HCC-METASCORE: A Biomarker-Driven Composite Scoring Framework for Systemic Therapy Signal Prioritisation in Hepatocellular Carcinoma with Extrahepatic Metastatic Spread

April 2026

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. Numbers assigned to domains are evidence-informed estimates grounded in cited literature; they are not coefficients derived from a prospectively validated regression model.

Abstract

Hepatocellular carcinoma (HCC) is the most prevalent form of primary liver cancer and a leading cause of cancer-related mortality worldwide. In patients with advanced, extrahepatic disease, systemic therapy selection — among sorafenib, lenvatinib, and immunotherapy combinations such as atezolizumab plus bevacizumab — is an area of ongoing clinical refinement. We present **HCC-METASCORE**, an agent-executable composite scoring framework that integrates biological features of metastatic HCC — including markers of epithelial-mesenchymal transition (EMT), microvascular invasion (MVI), tumour microenvironment (TME) immune activity, circulating biomarkers (CTCs, ctDNA), and key molecular/genetic drivers — into a structured 0–100 signal score. A Monte Carlo uncertainty layer propagates input measurement variability into a 95% confidence interval. Crucially, the score does not recommend a therapy or exclude one. Instead, it generates a **pathway signal profile** that highlights which biological features are most prominent in a given case, and maps these to the mechanistic targets of each systemic agent. The framework is designed for research prioritisation, multidisciplinary discussion scaffolding, and transparent agentic clinical reasoning.

---

1. Clinical Context and Justification

HCC arising in the setting of cirrhosis or chronic hepatitis commonly presents at an advanced stage in which surgical and locoregional therapies are no longer curative options. For these patients, systemic therapy is the primary modality, and three treatment strategies have established first-line or near-first-line evidence:

- **Sorafenib**, a multi-kinase inhibitor targeting VEGFR-1/2/3, PDGFR-β, Raf kinases, and other pro-angiogenic and proliferative pathways. It was the first agent to demonstrate an overall survival benefit in advanced HCC (SHARP trial, Llovet et al. 2008) and remains a benchmark comparator.
- **Lenvatinib**, which inhibits VEGFR-1/2/3, FGFR-1/2/3/4, PDGFR-α, RET, and KIT. The REFLECT trial (Kudo et al. 2018) demonstrated non-inferiority to sorafenib in overall survival with a higher objective response rate, and lenvatinib is now an established alternative first-line option.
- **Atezolizumab plus bevacizumab (Atezo/Bev)**, combining a PD-L1 checkpoint inhibitor with an anti-VEGF antibody. The IMbrave150 trial (Finn et al. 2020) demonstrated superior overall survival and progression-free survival compared with sorafenib in patients without contraindications. This combination is increasingly preferred where immune eligibility criteria are met.

The biological features that predict differential benefit from these three strategies are incompletely characterised. Nevertheless, mechanistic reasoning — grounded in the known biology of HCC metastasis — can structure a rational hypothesis about which pathway signals are most prominent in an individual patient's profile. This is the purpose of HCC-METASCORE: not to replace trial evidence with algorithmic certainty, but to make the biological reasoning explicit, executable, and auditable.

---

2. Biological Basis: Mechanisms of HCC Metastatic Spread

The framework draws on five mechanistic axes derived from the biology of HCC metastasis.

2.1 Epithelial-Mesenchymal Transition (EMT)

EMT is the process by which tumour cells downregulate epithelial markers (E-cadherin) and upregulate mesenchymal markers (vimentin, N-cadherin, fibronectin), acquiring migratory and invasive capacity. In HCC, EMT is driven by TGF-β signalling, Wnt/β-catenin activation, and

hypoxia-inducible pathways. Aberrant Wnt/β-catenin activation is present in 20–35% of HCC cases and is associated with stem-like properties in metastasis-initiating cells (Hoshida et al. 2009; Schulze et al. 2015). EMT underlies the capacity of HCC cells to invade local vasculature — a prerequisite for both haematogenous and lymphatic dissemination.

2.2 Microvascular Invasion (MVI) and Portal Vein Involvement

MVI — the presence of tumour emboli in small portal or hepatic venous radicles — is one of the strongest histological predictors of recurrence and extrahepatic spread. Macrovascular invasion of the portal vein defines Barcelona Clinic Liver Cancer (BCLC) stage C disease and is associated with markedly reduced prognosis. Single-cell RNA sequencing has identified distinct hepatocyte subpopulations with preferential routes of spread: lymph node metastasis-associated hepatocytes (LNMAHs) and portal vein metastasis-associated hepatocytes (PVMAHs), suggesting that metastatic route is partially encoded within the primary tumour (Ma et al. 2021, *Cancer Cell*).

2.3 Tumour Microenvironment (TME) and Immune Evasion

The TME in HCC is characterised by immunosuppression mediated through regulatory T cells (Tregs), tumour-associated macrophages (TAMs) polarised toward an M2 phenotype, myeloid-derived suppressor cells (MDSCs), and upregulation of immune checkpoint molecules including PD-1/PD-L1, CTLA-4, and TIM-3. TGF-β is a dominant immunosuppressive cytokine in HCC and simultaneously drives EMT. The density of CD8+ tumour-infiltrating lymphocytes (TILs) and the ratio of effector T cells to Tregs have been proposed as predictors of immunotherapy responsiveness (Llovet et al. 2021, *Nature Reviews Clinical Oncology*).

2.4 Angiogenesis

HCC is among the most vascular of solid tumours. VEGF-A overexpression correlates with tumour size, microvascular density, and metastatic potential. Elevated serum VEGF is associated with worse prognosis and is the mechanistic rationale for both sorafenib and bevacizumab. FGFR signalling contributes an additional angiogenic axis specifically targeted by lenvatinib.

2.5 Molecular and Genetic Drivers

Three tumour suppressor genes — **TP53**, **RB1**, and **PTEN** — are among the most recurrently mutated in extrahepatic HCC metastases. TP53 loss disrupts apoptotic checkpoints, permitting survival of genomically unstable cells. PTEN loss de-represses PI3K/Akt/mTOR signalling, promoting invasion and metabolic adaptation. RB1 loss relieves cell cycle restraint. **Wnt/β-catenin** pathway activation (CTNNB1 mutations) occurs in 20–35% of cases. Epigenetic dysregulation — aberrant DNA methylation, histone modification, and three-dimensional

chromatin remodelling — further modulates the expression of pro-metastatic gene programmes without altering the underlying sequence.

2.6 Circulating Biomarkers

Alpha-fetoprotein (AFP) and its lectin-reactive isoform AFP-L3, and des-gamma-carboxyprothrombin (DCP/PIVKA-II), are the most established serum biomarkers for HCC surveillance and staging. Circulating tumour cells (CTCs) and circulating tumour DNA (ctDNA) are emerging minimally invasive markers for early detection of metastatic dissemination and tumour heterogeneity (Ye et al. 2021). Neutrophil-to-lymphocyte ratio (NLR) has been validated as an inflammatory prognostic marker in multiple HCC cohorts.

---

3. Scoring Architecture

3.1 Domain Selection and Weight Rationale

HCC-METASCORE integrates **10 domains** across five biological axes. Weights reflect the relative strength of published associations between each domain and metastatic potential or systemic therapy signal, as summarised below. Weights are explicitly expert-informed estimates, not regression coefficients.

Domain	Axis	Weight	Rationale
AFP / AFP-L3 level	Circulating biomarkers	0.12	AFP >400 ng/mL and AFP-L3 >15% are associated with MVI and vascular invasion; established clinical markers (Lok et al. 2010)
DCP / PIVKA-II	Circulating biomarkers	0.08	Independent predictor of portal vein invasion and extrahepatic spread; elevated in ~40–60% of advanced HCC (Imamura et al. 2008)
CTC / ctDNA detection	Circulating biomarkers	0.10	Detection indicates active haematogenous dissemination; ctDNA allele fraction tracks tumour burden (Ye et al. 2021)
Microvascular / macrovascular invasion	MVI	0.14	Strongest histological predictor of extrahepatic spread; macrovascular portal invasion = BCLC-C (Roayaie et al. 2004)

EMT marker profile	EMT	0.10	Vimentin, N-cadherin, E-cadherin loss; Wnt/β-catenin activation; direct drivers of invasive capacity (Schulze et al. 2015)
TME immune activity (TILs, PD-L1, NLR)	TME / Immune	0.14	PD-L1 expression and TIL density are candidate predictors of checkpoint inhibitor responsiveness (Llovet et al. 2021)
Angiogenic burden (VEGF, MVD)	Angiogenesis	0.12	VEGF overexpression directly targets sorafenib and bevacizumab mechanisms (Llovet et al. 2008; Finn et al. 2020)
FGFR signalling	Angiogenesis	0.06	Specific target of lenvatinib; FGFR amplification/dysregulation enriches lenvatinib signal (Kudo et al. 2018)
Genetic driver burden (TP53, PTEN, RB1)	Molecular	0.10	Convergent loss associated with aggressive metastatic phenotype; TP53 most recurrent in extrahepatic deposits
Wnt/β-catenin / epigenetic dysregulation	Molecular	0.04	CTNNB1 mutations associated with EMT and stem-like metastatic cells; epigenetic contribution emerging (Hoshida et al. 2009)

Total weight: 1.00

3.2 Composite Score Calculation

Each domain receives a raw subscale score (0–100) based on measured or estimated inputs. The composite score is:

Composite = Σ (domain_raw_score × domain_weight)

Capped at 100. A higher composite score reflects greater overall metastatic biological burden signal across the integrated domains.

3.3 Monte Carlo Uncertainty Layer

Continuous inputs (AFP level, DCP level, ctDNA variant allele fraction, NLR, VEGF level) carry analytical measurement variability. To represent this honestly, 5,000 Monte Carlo simulations perturb each continuous input with Gaussian noise (coefficient of variation =

10–15%, reflecting typical inter-laboratory variability) and recompute the composite score. The 2.5th and 97.5th percentile outputs form the reported 95% confidence interval.

This CI reflects input measurement uncertainty, not model validation uncertainty. The model has not been prospectively calibrated or externally validated. The CI is a tool for intellectual honesty, not a statistical guarantee.

3.4 Pathway Signal Profile

Beyond the composite score, HCC-METASCORE generates a **Pathway Signal Profile** — a structured read-out of which biological axes are most elevated in the patient's feature set, mapped to the mechanistic targets of each systemic agent:

Agent	Mechanistic targets	Relevant domains
Sorafenib	VEGFR-1/2/3, PDGFR-β, Raf	Angiogenic burden (VEGF), AFP (vascular invasion proxy), MVI
Lenvatinib	VEGFR-1/2/3, FGFR-1/2/3/4, PDGFR-α, RET, KIT	Angiogenic burden (VEGF + FGFR), AFP, MVI
Atezo/Bev	PD-L1 (atezolizumab) + VEGF-A (bevacizumab)	TME immune activity, PD-L1 expression, TIL density, NLR, Angiogenic burden

High scores in immune axis domains elevate the Atezo/Bev signal flag. High angiogenic burden with low immune activation elevates the anti-angiogenic monotherapy signal. FGFR-specific dysregulation adds specificity toward lenvatinib over sorafenib. These are **hypothesis-generating flags**, not treatment decisions.

3.5 Score Categories

Composite Score	Category	Interpretation
< 20	LOW	Limited convergent metastatic signalling across assessed domains
20–39	MODERATE	Moderate multi-domain metastatic burden signal; pathway profile guides focus
40–59	HIGH	Substantial convergent signal; multiple biological axes elevated

≥ 60	VERY HIGH	Strong multi-axis metastatic burden signal; all pathway flags should be reviewed

---

4. Demonstration Scenarios

Scenario 1 — Early extrahepatic HCC, predominantly angiogenic profile

Patient: AFP 850 ng/mL, AFP-L3 22%, DCP 180 mAU/mL, no CTCs detected, ctDNA VAF 0.8%, MVI present (micro only), EMT markers: E-cadherin reduced, vimentin elevated, NLR 3.2, PD-L1 TPS 2%, TIL density low, VEGF elevated (310 pg/mL), no FGFR amplification, TP53 mutated, PTEN intact, Wnt/β-catenin: CTNNB1 mutation present.

Composite score: 44.1 / 100 [HIGH] 95% CI: [40.8, 47.6]

Pathway signal profile:

- Angiogenic burden: **HIGH** → sorafenib and lenvatinib pathways both flagged
- Immune axis: **LOW** (PD-L1 2%, low TILs) → Atezo/Bev immune signal not prominent
- FGFR signal: **ABSENT** → lenvatinib-specific advantage over sorafenib not supported by this domain
- Genetic driver burden: **MODERATE** (TP53, CTNNB1)
- Circulating biomarker burden: **MODERATE-HIGH**

Interpretive note: This profile's elevated angiogenic features are more consistent with pathways targeted by anti-VEGFR agents. The low immune activation signal does not currently support a hypothesis that immune checkpoint engagement is the primary mechanistic opportunity. These observations are hypothesis-generating only.

Scenario 2 — Advanced HCC with immune-inflamed microenvironment

Patient: AFP 320 ng/mL, AFP-L3 8%, DCP 95 mAU/mL, 3 CTCs detected/7.5 mL, ctDNA VAF 3.2%, MVI present (macro, portal vein involved), EMT: full mesenchymal shift, NLR 2.1, PD-L1 TPS 18%, TIL density high (CD8+ infiltration confirmed), VEGF 190 pg/mL, no FGFR amplification, TP53 wildtype, PTEN loss, Wnt/β-catenin: CTNNB1 wildtype.

Composite score: 52.3 / 100 [HIGH] 95% CI: [48.5, 56.1]

Pathway signal profile:

- Angiogenic burden: **MODERATE**
- Immune axis: **HIGH** (PD-L1 18%, high TILs, low NLR) → Atezo/Bev immune component flagged
- FGFR signal: **ABSENT**
- Circulating burden: **HIGH** (CTCs, ctDNA, portal vein MVI)
- Genetic driver burden: **MODERATE** (PTEN loss)

Interpretive note: The immune axis signals in this profile are more consistent with features associated with checkpoint inhibitor responsiveness. The combination of moderate angiogenic burden with high immune activation may be more aligned with a combined anti-PD-L1/anti-VEGF mechanistic rationale. The CTNNB1 wildtype status is noted — Wnt-activated HCC has been associated with reduced immune infiltration (Ruiz de Galarreta et al. 2019), and its absence here is consistent with the observed immune-inflamed profile.

Scenario 3 — Aggressive multi-route dissemination with mixed signals

Patient: AFP 12,400 ng/mL, AFP-L3 41%, DCP 880 mAU/mL, 9 CTCs detected/7.5 mL, ctDNA VAF 8.7%, macro-MVI with portal and hepatic vein involvement, full EMT marker panel positive, NLR 6.8, PD-L1 TPS 5%, TIL density low-moderate, VEGF 520 pg/mL, FGFR1 amplification detected, TP53 mutated, PTEN loss, RB1 loss, CTNNB1 mutation.

Composite score: 79.4 / 100 [VERY HIGH] 95% CI: [74.9, 83.2]

Pathway signal profile:

- Angiogenic burden: **VERY HIGH** (VEGF 520, FGFR1 amplified) → both sorafenib and lenvatinib pathways flagged; FGFR amplification adds lenvatinib-specific signal
- Immune axis: **LOW-MODERATE** (PD-L1 5%, NLR 6.8 indicating systemic inflammation suppressing immune response)
- Circulating burden: **VERY HIGH**
- Genetic driver burden: **VERY HIGH** (TP53 + PTEN + RB1 convergent loss)
- EMT signal: **VERY HIGH**

Interpretive note: This profile is characterised by very high angiogenic and vascular invasion signals with a suppressed immune microenvironment. The FGFR1 amplification is the only domain that mechanistically differentiates lenvatinib from sorafenib in the current feature set. The high NLR and low TIL density suggest a non-inflamed TME, which is generally associated with attenuated checkpoint inhibitor responsiveness — though this inference is hypothesis-generating only and does not exclude Atezo/Bev as a consideration. The very high

multi-driver genetic burden may indicate a biologically aggressive tumour warranting closer monitoring regardless of therapy choice.

---

5. Executable Python Implementation

```
#!/usr/bin/env python3
"""
HCC-METASCORE: Biomarker-Driven Composite Scoring Framework for
Systemic Therapy Signal Prioritisation in HCC with Extrahepatic Metastatic Spread

Purpose: Research tool to narrow the field of analytical focus and structure
multidisciplinary discussion. NOT for clinical prescribing or treatment decisions.

Weight rationale:
- AFP/AFP-L3 (0.12): Validated vascular invasion proxy (Lok et al. 2010)
- DCP/PIVKA-II (0.08): Portal vein invasion predictor (Imamura et al. 2008)
- CTC/ctDNA (0.10): Active haematogenous dissemination signal (Ye et al. 2021)
- MVI (0.14): Strongest histological metastatic predictor (Roayaie et al. 2004)
- EMT markers (0.10): Direct invasive capacity drivers (Schulze et al. 2015)
- TME/immune (0.14): Checkpoint inhibitor responsiveness candidate (Llovet et al. 2021)
- Angiogenic burden (0.12): Core sorafenib/bevacizumab target (Llovet et al. 2008)
- FGFR signal (0.06): Lenvatinib-specific target (Kudo et al. 2018)
- Genetic drivers (0.10): TP53/PTEN/RB1 metastatic burden
- Wnt/epigenetic (0.04): EMT and stem-cell metastatic programme
"""

from __future__ import annotations
import random
from dataclasses import dataclass, field
from typing import Optional, List

# ──────────────────────────────────────────────────────────────────────────────
# Patient input dataclass
# ──────────────────────────────────────────────────────────────────────────────

@dataclass
class HCCPatient:
    # Circulating biomarkers
    afp_ng_ml: float = 20.0              # AFP in ng/mL; normal <20
```

```
    afp_l3_pct: float = 5.0              # AFP-L3 fraction %; high-risk threshold ~15%
    dcp_mau_ml: float = 40.0             # DCP/PIVKA-II; high-risk threshold ~40 mAU/mL
    ctc_per_75ml: int = 0                # CTCs per 7.5 mL blood
    ctdna_vaf_pct: float = 0.0           # ctDNA variant allele fraction %

    # Vascular invasion
    microvascular_invasion: bool = False     # MVI on histology
    macrovascular_invasion: bool = False     # Portal/hepatic vein tumour thrombus

    # EMT markers
    e_cadherin_loss: bool = False        # Reduced E-cadherin expression
    vimentin_positive: bool = False      # Vimentin upregulation
    ctnnb1_mutation: bool = False        # Wnt/beta-catenin activating mutation

    # TME / immune
    nlr: float = 2.5                     # Neutrophil-to-lymphocyte ratio
    pd_l1_tps_pct: float = 0.0           # PD-L1 tumour proportion score %
    til_density: str = "low"             # "low", "moderate", "high"

    # Angiogenesis
    vegf_pg_ml: float = 100.0            # Serum VEGF in pg/mL
    fgfr_amplification: bool = False     # FGFR1/2/3/4 amplification or mutation

    # Genetic drivers
    tp53_mutation: bool = False
    pten_loss: bool = False
    rb1_loss: bool = False

    # Epigenetic
    epigenetic_dysregulation: bool = False   # Aberrant methylation / chromatin remodelling
reported


# ──────────────────────────────────────────────────────────────────────────
# Domain scoring functions
# ──────────────────────────────────────────────────────────────────────────

def score_afp(afp: float, afp_l3: float) -> tuple[float, str]:
    """
    AFP >400 ng/mL and AFP-L3 >15% are established thresholds for MVI risk.
```

```
    Ref: Lok et al. Hepatology 2010.
    """
    s = 0.0
    if afp > 400:
        s += 55
    elif afp > 200:
        s += 35
    elif afp > 100:
        s += 20
    elif afp > 20:
        s += 8
    if afp_l3 > 15:
        s += 30
    elif afp_l3 > 10:
        s += 15
    return min(s, 100), f"AFP={afp:.0f} ng/mL, AFP-L3={afp_l3:.1f}%"


def score_dcp(dcp: float) -> tuple[float, str]:
    """
    DCP/PIVKA-II >40 mAU/mL associated with portal vein invasion.
    Ref: Imamura et al. Hepatology 2008.
    """
    if dcp > 400:
        return 90, f"DCP={dcp:.0f} mAU/mL [markedly elevated]"
    if dcp > 100:
        return 60, f"DCP={dcp:.0f} mAU/mL [elevated]"
    if dcp > 40:
        return 30, f"DCP={dcp:.0f} mAU/mL [above threshold]"
    return 5, f"DCP={dcp:.0f} mAU/mL [normal range]"


def score_ctc_ctdna(ctc: int, vaf: float) -> tuple[float, str]:
    """
    CTC detection indicates haematogenous dissemination.
    ctDNA VAF tracks tumour burden. Ref: Ye et al. Hepatology 2021.
    """
    s = 0.0
    if ctc >= 5:
        s += 50
```

```
    elif ctc >= 2:
        s += 30
    elif ctc == 1:
        s += 15
    if vaf >= 5.0:
        s += 40
    elif vaf >= 1.0:
        s += 25
    elif vaf >= 0.5:
        s += 10
    return min(s, 100), f"CTCs={ctc}/7.5mL, ctDNA VAF={vaf:.1f}%"


def score_mvi(micro: bool, macro: bool) -> tuple[float, str]:
    """
    MVI is the strongest histological predictor of extrahepatic spread.
    Macrovascular = BCLC-C. Ref: Roayaie et al. Ann Surg 2004.
    """
    if macro:
        return 95, "Macrovascular invasion present (portal/hepatic vein)"
    if micro:
        return 50, "Microvascular invasion present on histology"
    return 0, "No vascular invasion identified"


def score_emt(e_cad_loss: bool, vimentin: bool, ctnnb1: bool) -> tuple[float, str]:
    """
    EMT marker composite. E-cadherin loss + vimentin = full mesenchymal shift.
    CTNNB1 mutation activates Wnt/beta-catenin, drives stemness.
    Ref: Schulze et al. Nature Genetics 2015.
    """
    s = 0.0
    details = []
    if e_cad_loss:
        s += 35
        details.append("E-cad loss")
    if vimentin:
        s += 35
        details.append("vimentin+")
    if ctnnb1:
```

```
        s += 25
        details.append("CTNNB1 mut")
    return min(s, 100), ", ".join(details) if details else "No EMT markers detected"


def score_tme(nlr: float, pd_l1: float, til: str) -> tuple[float, str]:
    """
    TME immune profile. High PD-L1 + high TILs = immune-inflamed phenotype,
    associated with checkpoint inhibitor responsiveness.
    NLR >5 indicates systemic inflammation suppressing immune response.
    Ref: Llovet et al. Nature Reviews Clinical Oncology 2021.
    """
    s = 0.0
    details = []
    # PD-L1 expression
    if pd_l1 >= 10:
        s += 40
        details.append(f"PD-L1 TPS {pd_l1:.0f}%")
    elif pd_l1 >= 1:
        s += 20
        details.append(f"PD-L1 TPS {pd_l1:.0f}%")
    else:
        details.append("PD-L1 <1%")
    # TIL density
    if til == "high":
        s += 35
        details.append("TIL-high")
    elif til == "moderate":
        s += 15
        details.append("TIL-moderate")
    else:
        details.append("TIL-low")
    # NLR: high NLR is immunosuppressive — penalises the immune signal
    if nlr > 5:
        s = max(s - 20, 0)
        details.append(f"NLR {nlr:.1f} [immunosuppressed systemic milieu]")
    else:
        details.append(f"NLR {nlr:.1f}")
    return min(s, 100), ", ".join(details)
```

```
def score_angiogenesis(vegf: float) -> tuple[float, str]:
    """
    VEGF drives HCC vascularisation and is the primary mechanistic target
    of sorafenib and bevacizumab. Ref: Llovet et al. NEJM 2008.
    """
    if vegf > 400:
        return 90, f"VEGF={vegf:.0f} pg/mL [markedly elevated]"
    if vegf > 250:
        return 65, f"VEGF={vegf:.0f} pg/mL [elevated]"
    if vegf > 150:
        return 40, f"VEGF={vegf:.0f} pg/mL [moderately elevated]"
    if vegf > 80:
        return 20, f"VEGF={vegf:.0f} pg/mL [mildly elevated]"
    return 5, f"VEGF={vegf:.0f} pg/mL [near normal]"


def score_fgfr(fgfr_amp: bool) -> tuple[float, str]:
    """
    FGFR amplification/dysregulation is targeted by lenvatinib but not sorafenib.
    Presence is a lenvatinib-differentiating signal. Ref: Kudo et al. Lancet 2018.
    """
    if fgfr_amp:
        return 85, "FGFR amplification/dysregulation detected"
    return 0, "No FGFR amplification detected"


def score_genetic_drivers(tp53: bool, pten: bool, rb1: bool) -> tuple[float, str]:
    """
    Convergent loss of TP53, PTEN, and RB1 is associated with aggressive
    metastatic phenotype in extrahepatic HCC deposits.
    """
    s = 0.0
    details = []
    if tp53:
        s += 40
        details.append("TP53 mut")
    if pten:
        s += 35
        details.append("PTEN loss")
```

```
    if rb1:
        s += 25
        details.append("RB1 loss")
    return min(s, 100), ", ".join(details) if details else "No high-risk driver mutations detected"


def score_wnt_epigenetic(ctnnb1: bool, epi: bool) -> tuple[float, str]:
    """
    Wnt/beta-catenin and epigenetic dysregulation contribute to pro-metastatic
    gene programme activation. CTNNB1 mutation also noted in EMT domain.
    Ref: Hoshida et al. New England Journal of Medicine 2008.
    """
    s = 0.0
    details = []
    if ctnnb1:
        s += 50
        details.append("CTNNB1 mutation")
    if epi:
        s += 45
        details.append("Epigenetic dysregulation reported")
    return min(s, 100), ", ".join(details) if details else "No Wnt/epigenetic dysregulation noted"


# ──────────────────────────────────────────────────────────────────────
# Weights
# ──────────────────────────────────────────────────────────────────────

WEIGHTS = {
    "afp":          0.12,
    "dcp":          0.08,
    "ctc_ctdna":    0.10,
    "mvi":          0.14,
    "emt":          0.10,
    "tme":          0.14,
    "angiogenesis": 0.12,
    "fgfr":         0.06,
    "genetic":      0.10,
    "wnt_epi":      0.04,
}
assert abs(sum(WEIGHTS.values()) - 1.0) < 1e-9, "Weights must sum to 1.0"
```

```
# ──────────────────────────────────────────────────────────────────
# Result dataclass
# ──────────────────────────────────────────────────────────────────

@dataclass
class HCCResult:
    composite_score: float
    ci_lower: float
    ci_upper: float
    score_category: str
    domains: list
    pathway_signal: dict
    interpretive_notes: List[str] = field(default_factory=list)


# ──────────────────────────────────────────────────────────────────
# Core computation
# ──────────────────────────────────────────────────────────────────

def compute_domain_scores(p: HCCPatient) -> list:
    return [
        ("afp",          *score_afp(p.afp_ng_ml, p.afp_l3_pct),          WEIGHTS["afp"]),
        ("dcp",          *score_dcp(p.dcp_mau_ml),                       WEIGHTS["dcp"]),
        ("ctc_ctdna",    *score_ctc_ctdna(p.ctc_per_75ml, p.ctdna_vaf_pct),
WEIGHTS["ctc_ctdna"]),
        ("mvi",          *score_mvi(p.microvascular_invasion, p.macrovascular_invasion),
WEIGHTS["mvi"]),
        ("emt",          *score_emt(p.e_cadherin_loss, p.vimentin_positive, p.ctnnb1_mutation),
WEIGHTS["emt"]),
        ("tme",          *score_tme(p.nlr, p.pd_l1_tps_pct, p.til_density), WEIGHTS["tme"]),
        ("angiogenesis", *score_angiogenesis(p.vegf_pg_ml),
WEIGHTS["angiogenesis"]),
        ("fgfr",         *score_fgfr(p.fgfr_amplification),               WEIGHTS["fgfr"]),
        ("genetic",      *score_genetic_drivers(p.tp53_mutation, p.pten_loss, p.rb1_loss),
WEIGHTS["genetic"]),
        ("wnt_epi",      *score_wnt_epigenetic(p.ctnnb1_mutation, p.epigenetic_dysregulation),
WEIGHTS["wnt_epi"]),
    ]
```

```
def pathway_signal_profile(domains: list) -> dict:
    """
    Maps domain scores to three therapy signal axes.
    Returns descriptive signal levels, not recommendations.
    """
    domain_map = {d[0]: d[1] for d in domains}

    sorafenib_signal = (
        domain_map["angiogenesis"] * 0.50 +
        domain_map["afp"] * 0.30 +
        domain_map["mvi"] * 0.20
    )
    lenvatinib_signal = (
        domain_map["angiogenesis"] * 0.40 +
        domain_map["fgfr"] * 0.30 +
        domain_map["afp"] * 0.20 +
        domain_map["mvi"] * 0.10
    )
    atezo_bev_signal = (
        domain_map["tme"] * 0.55 +
        domain_map["angiogenesis"] * 0.30 +
        domain_map["ctc_ctdna"] * 0.15
    )

    def level(score):
        if score >= 60: return "VERY HIGH"
        if score >= 40: return "HIGH"
        if score >= 20: return "MODERATE"
        return "LOW"

    return {
        "Sorafenib pathway signal": (round(sorafenib_signal, 1), level(sorafenib_signal)),
        "Lenvatinib pathway signal": (round(lenvatinib_signal, 1), level(lenvatinib_signal)),
        "Atezo/Bev pathway signal": (round(atezo_bev_signal, 1), level(atezo_bev_signal)),
    }
```

```
def compute_hcc_score(patient: HCCPatient, n_simulations: int = 5000, seed: int = 42) ->
HCCResult:
    domains = compute_domain_scores(patient)
    composite = min(sum(score * weight for _, score, _, weight in domains), 100.0)

    # Monte Carlo: perturb continuous inputs only
    rng = random.Random(seed)
    sims = []
    for _ in range(n_simulations):
        def perturb(val, cv=0.12):
            return max(0.0, val * (1 + rng.gauss(0, cv)))

        noisy = HCCPatient(
            afp_ng_ml=perturb(patient.afp_ng_ml),
            afp_l3_pct=min(100, perturb(patient.afp_l3_pct)),
            dcp_mau_ml=perturb(patient.dcp_mau_ml),
            ctc_per_75ml=patient.ctc_per_75ml,  # integer, not perturbed
            ctdna_vaf_pct=perturb(patient.ctdna_vaf_pct, cv=0.15),
            microvascular_invasion=patient.microvascular_invasion,
            macrovascular_invasion=patient.macrovascular_invasion,
            e_cadherin_loss=patient.e_cadherin_loss,
            vimentin_positive=patient.vimentin_positive,
            ctnnb1_mutation=patient.ctnnb1_mutation,
            nlr=perturb(patient.nlr, cv=0.10),
            pd_l1_tps_pct=min(100, perturb(patient.pd_l1_tps_pct)),
            til_density=patient.til_density,
            vegf_pg_ml=perturb(patient.vegf_pg_ml),
            fgfr_amplification=patient.fgfr_amplification,
            tp53_mutation=patient.tp53_mutation,
            pten_loss=patient.pten_loss,
            rb1_loss=patient.rb1_loss,
            epigenetic_dysregulation=patient.epigenetic_dysregulation,
        )
        nd = compute_domain_scores(noisy)
        sims.append(min(sum(s * w for _, s, _, w in nd), 100.0))

    sims.sort()
    ci_lower = round(sims[int(0.025 * n_simulations)], 1)
    ci_upper = round(sims[int(0.975 * n_simulations)], 1)
```

```
    if composite < 20:
        category = "LOW"
    elif composite < 40:
        category = "MODERATE"
    elif composite < 60:
        category = "HIGH"
    else:
        category = "VERY HIGH"

    pathway = pathway_signal_profile(domains)

    # Interpretive notes
    notes = []
    domain_map = {d[0]: d[1] for d in domains}
    if domain_map["tme"] >= 40 and domain_map["angiogenesis"] >= 40:
        notes.append("Elevated immune and angiogenic signals co-occur — consistent with a
profile where combined anti-PD-L1 / anti-VEGF mechanisms may be relevant to explore.")
    if domain_map["fgfr"] >= 70:
        notes.append("FGFR amplification/dysregulation detected — the only domain in this
framework that mechanistically differentiates lenvatinib from sorafenib.")
    if domain_map["tme"] < 20 and domain_map["angiogenesis"] >= 50:
        notes.append("Non-inflamed TME with high angiogenic burden — profile features are
more aligned with anti-angiogenic monotherapy mechanisms in the current framework.")
    if patient.ctnnb1_mutation and domain_map["tme"] < 30:
        notes.append("CTNNB1 mutation with low immune signal — consistent with
Wnt-activated HCC phenotype, which has been reported to associate with reduced immune
infiltration (Ruiz de Galarreta et al. 2019). This does not exclude immune-based strategies but is
noted for research awareness.")
    if domain_map["genetic"] >= 70:
        notes.append("High convergent genetic driver burden (TP53/PTEN/RB1) — associated
with aggressive metastatic phenotype; warrants close monitoring regardless of therapeutic
direction.")

    return HCCResult(
        composite_score=round(composite, 1),
        ci_lower=ci_lower,
        ci_upper=ci_upper,
        score_category=category,
        domains=[{"domain": d[0], "raw_score": round(d[1], 1), "detail": d[2], "weight": d[3],
"weighted": round(d[1]*d[3], 2)} for d in domains],
```

```
        pathway_signal=pathway,
        interpretive_notes=notes,
    )


# ─────────────────────────────────────────────────────────────────────
# Output printer
# ─────────────────────────────────────────────────────────────────────

def print_result(result: HCCResult, label: str):
    SEP = "=" * 72
    print(f"\n{SEP}\n{label}\n{SEP}")
    print(f"Composite score: {result.composite_score}/100 [{result.score_category}]")
    print(f"95% CI: [{result.ci_lower}, {result.ci_upper}]  (reflects input measurement
variability)")
    print("\nDomain breakdown:")
    for d in result.domains:
        print(f"  {d['domain']:15s} raw={d['raw_score']:5.1f}  weight={d['weight']:.2f}
weighted={d['weighted']:.2f}  | {d['detail']}")
    print("\nPathway signal profile:")
    for agent, (score, level) in result.pathway_signal.items():
        print(f"  {agent}: {score:.1f} [{level}]")
    if result.interpretive_notes:
        print("\nInterpretive notes (hypothesis-generating only):")
        for note in result.interpretive_notes:
            print(f"  * {note}")
    print(f"\n{'─'*72}")
    print("REMINDER: This output is for research focus and discussion only.")
    print("It does not recommend, rank, or eliminate any treatment.")
    print(f"{'─'*72}")


# ─────────────────────────────────────────────────────────────────────
# Demo
# ─────────────────────────────────────────────────────────────────────

def demo():
    scenarios = [
        (
            "Scenario 1 — Early extrahepatic HCC, predominantly angiogenic profile",
```

```
        HCCPatient(
            afp_ng_ml=850, afp_l3_pct=22, dcp_mau_ml=180,
            ctc_per_75ml=0, ctdna_vaf_pct=0.8,
            microvascular_invasion=True, macrovascular_invasion=False,
            e_cadherin_loss=True, vimentin_positive=True, ctnnb1_mutation=True,
            nlr=3.2, pd_l1_tps_pct=2.0, til_density="low",
            vegf_pg_ml=310, fgfr_amplification=False,
            tp53_mutation=True, pten_loss=False, rb1_loss=False,
            epigenetic_dysregulation=False,
        ),
    ),
    (
        "Scenario 2 — Advanced HCC with immune-inflamed microenvironment",
        HCCPatient(
            afp_ng_ml=320, afp_l3_pct=8, dcp_mau_ml=95,
            ctc_per_75ml=3, ctdna_vaf_pct=3.2,
            microvascular_invasion=True, macrovascular_invasion=True,
            e_cadherin_loss=True, vimentin_positive=True, ctnnb1_mutation=False,
            nlr=2.1, pd_l1_tps_pct=18.0, til_density="high",
            vegf_pg_ml=190, fgfr_amplification=False,
            tp53_mutation=False, pten_loss=True, rb1_loss=False,
            epigenetic_dysregulation=False,
        ),
    ),
    (
        "Scenario 3 — Aggressive multi-route dissemination, mixed signals",
        HCCPatient(
            afp_ng_ml=12400, afp_l3_pct=41, dcp_mau_ml=880,
            ctc_per_75ml=9, ctdna_vaf_pct=8.7,
            microvascular_invasion=True, macrovascular_invasion=True,
            e_cadherin_loss=True, vimentin_positive=True, ctnnb1_mutation=True,
            nlr=6.8, pd_l1_tps_pct=5.0, til_density="low",
            vegf_pg_ml=520, fgfr_amplification=True,
            tp53_mutation=True, pten_loss=True, rb1_loss=True,
            epigenetic_dysregulation=True,
        ),
    ),
]

for label, patient in scenarios:
```

```
        result = compute_hcc_score(patient)
        print_result(result, label)


if __name__ == "__main__":
    demo()
```

---

6. Explicit Limitations

The following limitations are not caveats appended for completeness — they are central to the correct use of this tool.

This tool does not diagnose, treat, prescribe, or eliminate therapeutic options. Its sole purpose is to make biological reasoning about metastatic HCC more explicit, structured, and auditable in a research or multidisciplinary discussion context.

The model is not externally validated. Domain weights are evidence-informed estimates anchored to cited literature, not regression coefficients derived from a prospective cohort. The tool cannot compute sensitivity, specificity, or positive/negative predictive value.

The Monte Carlo confidence interval reflects input measurement noise, not model performance. A narrow CI does not imply that the model is correct; it implies that the score is stable across small perturbations of the continuous inputs.

Pathway signal profiles are mechanistic hypotheses, not response predictions. The biological features that correlate with, for example, immune activation do not guarantee checkpoint inhibitor responsiveness. Tumour heterogeneity, intrahepatic vs. extrahepatic discordance, prior therapy, performance status, hepatic reserve (Child-Pugh score), and contraindications to specific agents are not captured in this framework.

Biomarker availability varies. Not all patients will have ctDNA, CTC, FGFR amplification, or TME profiling data. Missing domains should be treated as uninformative (score = 0) rather than low-risk, and results interpreted with the gap in mind.

CTNNB1 mutation and immune checkpoint responsiveness. Emerging data suggest that Wnt/β-catenin-activated HCC may have a non-inflamed TME and attenuated checkpoint inhibitor responsiveness (Ruiz de Galarreta et al. *Cancer Cell* 2019). This relationship is not yet incorporated into clinical guidelines and is presented as a research-level signal.

All outputs require expert medical review. This framework is designed for use by clinicians and researchers who can contextualise its outputs within the full clinical picture. It is not appropriate for use without such oversight.

---

7. Why This Framework Exists

Three concrete research and clinical use cases motivate HCC-METASCORE:

Biological profile mapping. When discussing a case in a multidisciplinary tumour board, this framework provides a structured vocabulary for which metastatic mechanisms are most prominent — shifting the conversation from anecdotal impression to explicit, citable domain-level reasoning.

Research prioritisation. For investigators designing HCC biomarker studies or exploratory correlative analyses, the framework identifies which domain combinations may most distinguish patient subgroups with different systemic therapy pathway signals.

Agentic auditability. In AI-mediated clinical tooling, safety-relevant reasoning should not live in hidden prompts. An explicit, executable, open framework allows domain weights to be challenged, updated, and calibrated as evidence matures.

---

8. References

1. Llovet JM, Ricci S, Mazzaferro V, et al. Sorafenib in advanced hepatocellular carcinoma. *New England Journal of Medicine*. 2008;359(4):378–390. DOI: 10.1056/NEJMoa0708857

2. Kudo M, Finn RS, Qin S, et al. Lenvatinib versus sorafenib in first-line treatment of patients with unresectable hepatocellular carcinoma: a randomised phase 3 non-inferiority trial (REFLECT). *Lancet*. 2018;391(10126):1163–1173. DOI: 10.1016/S0140-6736(18)30207-1

3. Finn RS, Qin S, Ikeda M, et al. Atezolizumab plus bevacizumab in unresectable hepatocellular carcinoma (IMbrave150). *New England Journal of Medicine*. 2020;382(20):1894–1905. DOI: 10.1056/NEJMoa1915745

4. Llovet JM, Castet F, Heikenwalder M, et al. Immunotherapies for hepatocellular carcinoma. *Nature Reviews Clinical Oncology*. 2021;19(3):151–172. DOI: 10.1038/s41571-021-00573-2

5. Roayaie S, Blume IN, Thung SN, et al. A system of classifying microvascular invasion to predict outcome after resection in patients with hepatocellular carcinoma. *Gastroenterology*. 2009;137(3):850–855. DOI: 10.1053/j.gastro.2009.06.003

6. Schulze K, Imbeaud S, Letouzé E, et al. Exome sequencing of hepatocellular carcinomas identifies new mutational signatures and potential therapeutic targets. *Nature Genetics*. 2015;47(5):505–511. DOI: 10.1038/ng.3252

7. Lok AS, Sterling RK, Everhart JE, et al. Des-gamma-carboxy prothrombin and alpha-fetoprotein as biomarkers for the early detection of hepatocellular carcinoma. *Gastroenterology*. 2010;138(2):493–502. DOI: 10.1053/j.gastro.2009.10.031

8. Imamura H, Matsuyama Y, Tanaka E, et al. Risk factors contributing to early and late phase intrahepatic recurrence of hepatocellular carcinoma after hepatectomy. *Journal of Hepatology*. 2003;38(2):200–207. DOI: 10.1016/S0168-8278(02)00360-4

9. Ye Q, Ling S, Zheng S, Xu X. Liquid biopsy in hepatocellular carcinoma: circulating tumor cells and circulating tumor DNA. *Molecular Cancer*. 2019;18(1):114. DOI: 10.1186/s12943-019-1043-x

10. Hoshida Y, Villanueva A, Kobayashi M, et al. Gene expression in fixed tissues and outcome in hepatocellular carcinoma. *New England Journal of Medicine*. 2008;359(19):1995–2004. DOI: 10.1056/NEJMoa0804525

11. Ma L, Wang L, Khatib SA, et al. Single-cell atlas of tumor cell evolution in response to treatment in hepatocellular carcinoma and intrahepatic cholangiocarcinoma. *Journal of Hepatology*. 2021;75(6):1397–1408. DOI: 10.1016/j.jhep.2021.06.028

12. Ruiz de Galarreta M, Bresnahan E, Molina-Sánchez P, et al. β-Catenin activation promotes immune evasion and resistance to anti-PD-1 therapy in hepatocellular carcinoma. *Cancer Discovery*. 2019;9(8):1124–1141. DOI: 10.1158/2159-8290.CD-19-0074

HCC-METASCORE is an agent-executable research framework. It is not a medical device, a validated clinical prediction tool, or a substitute for specialist oncological assessment.