78	0.84
MVI/EHS	0.90	1.00	0.90	0.89
AFP ≥ 400	0.78	0.90	1.00	0.95
China-only	0.84	0.89	0.95	1.00

B. Univariable meta-regression on overall-survival log-hazard ratio

Modifier	Slope (log-HR)	p value
HBV prevalence (per 10%)	-0.068	0.01

MVI/EHS proportion	—	0.03
AFP ≥400 proportion	—	0.21

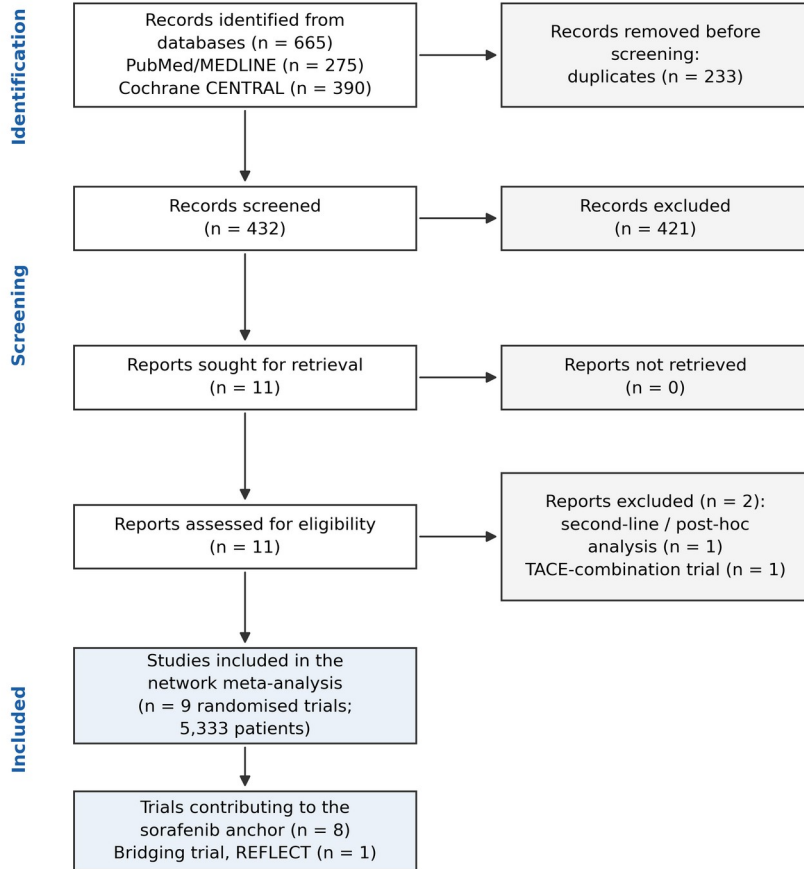
Univariable random-effects meta-regression of the trial-level overall-survival log-hazard ratio on each candidate effect modifier (seven trials with a direct sorafenib comparison). The three modifiers are strongly intercorrelated (Panel A) and each is strongly correlated with China-only trial status, so their individual associations with treatment effect cannot be separated from one another or from drug class and geography. They are reported as descriptive associations, not independent effect modifiers. MVI/EHS, macrovascular invasion or extrahepatic spread; AFP, alpha-fetoprotein; HBV, hepatitis B virus.

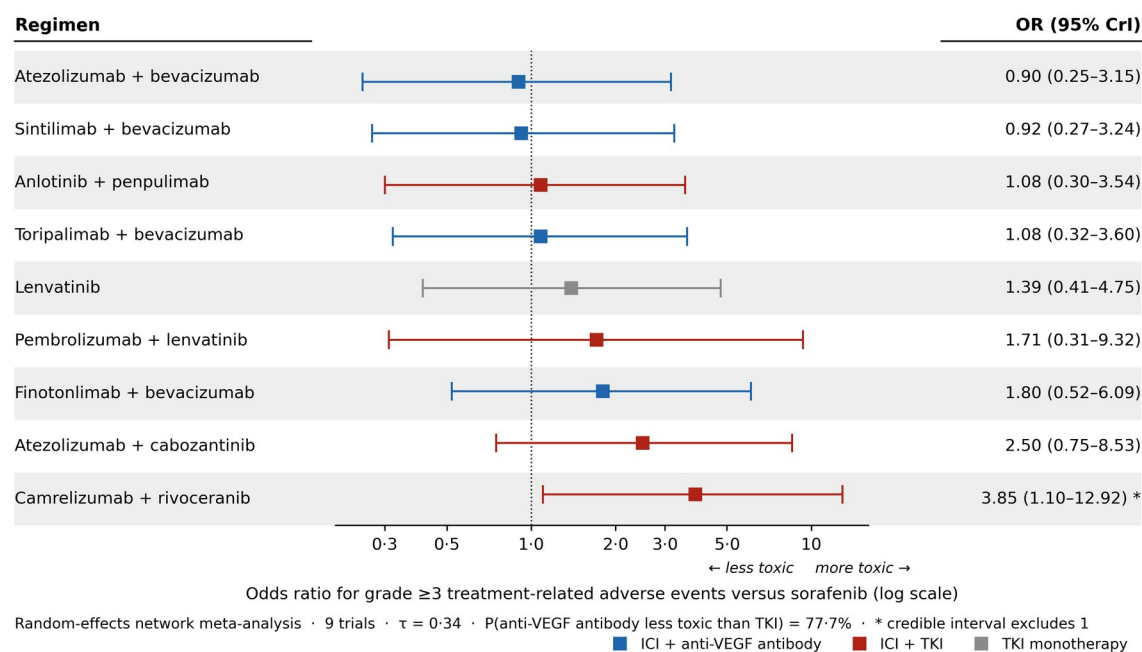
Supplementary Table S8. GRADE-NMA certainty of evidence for the class-level indirect comparison (ICI plus anti-VEGF antibody versus ICI plus VEGFR-TKI)

Endpoint	P(anti-VEGF > TKI)	Study limitations	Indirectness	Imprecision	Publication bias	Certainty
Overall survival	75.0%	Serious*	Very serious†	Serious‡	Undetected§	Very low
Progression-free survival	45.3%	Serious*	Very serious†	Very serious¶	Undetected§	Very low
Objective response	74.8%	Serious*	Very serious†	Serious‡	Undetected§	Very low
Grade ≥3 adverse events	77.7%	Serious*	Very serious†	Serious‡	Undetected§	Very low

Certainty rated with GRADE for network meta-analysis for the class-level contrast itself, rather than for individual regimens against sorafenib (Supplementary Table S4). All evidence derives from randomised trials and begins at high certainty. P(anti-VEGF > TKI) is the posterior probability that the best ICI-plus-anti-VEGF-antibody regimen is superior to the best ICI-plus-VEGFR-TKI regimen for the stated endpoint. *Study limitations: four of nine trials raised some concerns on RoB 2, and most trials were open-label, so progression-based endpoints, response assessment and adverse-event ascertainment are susceptible to assessment bias where blinded central review could not be confirmed. †Indirectness downgraded two levels on two independent grounds. First, the contrast is wholly indirect: no trial randomised patients between the two antiangiogenic backbones. Second, the trials contributing to the two classes differ systematically in characteristics that plausibly modify benefit — hepatitis B virus prevalence, macrovascular invasion or extrahepatic spread, and China-only enrolment are collinear with treatment class (pairwise $r = 0.81$ – 0.91 ; HBV–China $r = 0.87$; Supplementary Table S7) — so a contrast between the two classes cannot be separated from a contrast between the populations in which each class was tested. This confounding is not resolvable by reweighting, by alternative prior specification or by meta-regression on aggregate data. ‡Imprecision: the posterior probability does not exclude the absence of a class difference. ¶Very serious imprecision: a posterior probability of 45.3% is consistent with a difference in either direction. §Publication bias: with nine trials, formal assessment of small-study effects was not possible; no completed trial meeting the eligibility criteria was identified in ClinicalTrials.gov without a corresponding publication. Incoherence is not estimable: a star network with a single appended bridge contains no closed loops, so node-splitting and the Bucher method are structurally inapplicable; transitivity was assessed instead through the distribution of effect modifiers (Table 1). Between-study heterogeneity is not identifiable from these data and was specified from external empirical predictive distributions, so the imprecision assessments above are themselves conditional on that prior. ICI, immune checkpoint inhibitor; TKI, tyrosine kinase inhibitor; VEGF, vascular endothelial growth factor; VEGFR, VEGF receptor.

Supplementary Figure S1. PRISMA-NMA flow diagram.





Supplementary Figure S2. Grade ≥ 3 treatment-related adverse events: odds ratios versus sorafenib. Posterior median odds ratios (squares) with 95% credible intervals (horizontal lines) from the random-effects network meta-analysis, ordered from least to most toxic. Colour denotes treatment class. Odds ratios above 1 indicate greater toxicity than sorafenib. The asterisk marks the single regimen whose credible interval excluded unity (camrelizumab plus rivoceranib). Random-effects model, nine trials, $\tau = 0.34$; the posterior probability that the best-tolerated anti-VEGF-antibody regimen was less toxic than the best-tolerated TKI regimen was 77.7%.