Tribunal Session Summary

Session: conclave-20260514-153651-over-the-next-five-4497

Depth: T6

Advocates: 5 | Judges: 5 | Cost: $6.1024 | Time: 33m 44s

THE PROMPT: Over the next five years, which type of cancer is expected to see the most significant increase in survival rates? Please include a probability estimate, key factors influencing this change, and what evidence would most alter your view.

THE RULING: Non-small cell lung cancer (NSCLC) is expected to see the most significant absolute increase in 5-year survival rates over the next five years, with a 60% probability of achieving at least a 5 percentage-point gain by 2031.

The Prompt

Over the next five years, which type of cancer is expected to see the most significant increase in survival rates? Please include a probability estimate, key factors influencing this change, and what evidence would most alter your view.

The Ruling

Non-small cell lung cancer (NSCLC) is expected to see the most significant absolute increase in 5-year survival rates over the next five years, with a 60% probability of achieving at least a 5 percentage-point gain by 2031.

This answer is based on a synthesis of the strongest evidence from the deliberation, which shows NSCLC's unique advantage: its key therapeutic advances—perioperative immunotherapy, targeted therapies, and immunotherapy in advanced disease—were deployed early enough (2020–2024) to be captured in the SEER 2031 release, which reflects diagnoses from approximately 2019 to 2026. This timing advantage is decisive. While other cancers like pancreatic and bladder have biologically transformative therapies, they will not register in population-level 5-year survival statistics by 2031 due to data lags and late adoption. The 60% probability reflects high confidence in the durability of these therapies, tempered by uncertainty around screening uptake and real-world adherence.

Core Analysis

Hypothesis

Non-small cell lung cancer (NSCLC) will achieve the largest absolute increase in 5-year relative survival among major cancers between 2026 and 2031, with a 60% probability of a ≥5 percentage-point gain, driven by the convergence of four therapeutic and epidemiological levers already in clinical practice.

Probability Estimate

60% probability that NSCLC will achieve a ≥5 percentage-point absolute increase in 5-year relative survival by 2031, with a central estimate of +4 to 7 percentage points. This probability is anchored in the maturity and diffusion of existing therapies, not speculative future breakthroughs.

Key Factors Influencing This Change

1. Perioperative Immunotherapy in Resectable NSCLC

- **Mechanism**: Neoadjuvant and adjuvant immune checkpoint inhibitors (e.g., nivolumab, pembrolizumab) are now standard for resectable NSCLC (stages I–IIIA), reducing recurrence and converting relapses into

long-term remissions.

- **Evidence**: CheckMate 816 (neoadjuvant nivolumab + chemo) demonstrated a pathological complete response rate of 24% vs. 2.2% with chemo alone and an event-free survival (EFS) hazard ratio (HR) of 0.63. KEYNOTE-671 (neoadjuvant pembrolizumab + chemo) showed a 5-year overall survival (OS) of 64.6% vs. 53.6% for chemo alone.
- **Contribution**: +0.9 to +1.5 percentage points to all-stage 5-year survival.
- **Why it matters**: This is the largest contributor to the projected gain. Even with conservative assumptions about real-world translation, the EFS benefit is expected to mature into a durable OS gain, lifting the population survival curve.

2. Expansion of Targeted Therapies

- **Mechanism**: The fraction of NSCLC patients with actionable driver mutations (e.g., EGFR, KRAS G12C, ALK) continues to grow, and newer-generation drugs extend survival in both metastatic and adjuvant settings.
- **Evidence**: ADAURA (adjuvant osimertinib for EGFR-mutant NSCLC) reduced recurrence by 80% in stage IB–IIIA, with a 5-year OS benefit of ~10 percentage points. KRAS G12C inhibitors (sotorasib, adagrasib) are approved for ~13% of lung adenocarcinomas.
- **Contribution**: +1.5 percentage points to all-stage survival.
- **Why it matters**: Broad genomic profiling and therapeutic matching are now guideline-standard, meaning most patients with actionable mutations receive targeted agents, shifting survival across the population.

3. Advanced Immunotherapy in Metastatic NSCLC

- **Mechanism**: First-line immunotherapy combinations (e.g., pembrolizumab + chemo, nivolumab + ipilimumab) have become the standard of care for metastatic NSCLC, extending median OS and creating a durable tail of long-term survivors.
- **Evidence**: KEYNOTE-189 reported a 5-year OS rate of 19.4% vs. 11.3% for chemo alone. CheckMate 227 showed a 5-year OS of 24% vs. 14% for chemo in PD-L1 ≥1% patients.
- **Contribution**: +2.0 percentage points to all-stage survival.
- **Why it matters**: Metastatic disease accounts for ~50% of NSCLC cases. The incremental diffusion and quality-of-care gains of these regimens from 2019 to 2024 will be captured in SEER 2031, adding a measurable lift.

4. Modest Improvements in Low-Dose CT Screening Uptake

- **Mechanism**: Increased adoption of low-dose CT (LDCT) screening shifts diagnosis toward earlier, more treatable stages, where 5-year survival exceeds 60%.
- **Evidence**: The National Lung Screening Trial (NLST) showed a 20% reduction in lung-cancer–specific mortality. CISNET modeling indicates that a 10-percentage-point increase in LDCT uptake shifts diagnoses to stages I–II, yielding a 1.1–1.7 percentage-point gain in 5-year survival.
- **Contribution**: +1.0 to +1.4 percentage points to all-stage survival.
- **Why it matters**: Current US screening uptake is ~18% (NHIS 2022). Incremental improvements to 20–25% by 2025, driven by CMS bundled payments and HEDIS quality measures, will produce a modest but real stage shift.

Stage-Weighted Arithmetic

The projected +4 to +7 percentage-point gain is derived from a stage-weighted sum of the four levers, using conservative assumptions about real-world uptake and attenuation:

- **Resectable (30% of cases)**: 10% absolute 5-year OS gain from perioperative IO × 30% uptake = +0.9 percentage points.
- **Unresectable Stage III (15% of cases)**: 9.5% absolute 5-year OS gain from PACIFIC durvalumab × 80% uptake = +1.4 percentage points.
- **Metastatic (55% of cases)**: 8% absolute 5-year OS gain from chemo-IO × 90% uptake (accounting for prior diffusion) = +4.4 percentage points, attenuated to +2.0 points to avoid double-counting baseline therapies.
- **Screening stage shift**: 2–3% shift to earlier stages × 55-point survival gradient = +1.1 to +1.7 percentage points, attenuated to +1.0 to +1.4 points for real-world adherence.
- **Total**: +0.9 + 1.4 + 2.0 + 1.4 = **+5.7 percentage points** (central estimate ~6 points).

Why NSCLC Wins: The Timing Advantage

The decisive factor is not biological superiority but **measurement infrastructure**. The SEER 2031 release will reflect diagnoses from approximately 2019 to 2026. NSCLC's key therapies—KEYNOTE-189 (2018), ADAURA (2020), CheckMate 816 (2022), and KEYNOTE-671 (2023)—were approved early enough for their diffusion to be captured in this window. By contrast:

- **Pancreatic cancer**: Daraxonrasib, a pan-RAS inhibitor with a hazard ratio of 0.40, is not expected to be standard of care until 2027–2028. Diagnoses from 2027 onward will not appear in SEER 2031, making its impact invisible.
- **Bladder cancer**: Enfortumab vedotin + pembrolizumab (EV+P), with a hazard ratio of 0.47, became NCCN-preferred in 2024. The earliest cohort with broad exposure is 2024, which will only partially appear in SEER 2031.

What Evidence Would Most Alter This View

- **Negative perioperative IO OS data**: If 5-year OS updates from KEYNOTE-671 or CheckMate 816 fail to confirm the EFS benefit, the largest lever for NSCLC would collapse, reducing the projected gain to <3 points.
- **SEER 2027 or 2028 releases showing NSCLC gains <3 points**: Early data indicating stagnation would undermine confidence in the current trajectory.
- **Frontline daraxonrasib Phase 3 trial in PDAC with HR <0.50 and a durable survival tail**: This could position PDAC for future leadership, though not within the 2031 window.
- **RWD showing >50% real-world attenuation of EV+P benefit in bladder cancer**: This would further reduce bladder's already limited population impact.

How We Got Here

Scorecard

Judge	Magistral Medium	DeepSeek V4 Pro	Gemini 3.1 Pro	Claude Opus 4.7	GPT-5	Winner
DeepCogito v2.1 671B	REJECT	ACCEPT (modified)	REJECT	REJECT	REJECT	DeepSeek V4 Pro
Zhipu GLM-5	REJECT	ACCEPT	REJECT	ACCEPT (secondary)	ACCEPT (secondary)	DeepSeek V4 Pro
Mistral Large 3	REJECT	ADOPT	REJECT	ADOPT	REJECT	SYNTHESIZE (C + D)
GPT-OSS 120B	REJECT	ACCEPT	REJECT	REJECT	REJECT	DeepSeek V4 Pro
Qwen 3 235B (Cerebras)	REJECT	ACCEPT	REJECT	REJECT	REJECT	DeepSeek V4 Pro

Convergence Assessment

The convergence was epistemic, driven by verifiable corrections. Key evidence includes:

- Claude Opus 4.7’s challenge on SEER timing, which all advocates acknowledged and incorporated.
- DeepSeek V4 Pro’s provision of stage-weighted arithmetic, which became the new standard for quantification.
- Gemini 3.1 Pro’s concession on PDAC’s 5-year survivor math, which was based on a chronological impossibility. Position stability data shows that advocates revised their claims in response to these corrections, not social pressure.

Closing Statements

GPT-5 My final position is that NSCLC will post the largest absolute increase in all-stage 5-year survival, with a central estimate of +2 to +4 percentage points in the SEER-2031 release. I assign a 50% probability to NSCLC leading, 25% to bladder cancer, 15% to multiple myeloma, and 10% to other cancers. What changed from my opening statement is the recognition that the SEER measurement window limits the visibility of recent breakthroughs. My strongest surviving evidence is the stage-weighted arithmetic for NSCLC, which shows that mature gains from ADAURA, PACIFIC, and metastatic IO, combined with modest screening improvements, will lift survival by +2–4 points. I found the challenges on temporal compression and SEER lag most challenging, as they forced me to re-anchor my forecast to the measurement reality rather than the biological promise.

Gemini 3.1 Pro My final position is that NSCLC will generate the most significant absolute increase in 5-year survival rates by 2031, yielding a conservatively estimated, stage-weighted all-stage gain of ~6.1 percentage points. I learned from the debate that trial hazard ratios do not equal population survival deltas, and that a clinical breakthrough extending life from 8 to 15 months is an epidemiological non-factor for a 5-year survival metric. My strongest surviving evidence is the stage-weighted arithmetic model, which shows that perioperative IO, PACIFIC, and metastatic IO contribute +8.8 points before a 30% real-world discount. I found Claude Opus 4.7's challenge on PDAC's 5-year survivor math most challenging, as it forced me to confront the stage-weighted arithmetic and abandon a speculative position.

Claude Opus 4.7 My final position is that NSCLC is the most likely cancer to show the largest measurable increase in 5-year survival rates, with a projected gain of +2 to +4 percentage points under SEER-2031 and +3 to +6 under real-world data. I assign ~40–45% probability to NSCLC leading. I learned that measurement substrate matters as much as biology, and that trial hazard ratios do not equal population survival deltas. My strongest surviving evidence is the SEER lag architecture, which shows that therapies approved after 2023 contribute minimally to the 2031 readout. I found GPT-5's timing and arithmetic critique most challenging, as it forced me to confront the gap between Phase 3 readouts and SEER capture and abandon my initial PDAC hypothesis.

Magistral Medium My final position is that NSCLC is expected to see the most significant increase in survival rates, with a revised probability estimate of 55%. I learned that real-world effectiveness may be lower than clinical trial results due to patient frailty and treatment discontinuations. My strongest surviving evidence is the advancements in targeted therapies and immunotherapies, which continue to improve survival rates. I found the challenges on screening adoption barriers most challenging, as they forced me to adopt a more nuanced view of the potential impact of these therapies.

Reality Checks

Claim	Challenge	Result
Gemini 3.1 Pro: "EV+P guarantees a mathematical step-function increase in 5-year survival"	**DeepSeek V4 Pro:** A median OS of 31.5 months does not determine the 5-year survival rate; the tail beyond the median is unknown.	Full concession: Gemini 3.1 Pro acknowledged that extrapolating a median to a tail without mature data was flawed epidemiology.
Claude Opus 4.7: "Daraxonrasib will move 5-year survival from 13% to >20% by 2031"	**DeepSeek V4 Pro:** The 2031 SEER 5-year survival statistic will be based on patients diagnosed in 2021–2024, before daraxonrasib is standard.	Full concession: Claude Opus 4.7 admitted the timeline incompatibility and revised the probability of PDAC leading down to ~10–15%.
GPT-5: "Avelumab maintenance confers a population-level OS benefit that compounds with EV+pembro uptake"	**Gemini 3.1 Pro:** Avelumab maintenance requires induction chemotherapy, while EV+P replaces it; the pathways are mutually exclusive.	Full concession: GPT-5 acknowledged the contradiction and removed "compounding" from the framework.

Next Steps

Next Step	Why It Matters	Confidence
Monitor 5-year OS updates from KEYNOTE-671 and CheckMate 816	This will determine if perioperative IO's EFS benefit translates to durable cures, the largest lever for NSCLC's gain.	70%
Track LDCT screening uptake in the US via NHIS or SEER	This will determine the stage-shift contribution; current trends suggest a plateau near 18–20%.	50%
Watch for frontline daraxonrasib Phase 3 readouts in PDAC	This could position PDAC for future leadership, though not within the 2031 window.	30%
Evaluate real-world OS data from Flatiron/Tempus for 2024–2025 NSCLC cohorts	This will show if perioperative IO gains are being realized in community practice, potentially exceeding SEER estimates.	60%

Dissenting Opinions

No dissenting opinions were filed in this session.

Glossary

Term	Definition
5-year relative survival	A measure of survival that compares the observed survival of cancer patients to the expected survival of the general population, used to account for deaths from other causes.
Adjuvant therapy	Treatment given after primary therapy (e.g., surgery) to reduce the risk of cancer recurrence.
Event-free survival (EFS)	The length of time after primary treatment during which no signs of cancer are present and no new events (e.g., recurrence, death) occur.
Hazard ratio (HR)	A measure of the effect of an intervention on an outcome over time, where HR < 1 indicates a reduction in risk of the event (e.g., death).
Low-dose CT (LDCT) screening	A type of computed tomography scan that uses low radiation doses to detect lung cancer at early stages in high-risk individuals.
Neoadjuvant therapy	Treatment given before primary therapy (e.g., surgery) to shrink a tumor and make it easier to remove.
Overall survival (OS)	The length of time from either diagnosis or the start of treatment for a disease, such as cancer, that patients are still alive.
Perioperative immunotherapy	The use of immune checkpoint inhibitors around the time of surgery, either before (neoadjuvant) or after (adjuvant), to improve outcomes in resectable cancers.
SEER	The Surveillance, Epidemiology, and End Results program, a U.S. National Cancer Institute initiative that collects and publishes cancer incidence and survival data from population-based cancer registries.
Stage shift	A change in the distribution of cancer diagnoses toward earlier stages, often due to screening, which improves survival outcomes.

How The Conclave Works

The Conclave is a structured multi-model deliberation system. Instead of asking one AI model a question and trusting its answer, the Conclave convenes a panel of independent models, forces them to argue, and subjects their conclusions to judicial review. The process is adversarial by design — consensus must be earned through evidence, not assumed through agreement.

Advocates are independent AI models (e.g., Claude, GPT-5, Gemini, DeepSeek) that each receive the same question and produce a sealed submission without seeing each other's work. Each submission must include a falsifiable hypothesis, tagged evidence with reasoning type (deductive/inductive/abductive), provenance labels for any frameworks cited, and an honest self-assessment.

Challenges follow submissions. Each advocate reads all other submissions and directly attacks weak points — not polite reviews, but pointed cross-examination. They must identify the single factual crux that would settle each disagreement.

Debate Rounds (1–7 depending on depth) force advocates to defend, concede, or revise their positions under pressure. The system tracks position stability across rounds to detect sycophantic drift — when models change their stance to agree with others without citing new evidence.

Judges (called Justices on The Bench) are separate models that never participated as advocates. They evaluate the full record: submissions, challenges, and debate transcripts. They fact-check claims, audit framework provenance, assess whether convergence was epistemic (evidence-driven) or affective (social pressure), and render a verdict: ACCEPT one position, SYNTHESIZE the best elements, or REMAND for further debate.

Depth levels control rigor: T1/Spot Check (submissions only), T2/Standard Review (1 debate round, 1 judge), T3/Deep Review (3 rounds, 3 judges), T4/Full Panel (5 rounds, full panel), T5/Stress Test (5 rounds + Fresh Eyes review), T6/Red Team (7 rounds + mid-debate judicial checkpoint).

The system is deterministic code — it cannot be sycophantic. It dispatches prompts, collects responses, and enforces the adversarial structure. The models argue; the code referees.