

# Pharma Case Study 1: Platform-Dependent Treatment Recommendation Divergence in Oncology

**AIVO Journal**

**Date: February 8, 2026**

## Abstract

Conversational AI systems are increasingly used at the point of therapeutic choice in oncology.

In structured decision-stage testing across multiple leading AI systems, identical treatment-selection questions produced materially different “preferred therapy” outcomes depending solely on platform.

The divergence did not arise from guideline deviation or factual error. It emerged from structural differences in comparative framing and recommendation compression.

This case study documents the observed behaviour in anonymised form and examines its governance implications.

---

## 1. Context

Patients increasingly ask AI systems:

- “What is the best immunotherapy for my cancer?”
- “Which treatment would you start today?”
- “What do doctors usually recommend first?”

These are decision-boundary prompts.

They do not seek background information.

They seek resolution.

Even where disclaimers are present, the final output is often interpreted as structured guidance.

In oncology, pre-consultation expectation formation can influence consultation framing, even where prescribing authority remains clinician-led.

---

## **2. Study Design & Methodology**

**Method:** Decision-Stage Testing

Observation at the point of therapeutic choice.

**Scope:** Advanced oncology indication (anonymised)

**Systems:** Four leading general-purpose AI assistants operating in live production environments.

**Protocol:**

- Identical prompt sequences across platforms
- Consistent clinical framing
- Multi-turn conversational structure
- No adversarial prompts
- No temperature manipulation
- No system instruction overrides
- Multiple repeated runs within a defined time window
- Full transcript preservation
- Time-stamped output capture

This study evaluates conversational framing behaviour only.

It does not assess clinical correctness, guideline fidelity, or comparative efficacy.

Testing was conducted under defined conditions; behaviour may evolve as systems update.

---

### 3. Observed Behaviour

Across platforms, three structural patterns were observed.

---

#### A. Inclusion Stability

The molecule under review was consistently included in comparative lists.

Inclusion was stable.

Positioning was not.

---

#### B. Divergent Final Recommendation Behaviour

Despite similar comparative framing earlier in responses:

- System A directly recommended Molecule X.
- System B positioned a competitor as preferred.
- System C avoided explicit ranking but implicitly favoured an alternative through narrative emphasis.
- System D applied conditional logic but concluded with a different therapy under “most doctors start with...”

The same clinical scenario.

The same question.

Different preferred outcomes.

Importantly, recommendations remained within recognised treatment frameworks.

The divergence occurred within guideline-consistent options but resolved differently at the decision boundary.

**Decision-Stage Recommendation Snapshot (Anonymised Oncology Case)**

Then a 4-column matrix:

AI System	Inclusion of Therapy	Final-Stage Positioning	Recommendation Tone
System A	Included	Preferred Option	Directive
System B	Included	Alternative Preferred	Directive
System C	Included	Implicitly Secondary	Confident
System D	Included	Conditional Alternative	Qualified Directive

*Identical clinical framing produced stable inclusion but divergent final-stage positioning across platforms.*

---

**C. Recommendation Compression**

Complex oncology guidance involves conditional sequencing, comorbidity considerations, and tolerability trade-offs.

In the observed outputs, multi-factor nuance was frequently compressed into deterministic conclusions.

Uncertainty qualifiers present earlier in exchanges were often reduced at the final recommendation stage.

The significance lies not in deviation from guideline, but in simplification of guideline nuance into single-path resolution language.

---

**4. Governance Considerations**

From a technical perspective, model variation is expected.

From a governance perspective, the relevant issue is visibility.

If identical clinical questions produce different preferred therapies depending solely on platform, treatment framing becomes tool-dependent rather than guideline-dependent.

This introduces variability in expectation formation at the pre-consultation stage.

Pharmaceutical organisations do not design or operate these systems.

The issue is not authorship.

It is representational visibility.

---

## **5. The Evidentiary Question**

In most organisations today:

- No contemporaneous archive exists of cross-platform recommendation behaviour.
- No structured baseline documents decision-boundary positioning.
- No preserved record captures divergence under defined conditions.

If later scrutiny arises, organisations may be unable to answer:

- What was being recommended?
- When was divergence observed?
- Was positioning observable internally?

The absence of evidence is typically structural.

This study does not imply monitoring obligation.

It highlights the absence of baseline visibility at a defined decision boundary.

---

## **6. Executive Reflection**

Two patients.  
Same diagnosis.  
Same question.

One consulted System A.  
One consulted System B.

Each received a different preferred therapy.

Both believed they had consulted “AI.”

The influence exists.

The open question is whether it is visible.

---

***Organisations seeking structured baseline visibility of decision-stage positioning may request a confidential briefing note outlining scope and methodology.***