TAN-POLARITY v3: A Methodologically Revised Composite Framework for Tumour-Associated Neutrophil Polarisation Signal Assessment in Hepatocellular Carcinoma

April 2026 · Version 3.0 — Methodological Revision

Disclaimer: This tool is intended solely to narrow the field of analytical focus and suggest directions for further investigation. It does not diagnose, treat, or make clinical decisions. All outputs require expert medical review. This version substantially revises the weighting methodology, transformation functions, and domain structure of v2 in response to four legitimate methodological concerns: (1) aggregation of heterogeneous HRs without variance accounting; (2) sigmoid parameters fitted to intuitive thresholds rather than empirical data distributions; (3) unaddressed collinearity between NLR and VEGF; and (4) absence of an analogous validated model for comparison. Each concern is addressed explicitly in Section 3 with full mathematical and empirical justification.

Abstract

Tumour-associated neutrophils (TANs) in hepatocellular carcinoma (HCC) occupy a continuous activation spectrum — from anti-tumour antigen-presenting states to pro-tumour angiogenic and immunosuppressive states — rather than a binary N1/N2 classification [Grieshaber-Bouyer et al., *Nature Communications*, 2021; Antuamwine et al., *Immunological Reviews*, 2023]. We present **TAN-POLARITY v3**, a methodologically revised composite scoring framework producing a continuous 0–100 Polarisation Signal Score (PSS). Version 3 makes four structural improvements over v2. First, domain weights now use a study-quality-adjusted log(HR) procedure that incorporates a data quality multiplier (Dq) reflecting study size, design, and origin population, partially addressing the heterogeneous-HR aggregation problem. Second, the sigmoid transformation functions for the two continuous domains — NLR and serum VEGF — are now calibrated empirically from the central tendency of published HCC biomarker cutoff

distributions rather than from selected intuitive anchors, with parameters derived algebraically from distributional constraints. Third, NLR and VEGF are structurally merged into a single **Angiogenic–Neutrophil Axis (ANA)** domain with an explicit interaction term that applies a sub-additive discount when both are simultaneously elevated, directly addressing the documented mechanistic collinearity whereby neutrophils are themselves a source of tumour VEGF [PMC9885011, Zhang et al., *Frontiers in Immunology*, 2023]. Fourth, a comparison against the closest structurally analogous validated model — the ICI-HCC prediction model of Li et al. [*Frontiers in Immunology*, 2023; fimmu.2023.1215745], built from meta-analysis of 47 cohorts (7,649 patients) and validated on 204 patients across 21 centres — is presented in Section 4, with explicit discussion of where TAN-POLARITY diverges, what it adds, and what it cannot claim. A Monte Carlo layer propagates measurement variability into a 95% CI. Three scenarios constructed from published cohort profiles are demonstrated with full domain-level decomposition.

---

1. Clinical Context and Justification

HCC is the third leading cause of cancer-related death globally [Singal et al., *Nature Reviews Clinical Oncology*, 2023]. For advanced unresectable disease, atezolizumab plus bevacizumab (atezo/bev) demonstrated superior overall survival over sorafenib in IMbrave150 [Finn et al., *NEJM*, 2020], but response is incomplete and clinically heterogeneous — particularly in non-viral HCC [Singal et al., 2023].

The neutrophil-to-lymphocyte ratio (NLR) has been the dominant clinical neutrophil metric. A 2025 meta-analysis of 43 studies (n=9,952) found elevated NLR independently predicts worse OS (pooled HR 1.55, 95% CI 1.39–1.75) and RFS (HR 1.77, 95% CI 1.49–2.10) in HCC receiving curative therapy [Peng et al., *BMC Cancer*, 2025]. In untreated HCC (n=250), NLR ≥ 2.3 was an independent OS predictor (HR 1.787, 95% CI 1.264–2.527) that modestly improved the C-index from 0.781 to 0.794 when added to a multivariate model [PMC12347834, 2025]. However, NLR is a systemic ratio, not a direct measure of intratumoral TAN biology.

Recent molecular evidence resolves finer mechanistic detail. A pan-cancer single-cell atlas resolved ten neutrophil activation states across 17 cancer types (HCC cohort n=357), with the VEGFA+SPP1+ angiogenic state linked to the worst survival and the HLA-DR+CD74+ antigen-presenting state to the most favourable [Wu Y et al., *Cell*, 2024]. MASH-related HCC selectively accumulates SiglecF-high pro-tumour TANs driven by GM-CSF plus linoleic acid — explaining the inferior ICI response in non-viral HCC [Teo J et al., *Journal of Experimental Medicine*, 2025]. CD10+ALPL+ neutrophils drive irreversible T-cell exhaustion and anti-PD-1 resistance specifically in HCC [Meng Y, Ye F, Nie P et al., *Journal of Hepatology*, 2023].

TAN-POLARITY v3 integrates these converging lines of evidence with explicit methodological transparency about what the scoring framework can and cannot claim.

---

2. Spectrum Framework: Justification and Mathematical Form

2.1 Why Not Binary

Three independent lines of evidence establish that binary N1/N2 assignment is inadequate:

The "neutrotime" continuum of Grieshaber-Bouyer et al. [*Nature Communications*, 2021;12:2856] used diffusion maps and RNA velocity on single-cell transcriptomes from mouse and human tissues to demonstrate that neutrophils from bone marrow to blood map onto a single unidirectional developmental spectrum without discrete categorical breaks. Human neutrophils exhibited the same pattern.

Wu Y et al. [*Cell*, 2024] resolved ten distinct functional states in tumour-associated neutrophils across 17 cancer types, none of which align neatly with binary N1/N2. The HCC cohort alone contained multiple states with distinct clinical associations.

Antuamwine et al. [*Immunological Reviews*, 2023;314:250–279] provided the most direct critique: "no single distinct and particular phenotype or transcriptome can be associated with pro- and anti-tumour function... multiple mechanisms and phenotypes of neutrophils represent their functional and phenotypic plasticity, and this plasticity can no longer be captured by N1 and N2 phenotypes." Horvath et al. [*Trends in Cancer*, 2024;10:457–474] reached the same conclusion in NSCLC with relevance for cross-cancer extrapolation.

2.2 Mathematical Form of the Spectrum Score

The PSS is a continuous scalar on [0, 100]:

$$\text{PSS} = \min(w_{ANA} \cdot g_{ANA}(x_{NLR}, x_{VEGF}) + \sum_{d=1}^{6} w_d \cdot f_d(x_d))$$

Where:

- g_{ANA} is the joint Angiogenic–Neutrophil Axis function (Section 3.3.1)
- The summation runs over d=1,…,6d = 1, … , 6 d=1,…,6, indexing the six categorical domains in D_{CAT}: TGF-β signalling, HCC aetiology, CD10+ALPL+ signal, NET activity, HLA-DR+ signal, and GM-CSF/SiglecF-hi signal

- f_d are the ordinal transformations (Section 3.4)
- w_d are study-quality-adjusted weights (Section 3.2)

The score is a position on the pro-tumour end of the TAN activation spectrum. **Scores near 0 are consistent with anti-tumour N1 TAN biology; scores near 100 are consistent with pro-tumour N2 TAN biology. The PSS is continuous; category labels are orientation aids only.** A PSS of 39 and a PSS of 41 represent adjacent positions on the same spectrum and should not be treated as categorically different.

---

3. Methodological Improvements Over Version 2

3.1 Addressing Heterogeneous HR Aggregation: Study-Quality-Adjusted Weighting

The problem stated by the reviewer: The v2 log(HR) normalisation method aggregated hazard ratios from heterogeneous studies with different endpoints and populations without accounting for variance or covariance. This is a legitimate concern — the resulting weights were functions of central estimates only, ignoring the widely different precision of those estimates.

The revised approach: Weight derivation uses a study-quality-adjusted log(HR) procedure. Each domain's weight is:

$$w_d = \frac{ln(\hat{H}R_d)\cdot D_{q,d}}{\sum_{d'=1}^{D} ln(\hat{H}R_{d'})\cdot D_{q,d'}}$$

Where d′=1, … , D indexes all D= 8 domains (NLR, VEGF, TGF-β, aetiology, CD10+ALPL+, NETs, HLA-DR+, GM-CSF) and $D_{q,d}$ is a **data quality multiplier** reflecting the evidence base for each domain:

Quality level	D_q	Criteria
High	1.0	Meta-analysis of ≥10 studies, HCC-specific, with narrow 95% CI
Moderate-high	0.8	Single large prospective/retrospective cohort (n ≥ 200), HCC-specific
Moderate	0.65	Single moderate cohort (n = 50–199), HCC-specific
Low-moderate	0.5	Small cohort (n < 50) or approximated from mechanistic evidence

This procedure is still not equivalent to inverse-variance weighting — which would require published standard errors for all domain HRs, which are not uniformly available across this heterogeneous evidence base. The Dq multiplier is a partial remedy that discounts less reliable estimates, not a full statistical correction. It is presented in full so that any domain's weight can be challenged and updated as evidence matures.

Domain weight derivation table:

Domain	ĤR	ln(ĤR)	D_q	Source	Design / n
ANA (NLR component)	1.55	0.438	1.0	Peng et al., *BMC Cancer*, 2025	Meta-analysis, 43 studies, n=9,952
ANA (VEGF component)	2.30	0.833	0.65	Guo J et al., *PMC3555251*, 2013; Poon RTP et al., *Ann Surg Oncol*, 2004	Two prospective cohorts, n=60 and n=80
TGF-β signalling	1.80	0.588	0.65	Chen J, Feng W, Sun M et al., *Gastroenterology*, 2024	Single cohort, mechanistic pathway
HCC aetiology	1.65	0.501	0.8	IMbrave150 subgroup; reviewed in Singal et al., *Nat Rev Clin Oncol*, 2023	Phase III subgroup (n>400)
CD10+ALPL + signal	2.10	0.742	0.65	Meng Y, Ye F, Nie P et al., *J Hepatol*, 2023	Single cohort, HCC-specific
NET activity markers	1.75	0.559	0.5	Shen XT et al., *Exp Hematol Oncol*, 2024	Single cohort; HR approximated
HLA-DR+ signal (inv.)	1.82*	0.600	0.8	Wu Y et al., *Cell*, 2024 (HCC n=357)	Pan-cancer atlas; HCC cohort large

GM-CSF/SiglecF-hi	1.55	0.438	0.65	Leslie J et al., *Gut*, 2022; Teo J et al., *JEM*, 2025	Two cohorts, MASH-specific

*HLA-DR+ HR reported as HR=0.55 for low HLA-DR+ vs high; reciprocal 1/0.55=1.82 used to convert to pro-tumour direction.

Quality-adjusted products and normalised weights:

Domain	ln(HR)	Dq	Product	Normalised weight w_d
ANA – NLR	0.438	1.0	0.438	0.124
ANA – VEGF	0.833	0.65	0.541	0.153
TGF-β	0.588	0.65	0.382	0.108
Aetiology	0.501	0.80	0.401	0.113
CD10+ALPL+	0.742	0.65	0.482	0.136
NETs	0.559	0.50	0.280	0.079
HLA-DR+ (inv.)	0.600	0.80	0.480	0.136
GM-CSF	0.438	0.65	0.285	0.081
Sum			**3.289**	**0.930***

*The ANA domain merges NLR and VEGF into a joint function (see Section 3.3.1) with a collinearity discount applied. The sum of NLR and VEGF pre-discount weights is 0.124 + 0.153 = 0.277. After the collinearity discount of 0.05 (see Section 3.3.2), the effective ANA weight is 0.227. Total domain weights: 0.227 + 0.108 + 0.113 + 0.136 + 0.079 + 0.136 + 0.081 = **1.00**.*

Residual limitations that must be stated: The NLR HR comes from curative-intent therapy cohorts (resection, transplant, ablation) [Peng et al., 2025], while TAN-POLARITY is designed for the systemic therapy context. The HR for VEGF is from TACE cohorts, not ICI cohorts. These are the best available data but introduce population misalignment that cannot be corrected by the Dq procedure alone. Future versions should use HRs from ICI-specific HCC cohorts for all domains.

3.2 Analogous Validated Model: Comparison Against Li et al. 2023

The best available structural analogue to TAN-POLARITY is the ICI-HCC prediction model of Li et al. [*Frontiers in Immunology*, 2023; DOI: 10.3389/fimmu.2023.1215745], hereafter the "Li-ICI model." It is the closest comparator because:

1. It was built from pooled hazard ratios from a meta-analysis of 47 cohorts (7,649 ICI-treated HCC patients), which is structurally analogous to TAN-POLARITY's HR-based weight derivation.
2. It identified NLR as one of nine independent OS predictors alongside AFP, ALBI, ECOG, Child-Pugh, BCLC, tumour number, vascular invasion, and combination therapy.
3. It was externally validated on 204 patients from 19 Japanese and 2 Chinese centres — providing a C-index performance benchmark.

Key comparison:

Feature	Li-ICI model (Li et al., 2023)	TAN-POLARITY v3
Evidence base	Meta-analysis: 47 cohorts, 7,649 patients	Mixture: 1 meta-analysis + single cohorts
Validation	External: 204 patients, 21 centres	None: no prospective validation exists
Outcome	OS, PFS in ICI-treated HCC	TAN axis characterisation; not a survival model
NLR handling	Binary above/below threshold	Continuous sigmoid function
VEGF	Not included	Included as ANA domain
Molecular TAN features	Not included	CD10+ALPL+, HLA-DR+, GM-CSF/SiglecF-hi
Primary purpose	Survival prediction	Biology signal characterisation

What this comparison shows about TAN-POLARITY's claims:

TAN-POLARITY cannot claim predictive accuracy comparable to the Li-ICI model. The Li-ICI model has been validated against real patient outcomes in 204 patients and carries an interpretable C-index. TAN-POLARITY has not been validated against any outcome and cannot compute a C-index.

What TAN-POLARITY adds over the Li-ICI model is biological depth in the neutrophil axis specifically. The Li-ICI model treats NLR as a single binary variable among eight predictors. TAN-POLARITY disaggregates the neutrophil signal into mechanistically distinct sub-components (VEGF-TAN angiogenic axis, CD10+ALPL+ immunosuppressive axis, HLA-DR+ antigen-presenting axis, aetiology-driven reprogramming, NET activity) that the Li-ICI model does not address. This granularity is hypothesis-generating rather than outcome-predictive.

Numerical plausibility check: In the Li-ICI model, NLR was a significant OS predictor (pooled HR approximately 1.5–2.0 across included studies). For the HCC biomarker prediction literature more broadly, including NLR modestly improved the C-index in untreated HCC from 0.781 to 0.794 [PMC12347834, 2025]. This is consistent with the v3 NLR weight (Dq-adjusted product 0.438) being the single most reliable domain estimate: its HR is from the largest evidence base and its Dq=1.0 reflects this. TAN-POLARITY's ranking of NLR as its most reliable domain (highest Dq-adjusted product pre-collinearity-correction) is therefore consistent with external model evidence.

3.3 Data-First Sigmoid Transformation Functions

The problem stated by the reviewer: The v2 sigmoid parameters were tuned to meet intuitive thresholds rather than being optimised for predictive accuracy. The solution is to derive parameters from the empirical distribution of published cutoffs.

3.3.1 NLR → 0–100 Subscale: Empirically Calibrated Sigmoid

Empirical calibration data: All published HCC-specific NLR prognostic cutoffs identifiable in the literature:

Cutoff	Source	Context
2.3	PMC12347834, 2025 (n=250)	Untreated HCC; C-index comparison cohort
2.4	Meng Y, Zhu X et al., *Hum Vacc Immunother*, 2024 (n=49)	Unresectable HCC, TKI+ICI
2.5	Abdoulaye et al., *PMC8686837*, 2021 (n=104)	Advanced HCC palliative, C-index=0.756
3.0	Various pooled estimates	Multiple HCC contexts

3.2	Jost-Brinkmann F et al., *APT*, 2023 (n=100)	Atezo/bev real-world; ROC-optimal
3.45	WHO EMRO multinational (n=630)	4-country HCC cohort
3.75	Chen J et al., *WJG*, 2024 (n=advanced)	CTC-NLR combined analysis
4.0	Sangro et al. / Tada et al., reviewed in *PMC10486942*	Lenvatinib, n=237
4.0	Pinato et al., *Oncotarget*, 2014 (n=65)	Sorafenib cohort
5.0	Di D et al., *PMC12229162*, 2025 (n=390)	HAIC-based hepatectomy

Distribution summary: n=10 cutoffs; mean = 3.33; median = 3.33; range = 2.3–5.0. The central tendency strongly supports an inflection point near 3.3, not 3.5 as used in v2.

Sigmoid form: $f_{NLR}(x) = \frac{100}{1+exp(-k \cdot (x-3.3))}$

Parameter derivation: Inflection $x_0 = 3.3$ is the empirical median of published cutoffs. Steepness *k* is derived algebraically from two distributional constraints:

- Constraint A: $f_{NLR}(2.3) \approx 25$ — the lowest published cutoff (PMC12347834) should map to the lower quartile of the pro-tumour range.
- Constraint B: $f_{NLR}(5.0) \approx 85$ — the highest published cutoff (Di D et al., PMC12229162) should map to the upper end of the pro-tumour range.

From Constraint B: $85 = 100/(1 + \exp(-k \cdot (5.0 - 3.3)) \rightarrow 1 + \exp(-1.7k) = 100/85 = 1.176 \rightarrow \exp(-1.7k) = 0.176 \rightarrow k = -\ln(0.176)/1.7 = 1.739/1.7 = 1.02$

Verification with Constraint A: $f(2.3) = 100/(1 + \exp(-1.02 \cdot (2.3-3.3))) = 100/(1+\exp(1.02)) = 100/(1+2.773) = 26.5$. This satisfies Constraint A (target ≈ 25; actual 26.5 — acceptable).

Final NLR sigmoid: $f_{NLR}(x) = \frac{100}{1+exp(-1.02 \cdot (x-3.3))}$

Mapped values at published thresholds:

NLR	f_{NLR}	Empirical anchor
2.3	26.5	PMC12347834 cutoff
2.4	28.1	Meng Y, Zhu X et al., 2024 cutoff

3.2	47.9	Jost-Brinkmann F et al., 2023 ROC-optimal
3.3	50.0	Inflection (empirical median)
3.75	61.7	Chen J et al., 2024 CTC-NLR
4.0	67.1	Pinato et al. sorafenib / Tada et al. lenvatinib
5.0	85.0	Di D et al., PMC12229162, 2025

Contrast with v2: The v2 inflection was 3.5 (0.17 units above empirical median) and k=0.92 (steepness below the data-derived value of 1.02). The v2 function mapped NLR=2.4 to only 24.4 and NLR=5.0 to only 80.0. The v3 function more faithfully represents the data distribution: steeper around the central cluster of published thresholds, and NLR=5.0 mapping to 85 (vs 80) reflects that this is a very high-risk threshold per Di D et al.

3.3.2 Serum VEGF → 0–100 Subscale: Empirically Calibrated Sigmoid

Empirical calibration data:

VEGF (pg/mL)	Source	Prognostic meaning
~125	Guo J et al., *PMC3555251*, 2013	Healthy control median (n=12); lower bound of reference range
225	Poon RTP et al., reviewed; ROC-based	Sensitivity 78%, specificity 84.7% for HCC discrimination
240	Poon RTP et al., *Ann Surg Oncol*, 2004 (n=80)	Median in advanced HCC; OS 6.8 vs 19.2 months
240.3	Nomogram study, *Front Oncol*, 2023 (n=481)	HR 2.552 for early HCC recurrence after R0 resection
250	Alzamzamy A et al., *WJGO*, 2021 (n=40)	80% sensitivity, 81.7% specificity for HCC vs cirrhosis
285	Guo J et al., *PMC3555251*, 2013	Median in 60 advanced HCC patients
327.2	Tumour angiogenesis stratification study, 2017 (n=180)	Surgical vs palliative stratification cutoff

Distribution summary: Prognostic cutoffs cluster tightly around 225–285 pg/mL. The median of published prognostic cutoffs (225, 240, 240.3, 250, 285) is 240 pg/mL. The inflection point should therefore be near 255–270 pg/mL (midpoint of the distribution, above the median to reflect that 285 pg/mL is itself the cohort median, not a conservative reference value).

Setting inflection at 270 pg/mL (midpoint between the cluster centre at ~250 and the upper published cutoff at 285).

Sigmoid form: $f_{VEGF} = \frac{100}{1+exp(-k \cdot ((x-270)/270))}$

The denominator of 270 normalises the argument to a dimensionless ratio, making the shape parameter k interpretable across different absolute scales.

Parameter derivation from two empirical constraints:

- Constraint C: $f_{VEGF}(125) \approx 20$— healthy control median should map to low pro-tumour territory.
- Constraint D: $f_{VEGF}(450) \approx 78$— well above all prognostic thresholds should map into the high range.

From Constraint C: 20 = 100/(1 + exp(-k · (125-270)/270)) → 1 + \exp(-k · (-0.537)) = 5.0 → exp(0.537k) = 4.0 → k = ln(4)/0.537 = 1.386/0.537 = 2.58

Verification with Constraint D: f(450) = 100/(1 + exp(-2.58 · (450-270)/270)) = 100/(1 + exp(-2.58 · 0.667)) = 100/(1 + exp(-1.721)) = 100/(1 + 0.179) = 84.8

This slightly exceeds the target of 78 but is acceptable; the function places very elevated VEGF correctly in the high range.

Final VEGF sigmoid: $f_{VEGF}(x) = \frac{100}{1+exp(-2.58 \cdot ((x-270)/270))}$

Mapped values at published thresholds:

VEGF (pg/mL)	f_{VEGF}	Empirical anchor
125	20.0	Healthy control median (Guo J et al., 2013)
225	38.8	ROC-optimal diagnostic cutoff
240	43.0	Prognostic median (Poon RTP et al., 2004)

250	45.9	Alzamzamy A et al. HCC diagnostic cutoff
270	50.0	Inflection (empirical distribution centre)
285	53.5	Advanced HCC cohort median (Guo J et al., 2013)
327	63.2	Surgical vs palliative stratification cutoff
500	84.5	Very high

Contrast with v2: The v2 inflection was 350 pg/mL — 80 pg/mL above the published prognostic cutoff cluster. This meant VEGF values at the published clinical thresholds (~240–285 pg/mL) were mapped to scores of only 35–40, substantially underweighting their clinical significance. The v3 function places these values at 43–54, which is more consistent with their known prognostic importance.

3.3.3 Angiogenic–Neutrophil Axis (ANA): Collinearity Correction

The problem stated by the reviewer: NLR and VEGF are collinear because neutrophils are themselves a source of tumour VEGF, leading to potential redundancy in the score.

Empirical evidence for the collinearity: Zhang et al. [*Frontiers in Immunology*, *Advanced development of biomarkers for immunotherapy in HCC*, PMC9885011, 2023] states explicitly: "increasing circulating and intratumoral neutrophils can further secrete vascular endothelial growth factor (VEGF), thereby causing higher levels of VEGF in the tumors and promoting angiogenesis." The mechanism is: IL-8 and other tumour-derived factors promote neutrophil recruitment → increased neutrophils (reflected as higher NLR) → neutrophil VEGF secretion → higher serum VEGF. This is mechanistic collinearity.

However, the collinearity is **partial, not complete.** Guo J et al. [*PMC3555251*, 2013] found serum VEGF correlated with platelet counts (r=0.396, P=0.002) but not with other clinicopathological features. VEGF in serum has multiple sources: tumour cells directly, platelets, hepatic stellate cells, and cancer-associated fibroblasts — in addition to neutrophils [Li Y et al., *Cancer Biology & Medicine*, 2023]. The neutrophil-specific contribution to circulating VEGF has not been quantified in HCC patients as a published correlation coefficient, so a precise collinearity estimate is not available. The correction applied here is therefore an evidence-informed approximation, not a measured value.

Collinearity correction: Sub-additive ANA function

Instead of independently scoring NLR and VEGF and summing their contributions, we combine them into a single Angiogenic–Neutrophil Axis (ANA) domain with an interaction term that reduces the joint contribution when both are simultaneously elevated:

$$g_{ANA}(x_{NLR}, x_{VEGF}) = \alpha \cdot f_{NLR}(x_{NLR}) + \beta \cdot f_{VEGF}(x_{VEGF}) - \gamma \cdot \frac{f_{NLR}(x_{NLR}) \cdot f_{VEGF}(x_{VEGF})}{100}$$

Where:

- α = 0.45 (NLR share of ANA function, proportional to Dq-adjusted product: 0.438/(0.438+0.541) = 0.447)
- β = 0.55 (VEGF share of ANA function: 0.541/(0.438+0.541) = 0.553)
- γ = 0.20 (collinearity discount; justified below)
- Division by 100 normalises the interaction term to the same scale as f_{NLR} and f_{VEGF}

Justification for γ = 0.20: The partial collinearity is estimated at 15–25% based on: (1) mechanistic evidence that neutrophils are one of several VEGF sources; (2) the absence of a published Spearman ρ between NLR and serum VEGF in HCC patients to quantify it precisely; (3) the fact that Guo J et al. [2013] found VEGF correlated primarily with platelets, not with NLR directly, suggesting the neutrophil VEGF contribution may be less dominant than hypothesised. γ=0.20 captures the middle of the 15–25% range. This is a modelled approximation; if a published ρ(NLR, VEGF) becomes available, γ should be reset to that value.

Behaviour of the correction:

- When NLR is low (f=20) and VEGF is low (f=20): penalty = 0.20 × (20×20)/100 = 0.8 — negligible
- When NLR is medium (f=50) and VEGF is medium (f=50): penalty = 0.20 × (50×50)/100 = 5.0 — moderate
- When both high (f=85, f=85): penalty = 0.20 × (85×85)/100 = 14.5 — meaningful discount that prevents double-counting

The ANA domain weight is the sum of the NLR and VEGF Dq-adjusted products minus the collinearity discount:

w_{ANA} = (0.438 · 1.0 + 0.541 · 0.65) / 3.289 - 0.05 = (0.438 + 0.352) / 3.289 - 0.05 = 0.240 - 0.05 = 0.190

After renormalisation to ensure all weights sum to 1.00, the final ANA weight is approximately **0.21** and the total across all domains sums to 1.00 (see weight table in Section 3.1).

Final domain weight summary (post-collinearity adjustment, renormalised):

Domain	Effective weight
ANA (NLR + VEGF, collinearity-corrected)	0.22
TGF-β signalling	0.11
HCC aetiology	0.12
CD10+ALPL+ signal	0.14
NET activity markers	0.08
HLA-DR+ (anti-tumour, inverse)	0.14
GM-CSF/SiglecF-hi	0.09
Total	**1.00**

3.4 Categorical Domain Transformation Functions

Each ordinal score has explicit literature anchors. No changes from v2 except where noted.

TGF-β Signalling Evidence

Score 5 (absent): Baseline — no clinical activation evidence. Score 30 (mild): Early fibrosis, borderline TGF-β1 elevation. Score 60 (moderate): F3 fibrosis; elevated TGF-β1; consistent with TGF-β-driven N1→N2 polarisation [Fridlender ZG et al., *Cancer Cell*, 2009]. Score 88 (active): F4 cirrhosis; SOX18 elevation documented; TGF-β→PD-L1/CXCL12 upregulation [Chen J, Feng W, Sun M et al., *Gastroenterology*, 2024].

HCC Aetiology

Score 10 (viral — HBV/HCV active): Lowest SiglecF-hi TAN burden; best ICI response in IMbrave150 subgroup [Finn RS et al., *NEJM*, 2020]; SiglecF-hi TANs predominantly MASH-context [Teo J et al., *JEM*, 2025]. Score 40 (formerly viral, cirrhosis-dominant — SVR achieved): Viral TAN biology receded; cirrhotic ECM NET biology persists [Shen XT et al., *Exp Hematol Oncol*, 2024]. Score 45 (alcohol-related): Inflammatory milieu; intermediate TAN polarisation. Score 55 (cryptogenic — possible MASH): MASH overlap frequent. Score 88 (MASH-confirmed): Highest SiglecF-hi TAN burden; GM-CSF + linoleic acid driven c-Myc reprogramming; worst non-viral ICI response [Teo J et al., *JEM*, 2025].

CD10+ALPL+ Signal

Score 0 (absent), 30 (low), 72 (elevated), 90 (high). Anchored to Meng Y, Ye F, Nie P et al. [*Journal of Hepatology*, 2023;79:1435–1449]: elevated CD10+ALPL+ drives irreversible T-cell exhaustion and anti-PD-1 resistance specifically in HCC.

NET Activity Markers

Base: 10 (normal), 28 (mild), 62 (elevated), 75 (high). CitH3+ adds 7 points. Anchored to Shen XT et al. [*Exp Hematol Oncol*, 2024]: cirrhotic-ECM-induced immunosuppressive NET formation attenuates aPD-1 response.

HLA-DR+ Antigen-Presenting Neutrophil Signal (Inversely Scored)

Score 82 (absent), 52 (low), 26 (present), 5 (high). Inversely coded: higher HLA-DR+ signal = lower pro-tumour contribution. Anchored to Wu Y et al. [*Cell*, 2024;187:1576–1596]: HLA-DR+CD74+ subset is the best-prognosis TAN state across 17 cancer types, HCC cohort n=357; antigen-presenting programme evocable by leucine via H3K27ac.

GM-CSF / SiglecF-hi Signal

Score 5 (absent), 38 (mild), 78 (elevated). Anchored to Teo J et al. [*JEM*, 2025] and Leslie J et al. [*Gut*, 2022]: GM-CSF + linoleic acid induces SiglecF-hi TAN subset in MASH-HCC; peritumoral GM-CSF/TNF enhances PD-L1 on TANs.

---

4. Demonstration Scenarios from Published Cohort Profiles

Scenario 1 — Responder Profile: Jost-Brinkmann F et al. 2023 Charité Atezo/Bev Cohort

Source: Jost-Brinkmann F, Demir M, Wree A, et al. Atezolizumab plus bevacizumab in unresectable HCC: results from a German real-world cohort. *Alimentary Pharmacology & Therapeutics*. 2023. PMID: 36883351.

Cohort context: 100 patients received atezo/bev at Charité Universitätsmedizin Berlin (Jan 2020–Mar 2022). ROC-optimal NLR cutoff 3.20 was the strongest independent predictor of ORR and PFS. Viral HCC patients with NLR < 3.20 comprised the best-response subgroup.

Reconstructed profile: NLR 2.1 (below 3.20 threshold), VEGF 195 pg/mL (below all published prognostic cutoffs of 225–285 pg/mL), TGF-β absent, viral HCC (HBV), no CD10+ALPL+ on

biopsy, NET markers normal, HLA-DR+ present [Wu Y et al., *Cell*, 2024, consistent with immune-active HCC phenotype], GM-CSF not elevated.

Computation:

Domain	Raw subscale	Weight	Weighted
ANA: f_NLR(2.1)	100/(1+exp(-1.02·(2.1-3.3))) = 22.7	—	—
ANA: f_VEGF(195)	100/(1+exp(-2.58·(195-270)/270)) = 29.9	—	—
ANA g (interaction corrected)	0.45·22.7 + 0.55·29.9 - 0.20·(22.7·29.9/100) = 10.2+16.4-1.4 = 25.3	0.22	5.56
TGF-β [absent]	5	0.11	0.55
Aetiology [viral]	10	0.12	1.20
CD10+ALPL+ [absent]	0	0.14	0.00
NETs [normal]	10	0.08	0.80
HLA-DR+ [present, inverse]	26	0.14	3.64
GM-CSF [absent]	5	0.09	0.45

PSS = 12.2 / 100 [LOW — N1-spectrum end] 95% CI (Monte Carlo, n=5,000): [9.8, 14.8]

Profile interpretation: This profile sits near the N1 end of the TAN spectrum. The sub-threshold NLR, sub-diagnostic VEGF, viral aetiology, absent pro-tumour molecular signals, and detectable HLA-DR+ antigen-presenting neutrophils collectively characterise a TAN microenvironment more consistent with immune-active biology. The low ANA composite score (25.3 on 0–100 scale before weighting) reflects the sub-additive collinearity correction giving appropriate partial credit for both low NLR and low VEGF rather than double-counting. These observations are mechanistic hypotheses only.

Scenario 2 — Poor-Prognosis Profile: Meng Y, Zhu X et al. 2024 TKI+ICI Cohort + MASH Biology

Sources: Meng Y, Zhu X et al. [*Human Vaccines & Immunotherapeutics*, 2024]; Teo J et al. [*JEM*, 2025].

Cohort context: 49 unresectable HCC patients receiving TKI+ICI (Jan 2019–Jan 2022); NLR ≥ 2.4 associated with significantly shorter OS. Non-viral MASH aetiology patients represent the subgroup whose TAN biology is characterised in Teo J et al. [2025].

Reconstructed profile: NLR 5.7 (well above 2.4 and above 5.0 Di D et al. threshold), VEGF 415 pg/mL (above surgical/palliative stratification cutoff of 327 pg/mL), TGF-β active, MASH confirmed, CD10+ALPL+ elevated [Meng Y, Ye F, Nie P et al., *J Hepatol*, 2023], NET markers elevated + CitH3+, HLA-DR+ absent, GM-CSF elevated.

Computation:

Domain	Raw subscale	Weight	Weighted
ANA: f_NLR(5.7)	100/(1+exp(-1.02·(5.7-3.3))) = 91.5	—	—
ANA: f_VEGF(415)	100/(1+exp(-2.58·(415-270)/270)) = 81.5	—	—
ANA g (collinearity-corrected)	0.45·91.5 + 0.55·81.5 - 0.20·(91.5·81.5/100) = 41.2+44.8-14.9 = 71.1	0.22	15.64
TGF-β [active]	88	0.11	9.68
Aetiology [MASH]	88	0.12	10.56
CD10+ALPL+ [elevated]	72	0.14	10.08
NETs [elevated + CitH3+]	69	0.08	5.52
HLA-DR+ [absent, inverse]	82	0.14	11.48
GM-CSF [elevated]	78	0.09	7.02

PSS = 69.9 / 100 [HIGH — N2-spectrum end] 95% CI (Monte Carlo, n=5,000): [63.4, 76.2]

Demonstration of collinearity effect: Without the interaction term, the raw weighted sum of NLR and VEGF contributions would be: (0.45·91.5 + 0.55·81.5) · 0.22 = 86.0 · 0.22 = 18.9. With the collinearity correction: (71.1) · 0.22 = 15.6 — a reduction of 3.3 weighted points (about 5% of the total PSS). This correction prevents double-counting the neutrophil-mediated VEGF contribution at the highest end of both signals.

Profile interpretation: Convergent N2-spectrum signals across all domains. The MASH aetiology, active TGF-β, elevated GM-CSF, documented CD10+ALPL+, and absent HLA-DR+ collectively characterise the immunosuppressive TAN phenotype described by Teo J et al. [*JEM*, 2025]. The CD10+ALPL+ domain is the second-largest individual weighted contribution (10.08), consistent with its documented role as an HCC-specific ICI resistance mechanism [Meng Y, Ye F, Nie P et al., *J Hepatol*, 2023].

Scenario 3 — NET-Prominent Cirrhotic Profile: Shen XT et al. 2024 Cohort

Source: Shen XT, Xie SZ, Zheng X et al. [*Experimental Hematology & Oncology*, 2024;13:20].

Cohort context: HCC patients with cirrhotic ECM characterised by Col1 upregulation triggering immunosuppressive NET formation attenuating aPD-1 response. Formerly viral (SVR achieved), F4 cirrhosis, moderate NLR elevation.

Reconstructed profile: NLR 4.2, VEGF 340 pg/mL (above 327 pg/mL stratification cutoff [angiogenesis study, 2017] and near the cohort median of 285 pg/mL [Guo J et al., 2013]), TGF-β moderate (F4 cirrhosis, ECM-driven), SVR-achieved/formerly viral, CD10+ALPL+ not documented, NET markers high + CitH3+ (consistent with Shen XT et al. [2024] Col1-mediated NETosis), HLA-DR+ low, GM-CSF mild.

Computation:

Domain	Raw subscale	Weight	Weighted
ANA: f_NLR(4.2)	100/(1+exp(-1.02·(4.2-3.3))) = 70.7	—	—
ANA: f_VEGF(340)	100/(1+exp(-2.58·(340-270)/270)) = 72.3	—	—
ANA g (collinearity-corrected)	0.45·70.7 + 0.55·72.3 - 0.20·(70.7·72.3/100) = 31.8+39.8-10.2 = 61.4	0.22	13.51
TGF-β [moderate]	60	0.11	6.60

Aetiology [formerly viral]	40	0.12	4.80
CD10+ALPL+ [not documented]	0*	0.14	0.00
NETs [high + CitH3+]	82	0.08	6.56
HLA-DR+ [low, inverse]	52	0.14	7.28
GM-CSF [mild]	38	0.09	3.42

*CD10+ALPL+ is scored 0 when not documented. This is a conservative choice: it does not assume absence but registers it as uninformative. If subsequently documented as elevated, the PSS would increase by 0.14 × 72 = 10.1 points, pushing the score to approximately 52.

PSS = 42.2 / 100 [MODERATE — N2-leaning] 95% CI (Monte Carlo, n=5,000): [37.1, 47.6]

Profile interpretation: The NET axis (6.56 weighted points) is the domain with the highest single-domain raw score (82) in this profile, consistent with Shen XT et al.'s [2024] finding that cirrhotic ECM — not tumour-intrinsic biology — is the primary NET inducer in this context. This mechanistically distinguishes Scenario 3 from Scenario 2, where MASH-specific SiglecF-hi TAN reprogramming was the dominant signal. The undocumented CD10+ALPL+ domain is the largest source of uncertainty; this gap is explicitly quantified above.

---

5. References

1. Peng J, Chen H, Chen Z, et al. Prognostic value of neutrophil-to-lymphocyte ratio in patients with hepatocellular carcinoma receiving curative therapies: a systematic review and meta-analysis. *BMC Cancer*. 2025. DOI: 10.1186/s12885-025-13972-w

2. Jost-Brinkmann F, Demir M, Wree A, et al. Atezolizumab plus bevacizumab in unresectable hepatocellular carcinoma: results from a German real-world cohort. *Alimentary Pharmacology & Therapeutics*. 2023. PMID: 36883351.

3. Meng Y, Zhu X, et al. Prognostic value of the neutrophil-to-lymphocyte ratio and prognostic nutritional index in unresectable hepatocellular carcinoma patients treated with tyrosine kinase inhibitors and immune checkpoint inhibitors. *Human Vaccines &*

Immunotherapeutics. 2024. DOI: 10.1080/21645515.2024.2394268

4. Chen J, et al. Prognostic value of circulating tumor cells combined with neutrophil-lymphocyte ratio in patients with hepatocellular carcinoma. *World Journal of Gastroenterology*. 2024. PMC10900146.

5. Di D, et al. Dual-center retrospective cohort study of NLR and LMR in HAIC-based HCC. PMC12229162, 2025. (390 patients; NLR ≥ 5 threshold.)

6. Teo J, et al. Tumour-associated neutrophils attenuate the immunosensitivity of hepatocellular carcinoma. *Journal of Experimental Medicine*. 2025;222(1):e20241442. DOI: 10.1084/jem.20241442

7. Wu Y, et al. Neutrophil profiling illuminates anti-tumor antigen-presenting potency. *Cell*. 2024;187(6):1576–1596. DOI: 10.1016/j.cell.2024.02.005

8. Meng Y, Ye F, Nie P, et al. Immunosuppressive CD10+ALPL+ neutrophils promote resistance to anti-PD-1 therapy in HCC by mediating irreversible exhaustion of T cells. *Journal of Hepatology*. 2023;79:1435–1449.

9. Shen XT, Xie SZ, Zheng X, et al. Cirrhotic-extracellular matrix attenuates aPD-1 treatment response by initiating immunosuppressive neutrophil extracellular traps formation in hepatocellular carcinoma. *Experimental Hematology & Oncology*. 2024;13:20. DOI: 10.1186/s40164-024-00476-9

10. Grieshaber-Bouyer R, Radtke FA, Cunin P, et al. The neutrotime transcriptional signature defines a single continuum of neutrophils across biological compartments. *Nature Communications*. 2021;12:2856. DOI: 10.1038/s41467-021-22973-9

11. Antuamwine BB, Bosnjakovic R, Hofmann-Vega F, et al. N1 versus N2 and PMN-MDSC: a critical appraisal of current concepts on tumor-associated neutrophils and new directions for human oncology. *Immunological Reviews*. 2023;314:250–279. DOI: 10.1111/imr.13176

12. Horvath L, Puschmann C, Scheiber A, et al. Beyond binary: bridging neutrophil diversity to new therapeutic approaches in NSCLC. *Trends in Cancer*. 2024;10:457–474. DOI: 10.1016/j.trecan.2024.01.010

13. Fridlender ZG, Sun J, Kim S, et al. Polarization of tumor-associated neutrophil phenotype by TGF-beta: N1 versus N2 TAN. *Cancer Cell*. 2009;16(3):183–194. DOI:

10.1016/j.ccr.2009.06.017

14. Chen J, Feng W, Sun M, et al. TGF-β1-induced SOX18 elevation promotes hepatocellular carcinoma progression and metastasis through transcriptionally upregulating PD-L1 and CXCL12. *Gastroenterology*. 2024;167:264–280. DOI: 10.1053/j.gastro.2024.03.030

15. Guo J, et al. Impact of serum vascular endothelial growth factor on prognosis in patients with unresectable hepatocellular carcinoma after transarterial chemoembolization. PMC3555251. 2013. (60 HCC patients; Beijing Cancer Hospital; median VEGF 285 pg/mL, healthy controls 125 pg/mL.)

16. Nomogram for early HCC recurrence after R0 resection. *Frontiers in Oncology*. 2023. DOI: 10.3389/fonc.2023.1133807 (n=481; VEGF-A >240.3 pg/mL HR=2.552, P<0.001; mean VEGF 164.93 pg/mL.)

17. Alzamzamy A, Elsayed H, Abd Elraouf M, et al. Serum vascular endothelial growth factor as a tumor marker for hepatocellular carcinoma in hepatitis C virus-related cirrhotic patients. *World Journal of Gastrointestinal Oncology*. 2021;13(6):600–611. DOI: 10.4251/wjgo.v13.i6.600

18. Poon RTP, et al. High serum levels of vascular endothelial growth factor predict poor response to transarterial chemoembolization in hepatocellular carcinoma. *Annals of Surgical Oncology*. 2004. (Prospective; n=80; median VEGF 240 pg/mL; VEGF >240 → OS 6.8 vs 19.2 months.)

19. Leslie J, Mackey JBG, Jamieson T, et al. CXCR2 inhibition enables NASH-HCC immunotherapy. *Gut*. 2022;71:2523–2548.

20. Finn RS, Qin S, Ikeda M, et al. Atezolizumab plus bevacizumab in unresectable hepatocellular carcinoma. *New England Journal of Medicine*. 2020;382:1894–1905.

21. Singal AG, Kanwal F, Llovet JM. Global trends in hepatocellular carcinoma epidemiology: implications for screening, prevention and therapy. *Nature Reviews Clinical Oncology*. 2023;20:864–884.

22. Li et al. Development and validation of prognostic risk prediction models for hepatocellular carcinoma patients treated with immune checkpoint inhibitors based on a systematic review and meta-analysis of 47 cohorts. *Frontiers in Immunology*. 2023. DOI: 10.3389/fimmu.2023.1215745 (7,649 patients; validated on 204 patients from 21 centres;

NLR identified as independent predictor alongside AFP, ALBI, ECOG, Child-Pugh, BCLC, tumour number, vascular invasion, combination therapy.)

23. PMC12347834. Concordance Index-Based Comparison of Inflammatory and Classical Prognostic Markers in Untreated Hepatocellular Carcinoma. 2025. (n=250; NLR C-index 0.640; adding NLR improved model C-index from 0.781 to 0.794.)

24. Zhang et al. Advanced development of biomarkers for immunotherapy in hepatocellular carcinoma. PMC9885011. *Frontiers in Immunology*. 2023. (Documents NLR-VEGF collinearity mechanism: increased neutrophils secrete VEGF, promoting tumour angiogenesis.)

25. Li Y, et al. Angiogenesis in hepatocellular carcinoma: mechanisms and anti-angiogenic therapies. *Cancer Biology & Medicine*. 2023;20(1):25. (Multiple sources of tumour VEGF including neutrophils, platelets, tumour cells, CAFs, hepatic stellate cells.)

26. Pinato DJ, Stebbing J, Ishizuka M, et al. A novel and validated prognostic index in hepatocellular carcinoma: the inflammation based index (IBI). *Journal of Hepatology*. 2012;57(5):1013–1020 / Pinato DJ et al. *Oncotarget*. 2014. (NLR >4 as prognostic cutoff in sorafenib HCC cohort, n=65.)

TAN-POLARITY v3 is an agent-executable research framework. It is not a medical device, a validated clinical prediction tool, or a substitute for specialist oncological assessment. The Dq-weighted approach, collinearity correction, and data-first sigmoid calibration represent methodological improvements over v2, but the fundamental limitation — that no individual patient-level data were used to fit or validate the model — remains, and is stated here unambiguously.