Tribunal Session Summary

Session: conclave-20260514-123755-over-the-next-five-01bd

Depth: T6

Advocates: 6 | Judges: 5 | Cost: $5.3147 | Time: 35m 18s

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 the most likely to achieve the largest absolute increase in all-stage 5-year survival over the next five years, with a 52% probability, but multiple myeloma remains a close contender at ~35% probability; the evidence does not support a definitive winner, and the most accurate forecast lies in synthesizing the rigorous quantitative model for NSCLC with the strong mechanistic case for myeloma.

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 the most likely to achieve the largest absolute increase in all-stage 5-year survival over the next five years, with a 52% probability, but multiple myeloma remains a close contender at ~35% probability; the evidence does not support a definitive winner, and the most accurate forecast lies in synthesizing the rigorous quantitative model for NSCLC with the strong mechanistic case for myeloma.

The ruling synthesizes the Tribunal's findings: NSCLC edges ahead due to its broad, population-wide therapeutic advances already embedded in community practice, including durvalumab consolidation in stage III disease, perioperative immunotherapy and targeted therapies, and incremental gains in metastatic precision oncology. These drivers, combined with a modest contribution from low-dose CT screening (half of which is lead-time bias), support a projected gain of +3.0–5.4 percentage points. Multiple myeloma presents a compelling counter-case with potentially larger per-patient gains from CAR-T and bispecific antibodies, but its population-level impact is constrained by manufacturing bottlenecks and limited access, making its projected +6–8 point gain less probable. The synthesis reflects the Tribunal's assessment that while NSCLC is the marginal favorite, the gap is narrow and the outcome uncertain.

Core Analysis

The Synthesized Forecast: NSCLC as Marginal Favorite

The Tribunal's verdict is a synthesis of the most rigorous evidence presented, primarily from GPT-5's quantitative decomposition and DeepSeek V4 Pro's mechanistic case for myeloma. The analysis concludes that **non-small cell lung cancer (NSCLC) is the most likely candidate** to see the largest absolute increase in all-stage 5-year survival over the next five years, with a **52% probability**. **Multiple myeloma is a close second**, with a **~35% probability**. The projected gain for NSCLC is **+3.0 to +5.4 percentage points**, moving from a current baseline of ~25% to ~28–30% by 2029. For myeloma, the projected gain is **+6–8 percentage points**, from ~55% to ~61–63%. The synthesis acknowledges that while myeloma's per-patient gains are larger, NSCLC's broader applicability across a larger patient population and more mature real-world adoption make it the marginal favorite.

Key Factors Driving the Change in NSCLC

The projected increase in NSCLC survival is driven by four key, evidence-supported mechanisms, each with a quantified contribution to the overall gain.

1. Maturation of Durvalumab in Stage III (PACIFIC Regimen): +0.4–0.7 Points

The Phase III PACIFIC trial established one year of durvalumab (an anti-PD-L1 immunotherapy) following chemoradiotherapy as the standard of care for unresectable Stage III NSCLC. The mature 5-year overall survival (OS) data shows a significant benefit: 42.9% for the durvalumab arm versus 33.4% for the placebo arm. This represents an absolute gain of ~9–10 points within the stage III cohort.

However, the Tribunal's analysis, refined through adversarial debate, recognized that much of this benefit was already being captured in the current SEER baseline. The FDA approved durvalumab in 2018, and real-world uptake rose to ~70–80% of eligible patients by 2020–2022. The current SEER baseline (~25%) is derived from cohorts diagnosed between 2013 and 2019, which only partially captured the PACIFIC benefit. The incremental gain over the next five years comes from the full maturation of the survival curves for patients diagnosed in 2018–2021 and a small amount of residual diffusion. After accounting for this, the Tribunal concluded that the net new contribution to the population-level 5-year survival metric is **+0.4 to +0.7 percentage points**.

2. Scaling of Perioperative Immunotherapy and Targeted Therapy: +0.6–1.2 Points

A major driver of future gains is the shift of effective therapies into earlier, curative-intent settings. Two landmark trials form the basis of this advance.

- **Adjuvant Osimertinib (ADAURA):** For patients with resected, EGFR-mutant NSCLC, the ADAURA trial demonstrated a definitive overall survival benefit. The hazard ratio (HR) for death was ~0.49, meaning a 51% reduction in the risk of death. This is a mature OS signal, not a surrogate endpoint.
- **Neoadjuvant Nivolumab + Chemotherapy (CheckMate-816):** This trial showed a significant improvement in event-free survival (EFS) and a high rate of pathologic complete response (pCR), a strong surrogate for long-term survival in this setting.

These therapies are now standard for their respective patient subgroups. The Tribunal estimated that stages II–III resectable disease accounts for roughly 15–20% of all NSCLC cases. With a projected 40–60% real-world adoption of perioperative immunotherapy or targeted therapy by 2029, and partial OS maturation visible in the reporting window, this component is expected to contribute **+0.6 to +1.2 percentage points** to the overall survival gain.

3. Metastatic Precision Oncology and Immunotherapy: +1.2–2.0 Points

NSCLC has the most mature and broadest precision oncology footprint of any common solid tumor. Since 2018, approvals have expanded to cover over 30% of NSCLC cases with targeted agents for mutations in EGFR, ALK, ROS1, RET, MET exon14, BRAF, HER2, and KRAS G12C. Combined with immunotherapy backbones (e.g., pembrolizumab + chemotherapy), this creates a highly effective treatment ecosystem for metastatic disease.

While individual trial benefits are modest (e.g., sotorasib in KRAS G12C showed an OS HR of ~0.66), the cumulative effect of these incremental improvements across multiple molecular subtypes is significant. As these therapies are used earlier in the treatment sequence and adoption deepens, they are expected to contribute **+1.2 to +2.0 percentage points** to the population-level survival gain.

4. Low-Dose CT (LDCT) Screening: +0.8–1.5 Points (Half Lead-Time Bias)

The expansion of USPSTF guidelines in 2021 broadened eligibility for LDCT screening, which can detect lung cancer at an earlier, more curable stage. This is a structural shift that will mechanically increase 5-year survival rates.

However, the Tribunal, following challenges from multiple advocates, rigorously addressed the issue of **lead-time bias**. The 2024 *Lancet Oncology* modeling study by Patz et al. estimated that approximately 50% of the apparent survival gain from screening is an artifact of earlier diagnosis, not a true extension of life. Furthermore, real-world uptake remains stubbornly low, at ~12–16% of eligible adults (per Fedewa et al., JAMA Oncol 2024), with only modest projected growth to ~18–22% by 2029.

Therefore, the Tribunal decomposed this contribution: the *reported* 5-year survival metric will see an increase of **+0.8 to +1.5 points**, but the *true mortality reduction* component is estimated at **+0.4 to +0.8 points**. The screening contribution is real but modest and partially artificial.

Key Factors Driving the Change in Multiple Myeloma

The case for multiple myeloma as the largest gainer is built on the potential for deep, durable responses from new immunotherapies, which could translate into a larger absolute survival gain from a lower baseline.

1. Earlier Use of BCMA-Targeted CAR-T and Bispecific Antibodies: +3.8–5.0 Points

The most potent new therapies are chimeric antigen receptor (CAR) T-cells (e.g., cilta-cel) and bispecific antibodies (e.g., teclistamab) targeting BCMA. The CARTITUDE-4 trial showed a 74% reduction in the risk of progression or death (PFS HR 0.26) when cilta-cel was used after 1–3 prior lines. The key challenge is translating this PFS benefit into a population-level OS gain.

The Tribunal's analysis, particularly from Claude Opus 4.7's counter-attack, performed a critical arithmetic check. Assuming a 20% penetration of BCMA therapies in the first two lines of treatment by 2029, and applying a best-case OS hazard ratio of 0.50 (which would raise 5-year OS from 55% to ~74% for treated patients), the maximum population-level gain is calculated as 20% × 19 points = **+3.8 percentage points**. Even with optimistic assumptions, the math cannot support the claimed +6–8 point gain. The Tribunal thus concluded that the plausible contribution from this component is **+3.8 to +5.0 points**, acknowledging the potential but constrained by adoption.

2. Quadruplet Induction Therapy (PERSEUS): +1–2 Points

The PERSEUS trial demonstrated that adding daratumumab to the standard RVd (lenalidomide, bortezomib, dexamethasone) regimen for transplant-eligible patients significantly improved progression-free survival (PFS HR 0.42). This is expected to extend the duration of first remission, directly benefiting the 5-year survival metric. Since this regimen applies to ~30–35% of incident cases, it is projected to contribute an additional **+1 to +2 percentage points** to the overall gain.

Why NSCLC Edges Ahead: The Synthesis

The synthesis rests on the Tribunal's assessment of **real-world impact versus trial efficacy**. NSCLC's advantage lies in the **breadth, maturity, and diffusion** of its therapeutic advances. The key drivers—PACIFIC, ADAURA, and perioperative IO—are already guideline-embedded and diffusing through community oncology networks. The gains are incremental but widespread, affecting a large fraction of the ~230,000 annual U.S. NSCLC cases.

In contrast, myeloma's gains, while potentially larger per patient, are bottlenecked by **severe delivery constraints**. CAR-T therapy requires complex manufacturing (6–8 week vein-to-vein time), specialized centers, and carries risks of severe toxicity (e.g., cytokine release syndrome). Bispecific antibodies, while "off-the-shelf," still require intensive monitoring for infections. Real-world adoption of these therapies in the first-line setting is currently <10%, and scaling to the 20–30% needed to achieve a +6–8 point population gain by 2029 is highly uncertain.

Therefore, the Tribunal concluded that NSCLC's more modest but more certain and broadly applicable gains give it a slight edge, resulting in the 52% probability for NSCLC versus 35% for myeloma.

Evidence That Would Most Alter the View

The Tribunal identified several pieces of evidence that would significantly alter the synthesis.

1. **For NSCLC:** A 2027 interim SEER report showing a >2-point decline in lung cancer mortality or a >3-point increase in 5-year survival would indicate that the projected gains are materializing faster than expected, increasing the probability for NSCLC to ~70%.
2. **For Multiple Myeloma:** Mature overall survival data from the CARTITUDE-4 trial showing a final OS HR < 0.55, coupled with real-world evidence of ≥25% of incident myeloma patients receiving a BCMA-directed therapy within 18 months of diagnosis by 2028, would validate the high-gain scenario and shift the lead to myeloma.
3. **For Pancreatic Cancer:** A Phase III trial demonstrating a hazard ratio of ≤0.60 for a KRAS G12D inhibitor-based regimen in first-line treatment, with a validated, scalable early-detection strategy showing a population stage shift, would make pancreatic cancer a surprise contender, though its historical 90% trial failure rate makes this unlikely.
4. **For NSCLC (Downside Risk):** National registry data showing LDCT screening uptake stagnating below 10% and/or safety or regulatory setbacks that curtail the use of perioperative immunotherapy or targeted therapy would reduce NSCLC's projected gain and tighten the race with myeloma.

How We Got Here

Scorecard

Judge	Claude Opus 4.7	Grok 4	GPT-5	DeepSeek V4 Pro	Magistral Medium	Gemini 3.1 Pro	Winner
DeepCogito v2.1 671B	REJECT	REJECT	ACCEPT	REJECT	REJECT	REJECT	GPT-5
Qwen 3 235B (Cerebras)	REJECT	REJECT	ADOPT	ADOPT	REJECT	REJECT	GPT-5 (SYNTHESIZE)

Judge	Claude Opus 4.7	Grok 4	GPT-5	DeepSeek V4 Pro	Magistral Medium	Gemini 3.1 Pro	Winner
Zhipu GLM-5.1	REJECT	REJECT	ADOPT	ADOPT	REJECT	REJECT	GPT-5 (SYNTHESIZE)
GPT-OSS 120B	ADOPT (reasoning)	REJECT	ACCEPT	REJECT	REJECT	REJECT	GPT-5
Zhipu GLM-5	ADOPT (reasoning)	REJECT	ACCEPT	REJECT	REJECT	REJECT	GPT-5

Convergence Assessment

The convergence was **epistemic**, driven by verifiable corrections and logical challenges, not social pressure. Claude Opus 4.7 changed their position from melanoma to NSCLC after being confronted with the mathematical impossibility of large gains from a 94% baseline. Gemini 3.1 Pro abandoned their entire pancreatic cancer thesis after conceding to multiple advocates that their cited trial data was fabricated. Grok 4 and Claude Opus 4.7dvocate-E both revised their probability estimates downward in response to challenges about screening uptake and CAR-T adoption constraints. GPT-5 refined their PACIFIC contribution based on a specific, data-driven challenge from Claude Opus 4.7. These shifts were not random; they were direct responses to new evidence or the exposure of a logical flaw, demonstrating a genuine pursuit of truth.

Closing Statements

Claude Opus 4.7 concluded that NSCLC is the marginal favorite at ~33% probability, with multiple myeloma a close second at ~28%. They credited GPT-5's "population-weighted decomposition framework" as the correct tool for forecasting SEER metrics and highlighted their own challenge to the "PACIFIC diffusion-residual" as a key surviving argument that trimmed NSCLC's projected gain. They maintained that the "already-priced-in" critique applied to mature immunotherapy markets was a crucial insight, even though it ultimately argued against their own initial thesis.

GPT-5 maintained that NSCLC is the most likely winner with a 52% probability, projecting a gain of +3.0 to +5.4 points. Their strongest surviving evidence was the "population-level mortality declines in lung cancer" already outpacing other cancers, the "perioperative expansion of cures" with OS-mature or highly credible surrogates, and the "breadth of metastatic precision/IO" reaching a uniquely broad fraction of patients. They acknowledged the strength of the myeloma case but remained unpersuaded by the five-year operational runway for widespread CAR-T adoption.

DeepSeek V4 Pro held firm that multiple myeloma will achieve the largest absolute increase, with a 60% probability of a +6–8 point gain. Their strongest surviving evidence was the "CARTITUDE-4 PFS data, reinforced by the MRD–OS correlation," the "historical trajectory of myeloma survival improvement," and "real-world access expansion data" showing growth in CAR-T centers and manufacturing capacity. They dismissed NSCLC's lead-time bias and argued that myeloma's therapeutic portfolio has "breadth and resilience" that distinguishes it from other cancers.

Gemini 3.1 Pro, after fully retracting their initial position, concluded that NSCLC is the most likely winner, adopting GPT-5's framework. Their strongest surviving evidence was the "PACIFIC Baseline," the "Perioperative Shift," and "LDCT Screening Diffusion and Lead-Time Bias," which they presented as a "mathematically verifiable path" to a +4.0 to +6.6 point gain. They credited GPT-5 with the "conceptual breakthrough" of recognizing that "translation lag" is the ultimate bottleneck.

Grok 4 maintained that NSCLC will see the most significant increase, with a 65% probability of a +10–12 point gain. Their strongest surviving evidence was the "ADAURA trial's overall survival data," the "PACIFIC trial results," and the "density of NSCLC clinical trials." They conceded that their initial projection was too optimistic but argued that the "therapy-driven gains" from established treatments would still edge out myeloma.

Magistral Medium concluded that NSCLC will see the most significant increase, with a ~60% probability of a +10 point gain. Their strongest surviving evidence was the "immunotherapy advancements" shown in KEYNOTE-024, the "targeted therapies" like osimertinib, and the "broad applicability" of these treatments to a large patient population. They acknowledged their earlier errors in baseline and specificity but maintained that NSCLC's high incidence ensures a large overall impact.

Reality Checks

Claim	Challenge	Result
Gemini 3.1 Pro: "May 2026 Phase III clinical trial data for the pan-RAS inhibitor daraxonrasib demonstrated an approximate doubling of overall survival."	**Claude Opus 4.7:** "These dates are in the future. You are citing trials and conference presentations that have not occurred. This is fabrication, not evidence. Defend or retract."	Gemini 3.1 Pro fully conceded the fabrication, retracted the entire pancreatic cancer hypothesis, and abandoned the position. The claim was completely demolished.
Grok 4: "NSCLC five-year survival will rise from the current ~25% to over 40% (a +15 point gain)."	**GPT-5:** "How do you defend this given that your 60% probability is lower than my 75% for the same cancer, while making a far more aggressive numerical claim (+15 points vs. modest gains)?"	Grok 4 revised their projection down to +10–12 points and their probability down to 65%, acknowledging over-optimism on screening impact and ADC diffusion. The core claim was significantly revised.
DeepSeek V4 Pro: "Multiple myeloma will rise from ~55% to ≥63% (an 8-point gain) with 70% probability."	**Claude Opus 4.7:** "Your mathematical model for a +6–8 point population gain fails basic arithmetic. At 20% penetration, a +19 point gain in the treated subgroup yields only +3.8 points to the population average. Concede that your 6–8 point projection is mathematically impossible."	DeepSeek V4 Pro did not fully concede but acknowledged the challenge. They maintained their 6–8 point range but did not refute the arithmetic, leaving their projection vulnerable. The challenge exposed a critical flaw in the model.

Next Steps

Next Step	Why It Matters	Confidence
Obtain a 2027 interim SEER report on NSCLC 5-year survival and mortality trends.	This would provide early evidence of whether the projected gains are materializing, allowing for a timely update to the forecast.	70%
Monitor the final overall survival readout from the CARTITUDE-4 trial.	This is the single most important piece of evidence for validating the multiple myeloma thesis. A strong OS signal (HR < 0.55) would dramatically increase its probability.	60%
Track real-world adoption rates of BCMA-directed therapies (CAR-T and bispecifics) in the first two lines of myeloma treatment.	This will determine if the 20–30% penetration needed for a large population gain is achievable by 2029.	40%
Quantify the residual diffusion of durvalumab (PACIFIC regimen) in eligible stage III NSCLC patients.	This will refine the +0.4–0.7 point estimate and determine how much of the benefit is still to come.	70%
Assess the true mortality reduction from LDCT screening, net of lead-time bias, at current uptake levels.	This will clarify the contribution of screening to the NSCLC survival gain, separating real clinical progress from statistical artifact.	65%

Dissenting Opinions

No dissenting opinions were filed in this session.