Tribunal Session Summary

Session: conclave-20260514-191556-over-the-next-five-255e

Depth: T6

Advocates: 6 | Judges: 5 | Cost: $5.3583 | Time: 29m 41s

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 developments could alter this forecast.

THE RULING: Lung cancer is expected to see the most significant absolute increase in 5-year relative survival rates over the 2026–2031 diagnosis period, with a 50% probability and a projected gain of +4 to +7 percentage points (from ~28% to 32–35%).

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 developments could alter this forecast.

The Ruling

Lung cancer is expected to see the most significant absolute increase in 5-year relative survival rates over the 2026–2031 diagnosis period, with a 50% probability and a projected gain of +4 to +7 percentage points (from ~28% to 32–35%).

This answer is based on the convergence of evidence from the American Cancer Society's 2026 report and the maturation of perioperative immunotherapy regimens (e.g., CheckMate-816, KEYNOTE-671) that were not widely available during the 2015–2021 baseline period. The strongest evidence is the documented 13-point historical gain in lung cancer survival (15% to 28%) and the 8-point gain in distant-stage survival (2% to 10%), which establishes both momentum and headroom. The main caveat is the significant downside risk from federal funding and insurance access cuts, which the ACS CEO explicitly warned could "reverse this progress," and which justifies the 50% probability rather than a higher estimate.

Core Analysis

Thesis Statement

Lung cancer is the most likely cancer to achieve the largest absolute improvement in 5-year relative survival for patients diagnosed between 2026 and 2031, with a 50% probability and a projected gain of +4 to +7 percentage points. This forecast is driven by the population-level impact of newly approved perioperative immunotherapy regimens and modest gains from expanded low-dose CT (LDCT) screening, but it is tempered by substantial risks from federal funding and insurance access cuts that could stall progress.

Market Opportunity and Value Creation Mechanism

The "market" in this context is the population of cancer patients in the United States, with lung cancer representing the largest single source of cancer mortality. The American Cancer Society's *Cancer Statistics, 2026* report establishes that lung cancer will cause more deaths in 2026 than colorectal and pancreatic cancers combined, and its 5-year relative survival rate for people diagnosed during 2015–2021 was 28%, up from 15% in the mid-1990s. This combination of high incidence (~235,000 new cases per year) and a low baseline survival rate creates the largest potential for absolute population impact. A 10-point gain in lung cancer survival saves more lives than a 20-point gain in a rarer cancer like myeloma, which is why "most significant" is interpreted as the largest absolute percentage-point increase in population-weighted 5-year relative survival.

The value creation mechanism is the additive effect of multiple, already-approved interventions maturing their impact within the 2026–2031 window. The key insight is that the 28% baseline from the 2015–2021 cohort does not include the full effect of perioperative immunotherapy (CheckMate-816 approved in 2022, KEYNOTE-671 in 2023), which is now entering routine practice. This is not a speculative future therapy; it is a current standard of care whose 5-year overall survival (OS) data will mature for the 2026 diagnosis cohort by 2031. The value is created by the diffusion of this proven therapy into the community oncology setting, combined with incremental gains from metastatic therapies and a modest stage shift from LDCT screening.

Unit Economics and Competitive Advantages

The unit economics of the survival gain are broken down into population-weighted contributions from four key drivers:

1. **Perioperative Immunotherapy (1.0–2.0 percentage points):** This applies to the ~25% of lung cancer patients diagnosed with resectable (stage I–IIIA) disease. Trials like CheckMate-816 and KEYNOTE-671 have shown event-free survival hazard ratios (HR) of ~0.58–0.63. Assuming a real-world adoption rate of 50–60% among eligible patients and a conservative 8–10 point absolute OS benefit, the contribution to the all-stage survival rate is 0.25 × 0.55 × 9 = 1.2 percentage points, with a range of 1.0–2.0 points accounting for variability in adoption and benefit.
2. **Adjuvant EGFR Inhibition (0.4–0.6 percentage points):** For the ~10–15% of non-small cell lung cancer (NSCLC) patients with EGFR mutations, the ADAURA trial demonstrated a 5-year OS benefit of 10 percentage points (88% vs. 78%). This regimen is now standard of care. The population-weighted contribution is 0.12 × 0.8 × 10 = 0.96 points, but given that not all patients are tested or eligible, a conservative estimate of 0.4–0.6 points is used.
3. **Metastatic Therapy Maturation (2.0–3.0 percentage points):** The distant-stage cohort (~53% of cases) has seen its 5-year survival rise from 2% to 10% since the mid-1990s, driven by first-line chemo-immunotherapy, next-generation tyrosine kinase inhibitors (TKIs), and antibody-drug conjugates (ADCs). An additional 5–8 point gain in this cohort over the next five years is plausible, based on the acceleration observed in the last decade. The population-weighted contribution is 0.53 × 6 = 3.18 points, but a conservative range of 2.0–3.0 points is used to reflect the uncertainty of translating trial results to real-world populations.
4. **LDCT Screening Stage Shift (1.0–1.5 percentage points):** The National Lung Screening Trial (NLST) demonstrated a 20% reduction in lung cancer mortality with LDCT. Current uptake is estimated at 16–18% of eligible individuals. A realistic projection is an increase to 22–25% by 2031, driven by the 2021 USPSTF guideline expansion and AI-assisted nodule detection. This is expected to shift ~4% of cases from distant to earlier stages. Given the survival gap (distant: 10%, regional: 37%, localized: 65%), this contributes approximately 1.0–1.5 percentage points to the all-stage survival rate.

The sum of these components is 4.4–7.1 percentage points, which rounds to the final projection of +4 to +7 points.

Lung cancer's competitive advantage over other high-potential cancers is its combination of scale and timing. Myeloma, with its 32% to 62% survival gain, is the strongest competitor, but its transformative therapies (CAR-T, bispecific antibodies) face significant diffusion bottlenecks—manufacturing capacity, toxicity management, and eligibility (median age ~70)—that limit rapid, broad adoption within a five-year window. Liver cancer, with its 7% to 22% gain, is constrained by the slow pace of screening adoption and the complexity of managing cirrhosis. Melanoma, with its 16% to 35% distant-stage gain, already has a very high overall survival rate (~94%), so the absolute headroom for improvement is small. Lung cancer, starting from

a lower baseline and with therapies already approved and diffusing, has a more favorable risk-reward profile.

Key Risks and Mitigations

The primary risk to this forecast is **federal funding and insurance access cuts**. The ACS CEO, Shane Jacobson, explicitly warned that such cuts could "reverse this progress and stall future breakthroughs." This risk is asymmetric: it disproportionately affects lung cancer because its gains rely heavily on expensive, multi-year treatment sequences (immunotherapy, ADCs) and a public-health infrastructure for LDCT screening. A contraction in Medicaid or the ACA could reduce access for high-incidence rural populations, stalling both screening uptake and therapy diffusion. The 50% probability assigned to lung cancer's leadership is a direct reflection of this risk, pricing it as a 50/50 coin flip.

A secondary risk is **immunotherapy resistance**. If a higher-than-expected percentage of patients develop hyper-progressive disease or acquired resistance to perioperative immunotherapy, the projected 5-year survival curve will flatten prematurely. This risk is mitigated by the breadth of the therapeutic pipeline—KRAS inhibitors, ADCs, and bispecific T-cell engagers provide multiple "shots on goal" that reduce the impact of any single failure.

A third risk is **slower-than-expected adoption of perioperative regimens**. The projected gain assumes 50–60% adoption among eligible patients, but real-world data show disparities in biomarker testing, specialist referral, and treatment initiation, particularly in community and rural settings. This risk is mitigated by the fact that these regimens use widely available PD-1/PD-L1 antibodies, which are already scaled in metastatic NSCLC, making diffusion easier than for complex therapies like CAR-T.

Implementation Roadmap

The forecast will be validated through the following milestones:

1. **2026–2027:** Monitor real-world adoption rates of perioperative immunotherapy in national cancer registries (e.g., SEER, NCDB). A rate of >40% among eligible patients by 2027 would confirm the upper end of the adoption range.
2. **2027–2029:** Track the maturation of 5-year OS data from perioperative trials (KEYNOTE-671, CheckMate-816, AEGEAN). A hazard ratio for OS of <0.8 would validate the translation from event-free survival to overall survival.
3. **2026–2031:** Measure LDCT screening uptake through the American Lung Association's annual "State of Lung Cancer" report. An increase to >22% by 2028 would support the stage-shift projection.
4. **Ongoing:** Monitor federal policy for any enacted cuts to the National Cancer Institute (NCI) budget or changes to the ACA that could affect insurance coverage for cancer screening and treatment.

Exit Paths

The forecast will be considered successful if, for the cohort of patients diagnosed between 2026 and 2031, the 5-year relative survival rate for lung cancer increases by at least 4 percentage points. A gain of 7 or more points would indicate that the upper bound of the projection was achieved. The forecast will be considered a failure if the gain is less than 4 points, which would suggest that the downside risks (funding cuts, slow adoption) dominated the upside drivers.

How We Got Here

Scorecard

Judge	Claude Opus 4.7	Grok 4	Magistral Medium	GPT-5	DeepSeek V4 Pro	Gemini 3.1 Pro	Winner
MiniMax M2.7	ADOPT	ADOPT	REJECT	ADOPT	ACCEPT	ADOPT	DeepSeek V4 Pro
Qwen 3 235B (Cerebras)	ACCEPT	ADOPT	REJECT	ADOPT	REJECT	ADOPT	Claude Opus 4.7
DeepCogito v2.1 671B	REJECT	REJECT	REJECT	REJECT	ACCEPT	REJECT	DeepSeek V4 Pro
Zhipu GLM-5.1	ACCEPT	ADOPT	REJECT	ADOPT	ADOPT	ADOPT	Claude Opus 4.7
GPT-OSS 120B	ACCEPT	REJECT	REJECT	REJECT	REJECT	REJECT	Claude Opus 4.7

Convergence Assessment

The convergence was epistemic, not affective. All advocates changed their positions in direct response to verifiable corrections and challenges. Claude Opus 4.7 retracted the unsourced 5–7% stage-shift elasticity multiplier after being challenged by DeepSeek V4 Pro. Gemini 3.1 Pro conceded the misapplication of the Will Rogers Phenomenon after being corrected by multiple advocates. DeepSeek V4 Pro revised the screening uptake baseline from "<10%" to "16–18%" after being presented with data from the American Lung Association. GPT-5 revised the probability distribution to correct an internal mathematical contradiction after being challenged by Gemini 3.1 Pro. The only advocate who showed signs of affective convergence was Magistral Medium, whose probability dropped from 70% to 60% without a corresponding change in their evidence-based argument, suggesting a possible alignment with the group rather than a deep internal recalibration. The majority of position shifts were triggered by hard evidence, demonstrating that the convergence was driven by the adversarial process, not social pressure.

Closing Statements

Claude Opus 4.7: Claude Opus 4.7's final position was that lung cancer is the single most likely cancer to post the largest absolute percentage-point gain in population-weighted 5-year relative survival over the 2026–2031 diagnosis cohort, with a 50% probability and a projected gain of +4 to +7 points. This was a refinement of their opening 65% probability and +8 to +12 point range, driven by the retraction of the unsourced screening elasticity multiplier and the adoption of a more realistic screening uptake trajectory. Their strongest surviving evidence was the timing of the perioperative immunotherapy approvals (2022–2023), which means their 5-year OS contributions are not in the 2015–2021 ACS denominator but are in the 2026–2031 forecast cohort. They were most challenged by the "friction anchor" argument, which they acknowledged but countered by emphasizing that the same large denominator delivers outsized population impact. Their closing statement was a model of intellectual honesty, explicitly resisting sycophantic drift toward the 60% median and holding firm on their 50% calibration.

Gemini 3.1 Pro: Gemini 3.1 Pro's final position was that lung cancer has a 45% probability of realizing the highest absolute percentage-point increase in 5-year survival rates, with an expected magnitude of +4.0 to +6.0 points. This was a significant revision from their opening 75% probability, driven by the realization that their initial optimism was "mathematically unmoored." Their strongest surviving evidence was the "friction anchor" argument, which they developed by challenging DeepSeek V4 Pro's +7 to +10 point projection. They calculated that even a 15-point OS benefit from perioperative IO in 60% of resectable patients would only add 2.25 points to the all-stage survival rate, demonstrating the structural difficulty of large gains. They were most challenged by Claude Opus 4.7 on the weighting of systemic downside risks, which they defended by arguing that the downside (insurance cuts) affects the 75% metastatic cohort, while the upside (curative-intent IO) is restricted to the 25% resectable pool. Their closing statement showed remarkable intellectual growth, acknowledging that the deliberation process had "stripped away my initial exuberance" but left them with a "deeply resilient, mathematically validated model."

DeepSeek V4 Pro: DeepSeek V4 Pro's final position was that lung cancer will record the largest absolute increase in five-year survival rate with a 60% probability, projecting a +7 to +10 point gain. This was a revision from their opening 75% probability, driven by the retraction of an outdated screening baseline and an unsupported claim about a "demographic tailwind" from smoking cessation. Their strongest surviving evidence was the deterministic stage-shift from increased screening and the additive therapy ledger they built in response to Gemini 3.1 Pro's challenge. They argued that the survival gap between regional and distant disease is so large that even a modest doubling of screening penetration would add 2–3 points to the population survival rate. They were most challenged by Gemini 3.1 Pro on the metastatic therapy contribution, which they defended with a detailed breakdown of trial data and adoption estimates. Their closing statement was a testament to the value of adversarial testing, as they acknowledged that the debate "transformed my claim from a plausible extrapolation into a transparent, testable model."

Reality Checks

Claim	Challenge	Result
Claude Opus 4.7: Each 10-point increase in LDCT screening uptake historically shifts ~5–7% of diagnoses from distant to localized stage.	**DeepSeek V4 Pro:** This appears asserted without provenance and likely overstated. Provide the empirical source linking LDCT uptake increments to a 5–7% stage redistribution at the population level.	**Retracted.** Claude Opus 4.7 conceded the claim was an unsourced back-of-the-envelope estimate and revised the stage-shift contribution downward from 4–6 points to 1.0–1.5 points.
Gemini 3.1 Pro: The synergistic application of systemic therapies to earlier stages results in exponential survival gains (the "Adjuvant Stage-Shift Multiplier").	**DeepSeek V4 Pro:** There is no "exponential" dynamic—survival gains are additive, not multiplicative, on the absolute risk scale.	**Retracted.** Gemini 3.1 Pro conceded the term "exponential" was incorrect and retracted the "Adjuvant Stage-Shift Multiplier" concept, acknowledging it was "rhetoric" without a solid mathematical basis.
Grok 4: Distant-stage lung cancer will see gains exceeding 15 points over the 2026–2031 period.	**Gemini 3.1 Pro:** How do you defend moving distant-stage lung cancer from ~10% to ≥25% 5-year survival in five years, given no phase 3 data to date shows that magnitude of OS improvement at the population level?	**Retracted.** Grok 4 conceded the claim was an over-extrapolation and revised the distant-stage gain projection to +4 to +7 points, acknowledging that no mechanism justifies an 11-fold acceleration in the historical trend.
DeepSeek V4 Pro: Real-world screening rates languish below 10% of eligible individuals.	**Gemini 3.1 Pro:** The 10% figure is an aspirational projection, not a documented current rate—the most reliable real-world estimates place it at ~16–18%.	**Retracted.** DeepSeek V4 Pro conceded the screening baseline was outdated and revised it to 16–18%, which reduced the projected stage-shift contribution from 3–5 points to 2–3 points.
Gemini 3.1 Pro: The Will Rogers Phenomenon explains how LDCT screening improves survival statistics by shifting patients from the "healthy" or "undiagnosed" bucket into the early-stage lung cancer bucket.	**Claude Opus 4.7:** This is wrong. The Will Rogers Phenomenon is about stage reclassification, not detection of previously-undiagnosed disease. What you're describing is lead-time bias and length-time bias.	**Retracted.** Gemini 3.1 Pro conceded the error, acknowledging it was a "textbook epidemiological error" and that the correct terms are lead-time and length-time bias.

Next Steps

Next Step	Why It Matters	Confidence
Monitor real-world adoption rates of perioperative immunotherapy in national cancer registries (e.g., SEER, NCDB).	This is the single largest source of upside for lung cancer. If adoption is >40% by 2027, the upper end of the +4 to +7 point gain is achievable.	70%

Next Step	Why It Matters	Confidence
Track the maturation of 5-year OS data from perioperative trials (KEYNOTE-671, CheckMate-816, AEGEAN).	A hazard ratio for OS of <0.8 would validate the translation from event-free survival to overall survival, confirming the projected gain.	70%
Measure LDCT screening uptake through the American Lung Association's annual "State of Lung Cancer" report.	An increase to >22% by 2028 would support the stage-shift projection of 1.0–1.5 percentage points.	60%
Monitor federal policy for any enacted cuts to the National Cancer Institute (NCI) budget or changes to the ACA.	This is the primary downside risk. If cuts occur, they could stall both LDCT screening and perioperative IO diffusion, reducing the gain to the lower end of the range.	50%
Assess the real-world impact of metastatic therapy advances (ADCs, KRAS inhibitors) on survival.	This is the second-largest source of gain. If the metastatic cohort sees a 5–8 point survival improvement, it will contribute 2.0–3.0 percentage points to the all-stage rate.	60%

Dissenting Opinions

No dissenting opinions were filed in this session.

Glossary

Term	Definition
5-year relative survival	A standard oncology metric that estimates the percentage of patients who will live at least 5 years after diagnosis compared to people without the disease.
Adjuvant therapy	Treatment given after primary surgery to reduce the risk of cancer recurrence.
Antibody-Drug Conjugates (ADCs)	A class of targeted biopharmaceutical drugs that combine monoclonal antibodies with cytotoxic agents to deliver chemotherapy directly to cancer cells.
Event-Free Survival (EFS)	A clinical trial endpoint that measures the time from treatment until the cancer progresses, the patient dies, or another event occurs.
Hazard Ratio (HR)	A measure of how often a particular event happens in one group compared to another, commonly used in survival analysis.
Low-Dose CT (LDCT)	A type of computed tomography scan that uses lower radiation doses to screen for lung cancer in high-risk individuals.
Neoadjuvant therapy	Treatment given before primary surgery to shrink a tumor and make it easier to remove.
Objective Response Rate (ORR)	A standard oncological metric defining the percentage of patients whose tumor is destroyed or significantly reduced by a drug.
Overall Survival (OS)	A clinical trial endpoint that measures the length of time patients live after treatment.
Perioperative immunotherapy	The use of immunotherapy drugs both before (neoadjuvant) and after (adjuvant) surgery to treat cancer.
Will Rogers Phenomenon	An epidemiological paradox where moving an item from one set to another increases the average values of both sets, often due to improved detection or staging.

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.