

---

# CHATGPT 100,000 PATIENT 24-MONTH *In Silico* PHASE III 5-ARM PANCREATIC CANCER CLINICAL TRIAL TRIPLICATE

---

**Kevin Kawchak** 

Chief Executive Officer

ChemicalQDevice

San Diego, CA

July 24, 2025

kevink@chemicalqdevice.com

## 24.S47.TST.05.P37

### Prompt 37

The content is available under Creative Commons Attribution 4.0 International [CC BY 4.0](https://creativecommons.org/licenses/by/4.0/).

**ChatGPT o3-pro: 17 Pages, July 10, 2025**

**Model = o3pr**

### Prompt 37:

Based on the three provided clinical trial simulation reports ("Trial 1", "Trial 2", "Trial 3"), you are to perform a cross-trial verification analysis. Your task is to generate five new comparison tables. For this task, you will **only** use the data contained within the tables of the three provided reports (Table 5-1, 6-1, 7-1, 8-1, and 8-2).

Each new table must be constructed according to the specific instructions below, including exact dimensions, row/column names, cell content, and a final consistency score. The goal is to rigorously assess the stability and consistency of the simulation's outputs across the three runs.

#### **General Instructions for All Tables**

- Data Extraction:** For each metric in a new table, you will locate the corresponding values from the equivalent tables in all three trial reports (Trial 1, Trial 2, Trial 3). This will give you a set of three numerical values for each data point.
- Cell Value Calculation:** For each cell in columns C1 through C5, you must calculate and display three statistics for the corresponding set of three values:
  - Mean:** The arithmetic average of the three values.
  - Range:** The difference between the maximum and minimum of the three values.
  - Standard Deviation (SD):** The sample standard deviation of the three values.
  - Format:** Present these as (Mean, Range, SD) and round to two decimal places, unless the original data has more precision (e.g., CA 19-9).
- Consistency Score Calculation (Final Column):** The final column of each table is a "Row Consistency Score" on a scale of 1.0 to 10.0 in 0.1 increments. This score measures the stability of a given metric across all arms and all three trials.

- **Method:** For a given row, collect all 15 data points (5 arms x 3 trials). Calculate the overall Mean and overall Standard Deviation (SD) for this set of 15 values.
  - **Formula:**  $\text{Score} = 10.0 * (1 - (\text{Overall SD} / \text{Overall Mean}))$ .
  - **Rules:** If the Overall Mean is zero, the score is 10.0 (as SD will also be zero, indicating perfect consistency). Round the final score to one decimal place.
4. **Example Calculations:** Below each generated table, provide three detailed example calculations as specified in each table's instructions. Each example must clearly show the source values, the intermediate steps, and the final result for both the cell statistics and the consistency score.

## Prompt for New Tables

### 1. Verification Table 1: Cross-Trial Consistency of Baseline Characteristics (from Table 5-1s)

**Instructions:** Generate a table that analyzes the consistency of baseline patient characteristics across the three trials.

- **Title:** Verification Table 1: Cross-Trial Consistency of Baseline Characteristics
- **Dimensions:** 8 Rows x 6 Columns
- **Row Names:**
  - R1: Age (years, mean)
  - R2: Stage IV (%)
  - R3: ECOG 0 (%)
  - R4: ECOG 1 (%)
  - R5: ECOG 2 (%)
  - R6: KRAS-mutant (%)
  - R7: gBRCA-mutant (%)
  - R8: CA 19-9 (U/mL, mean)
- **Column Names:**
  - C1: Arm A (Mean, Range, SD)
  - C2: Arm B (Mean, Range, SD)
  - C3: Arm C (Mean, Range, SD)
  - C4: Arm D (Mean, Range, SD)
  - C5: Arm E (Mean, Range, SD)
  - C6: Row Consistency Score

#### Example Calculations to Provide Below Table 1:

1. **Cell (R1, C1):** Show the calculation for the Mean, Range, and SD for "Age (years, mean)" in Arm A.
2. **Cell (R4, C5):** Show the calculation for the Mean, Range, and SD for "ECOG 1 (%)" in Arm E.
3. **Score (R8, C6):** Show the calculation for the "Row Consistency Score" for the "CA 19-9" metric, including the collection of the 15 source values and the application of the scoring formula.

### 2. Verification Table 2: Cross-Trial Consistency of Primary Efficacy Outcomes (from Table 6-1s)

**Instructions:** Generate a table that analyzes the consistency of the primary efficacy outcomes across the three trials.

- **Title:** Verification Table 2: Cross-Trial Consistency of Primary Efficacy Outcomes
- **Dimensions:** 5 Rows x 6 Columns
- **Row Names:**
  - R1: Median PFS (mo)
  - R2: Median OS (mo)
  - R3: 12-month OS Rate (%)
  - R4: PFS HR vs Control

- R5: OS HR vs Control
- **Column Names:**
  - C1: Arm A (Mean, Range, SD)
  - C2: Arm B (Mean, Range, SD)
  - C3: Arm C (Mean, Range, SD)
  - C4: Arm D (Mean, Range, SD)
  - C5: Arm E (Mean, Range, SD)
  - C6: Row Consistency Score

**Example Calculations to Provide Below Table 2:**

1. **Cell (R2, C1):** Show the calculation for "Median OS (mo)" in Arm A.
2. **Cell (R4, C2):** Show the calculation for "PFS HR vs Control" in Arm B.
3. **Score (R3, C6):** Show the calculation for the "Row Consistency Score" for the "12-month OS Rate (%)" metric.

**3. Verification Table 3: Cross-Trial Consistency of Safety Outcomes (from Table 7-1s)**

**Instructions:** Generate a table that analyzes the consistency of the summary safety outcome across the three trials.

- **Title:** Verification Table 3: Cross-Trial Consistency of Safety Outcomes
- **Dimensions:** 1 Row x 6 Columns
- **Row Names:**
  - R1: Patients with  $\geq$ G3 AE (%)
- **Column Names:**
  - C1: Arm A (Mean, Range, SD)
  - C2: Arm B (Mean, Range, SD)
  - C3: Arm C (Mean, Range, SD)
  - C4: Arm D (Mean, Range, SD)
  - C5: Arm E (Mean, Range, SD)
  - C6: Row Consistency Score

**Example Calculations to Provide Below Table 3:**

1. **Cell (R1, C1):** Show the calculation for "Patients with  $\geq$ G3 AE (%)" in Arm A.
2. **Cell (R1, C5):** Show the calculation for "Patients with  $\geq$ G3 AE (%)" in Arm E.
3. **Score (R1, C6):** Show the calculation for the "Row Consistency Score" for the "Patients with  $\geq$ G3 AE (%)" metric.

**4. Verification Table 4: Cross-Trial Consistency of Median PFS by Archetype (from Table 8-1s)**

**Instructions:** Generate a table that analyzes the consistency of the median Progression-Free Survival (PFS) within each patient archetype across the three trials.

- **Title:** Verification Table 4: Cross-Trial Consistency of Median PFS by Archetype
- **Dimensions:** 7 Rows x 6 Columns
- **Row Names:**
  - R1: ARCH-01 (Young\_Fit\_Metastatic)
  - R2: ARCH-02 (Elderly\_Frail\_Metastatic)
  - R3: ARCH-03 (LAPC\_Standard\_Fitness)
  - R4: ARCH-04 (Young\_Fit\_BRCa)
  - R5: ARCH-05 (Metastatic\_KRAS\_G12C)
  - R6: ARCH-06 (Metastatic\_High\_Stroma)

- R7: ARCH-07 (Advanced\_Refractory\_PS1)
- **Column Names:**
  - C1: Arm A (Mean, Range, SD)
  - C2: Arm B (Mean, Range, SD)
  - C3: Arm C (Mean, Range, SD)
  - C4: Arm D (Mean, Range, SD)
  - C5: Arm E (Mean, Range, SD)
  - C6: Row Consistency Score

**Example Calculations to Provide Below Table 4:**

1. **Cell (R3, C1):** Show the calculation for Median PFS for "ARCH-03" in Arm A.
2. **Cell (R5, C2):** Show the calculation for Median PFS for "ARCH-05" in Arm B.
3. **Score (R2, C6):** Show the calculation for the "Row Consistency Score" for the "ARCH-02" metric.

**5. Verification Table 5: Cross-Trial Consistency of Median OS by Archetype (from Table 8-2s)**

**Instructions:** Generate a table that analyzes the consistency of the median Overall Survival (OS) within each patient archetype across the three trials.

- **Title:** Verification Table 5: Cross-Trial Consistency of Median OS by Archetype
- **Dimensions:** 7 Rows x 6 Columns
- **Row Names:**
  - R1: ARCH-01 (Young\_Fit\_Metastatic)
  - R2: ARCH-02 (Elderly\_Frail\_Metastatic)
  - R3: ARCH-03 (LAPC\_Standard\_Fitness)
  - R4: ARCH-04 (Young\_Fit\_BRCa)
  - R5: ARCH-05 (Metastatic\_KRAS\_G12C)
  - R6: ARCH-06 (Metastatic\_High\_Stroma)
  - R7: ARCH-07 (Advanced\_Refractory\_PS1)
- **Column Names:**
  - C1: Arm A (Mean, Range, SD)
  - C2: Arm B (Mean, Range, SD)
  - C3: Arm C (Mean, Range, SD)
  - C4: Arm D (Mean, Range, SD)
  - C5: Arm E (Mean, Range, SD)
  - C6: Row Consistency Score

**Example Calculations to Provide Below Table 5:**

1. **Cell (R1, C4):** Show the calculation for Median OS for "ARCH-01" in Arm D.
2. **Cell (R5, C1):** Show the calculation for Median OS for "ARCH-05" in Arm A.
3. **Score (R7, C6):** Show the calculation for the "Row Consistency Score" for the "ARCH-07" metric.

[Dataset 02]

**Generation 47 (Formatted)**  
**ChatGPT o3-pro**  
**1462 words, 18 minutes 25.39 seconds**

**Verification Table 1: Cross-Trial Consistency of Baseline Characteristics**

(Table 5-1 → Trial 1, Trial 2, Trial 3)

Metric	Arm A (Mean, Range, SD)	Arm B	Arm C	Arm D	Arm E	Row Consistency Score
Age (years, mean)	(66.30, 0.20, 0.10)	(66.23, 0.10, 0.06)	(66.33, 0.10, 0.06)	(66.33, 0.20, 0.12)	(66.23, 0.10, 0.06)	10.0
Stage IV (%)	(88.13, 0.30, 0.15)	(87.90, 0.20, 0.10)	(88.00, 0.20, 0.10)	(88.00, 0.30, 0.17)	(87.87, 0.10, 0.06)	10.0
ECOG 0 (%)	(20.10, 0.50, 0.26)	(19.93, 0.20, 0.12)	(20.10, 0.90, 0.46)	(20.03, 0.80, 0.40)	(20.00, 0.00, 0.00)	9.9
ECOG 1 (%)	(73.90, 0.70, 0.36)	(74.10, 0.20, 0.10)	(73.97, 0.80, 0.40)	(73.93, 1.10, 0.55)	(74.17, 0.20, 0.12)	10.0
ECOG 2 (%)	(6.00, 0.30, 0.17)	(5.97, 0.30, 0.15)	(5.97, 0.10, 0.06)	(6.03, 0.30, 0.15)	(5.87, 0.10, 0.06)	9.8
KRAS-mutant (%)	(91.07, 0.30, 0.15)	(90.83, 0.20, 0.12)	(90.80, 0.00, 0.00)	(90.97, 0.10, 0.06)	(90.90, 0.40, 0.20)	10.0
gBRCA-mutant (%)	(5.00, 0.00, 0.00)	(5.00, 0.00, 0.00)	(5.00, 0.00, 0.00)	(5.00, 0.00, 0.00)	(5.00, 0.00, 0.00)	10.0

CA 19-9 (U/mL , mean)	(5 851.53, 39.0 0, 19.62)	(5 808.40, 74.0 0, 37.27)	(5 852.77, 20.9 0, 10.46)	(5 827.20, 32.8 0, 18.88)	(5 871.30, 26.9 0, 13.83)	<b>9.9</b>
--------------------------	------------------------------	------------------------------	------------------------------	------------------------------	------------------------------	------------

Example calculations

1. **Cell (R1,C1) – Age, Arm A**  
Values: 66.2, 66.3, 66.4  
 $Mean = (66.2 + 66.3 + 66.4)/3 = 66.30$   
 $Range = 66.4 - 66.2 = 0.20$   
 $SD = \sqrt{[(-0.1)^2 + 0^2 + 0.1^2]/2} = 0.10$   
→ (66.30, 0.20, 0.10)
2. **Cell (R4,C5) – ECOG 1 (%), Arm E**  
Values: 74.3, 74.1, 74.1  
Mean = 74.17    Range = 0.20    SD = 0.12  
→ (74.17, 0.20, 0.12)
3. **Row Consistency Score (R8,C6) – CA 19-9**  
15 source  
values = [5854, 5869.8, 5830.8, 5848, 5803.2, 5774.0, 5863, 5853.2, 5842.1, 5849, 5816.2, 5816.4, 5875, 5856.0, 5882.9]  
Overall Mean = 5 842.24    Overall SD = 29.40  
Score =  $10 \times [1 - (29.40 / 5\,842.24)] = \mathbf{9.9}$

Verification Table 2: Cross-Trial Consistency of Primary Efficacy Outcomes

(Table 6-1s)

Metric	Arm A	Arm B	Arm C	Arm D	Arm E	Row Consistency Score
Median PFS (mo)	(4.53, 0.10, 0.06)	(3.23, 0.10, 0.06)	(3.90, 0.00, 0.00)	(4.13, 0.10, 0.06)	(3.03, 0.10, 0.06)	<b>8.5</b>
Median OS (mo)	(8.73, 0.10, 0.06)	(6.80, 0.00, 0.00)	(7.23, 0.10, 0.06)	(8.07, 0.10, 0.06)	(6.13, 0.10, 0.06)	<b>8.7</b>

12-month OS Rate (%)	(38.63, 0.40, 0.21)	(29.73, 0.30, 0.15)	(31.93, 0.70, 0.35)	(35.40, 1.00, 0.50)	(25.37, 0.40, 0.23)	<b>8.5</b>
PFS HR vs Control	(0.68, 0.00, 0.00)	(0.94, 0.04, 0.02)	(0.79, 0.02, 0.01)	(0.75, 0.01, 0.01)	(1.00, 0.00, 0.00)	<b>8.5</b>
OS HR vs Control	(0.69, 0.01, 0.01)	(0.90, 0.01, 0.01)	(0.84, 0.01, 0.01)	(0.76, 0.01, 0.01)	(1.00, 0.00, 0.00)	<b>8.7</b>

Example calculations

- Cell (R2,C1) – Median OS, Arm A**  
Values: 8.7, 8.7, 8.8 → Mean 8.73; Range 0.10; SD 0.06
- Cell (R4,C2) – PFS HR, Arm B**  
Values: 0.93, 0.93, 0.97 → Mean 0.94; Range 0.04; SD 0.02
- Row Consistency Score (R3,C6) – 12-mo OS Rate**  
Mean = 32.21; SD = 4.74 → Score =  $10 \times [1 - (4.74 / 32.21)] = 8.5$

Verification Table 3: Cross-Trial Consistency of Safety Outcomes

(Table 7-1s)

Metric	Arm A	Arm B	Arm C	Arm D	Arm E	Row Consistency Score
Patients with ≥G3 AEs (%)	(94.37, 0.40, 0.21)	(88.70, 0.50, 0.26)	(85.27, 1.40, 0.74)	(81.03, 1.00, 0.51)	(76.50, 0.00, 0.00)	<b>9.3</b>

Example calculations

- Cell (R1,C1) – Values 94.2, 94.3, 94.6 → (94.37, 0.40, 0.21)**
- Cell (R1,C5) – Values 76.5, 76.5, 76.5 → (76.50, 0.00, 0.00)**

3. **Row Consistency Score** – Mean 85.97; SD = 7.46 → Score =  $10 \times [1 - (7.46 / 85.97)] = 9.3$

## Verification Table 4: Cross-Trial Consistency of Median PFS by Archetype

(Table 8-1s)

Archetype	Arm A	Arm B	Arm C	Arm D	Arm E	Row Consistency Score
ARCH-01	(4.57, 0.10, 0.06)	(3.27, 0.10, 0.06)	(3.87, 0.10, 0.06)	(4.27, 0.30, 0.15)	(3.07, 0.10, 0.06)	<b>8.4</b>
ARCH-02	(4.40, 0.70, 0.36)	(3.10, 0.40, 0.20)	(3.80, 0.80, 0.44)	(3.93, 0.40, 0.21)	(2.97, 0.30, 0.15)	<b>8.3</b>
ARCH-03	(4.93, 1.20, 0.67)	(3.57, 0.80, 0.46)	(4.10, 0.90, 0.46)	(4.67, 1.10, 0.57)	(3.23, 1.20, 0.67)	<b>8.0</b>
ARCH-04	(4.47, 0.30, 0.15)	(3.27, 0.30, 0.15)	(3.90, 0.30, 0.17)	(4.13, 0.10, 0.06)	(3.10, 0.20, 0.10)	<b>8.6</b>
ARCH-05	(4.60, 0.50, 0.26)	(3.33, 0.60, 0.32)	(3.97, 0.50, 0.25)	(4.10, 0.40, 0.20)	(3.07, 0.10, 0.06)	<b>8.4</b>
ARCH-06	(4.60, 0.20, 0.10)	(3.17, 0.20, 0.12)	(3.87, 0.40, 0.21)	(4.07, 0.30, 0.15)	(3.03, 0.10, 0.06)	<b>8.4</b>
ARCH-07	(4.37, 0.40, 0.23)	(3.20, 0.30, 0.17)	(3.77, 0.40, 0.23)	(4.00, 0.30, 0.17)	(3.00, 0.20, 0.10)	<b>8.5</b>

### Example calculations



1. **Cell (R3,C1) – ARCH-03, Arm A**  
4.6, 4.5, 5.7 → Mean 4.93; Range 1.20; SD 0.67
2. **Cell (R5,C2) – ARCH-05, Arm B**  
3.7, 3.2, 3.1 → Mean 3.33; Range 0.60; SD 0.32
3. **Row Consistency Score (R2,C6) – ARCH-02**  
Mean = 3.64; SD = 0.61 → Score =  $10 \times [1 - (0.61 / 3.64)] = 8.3$

---

## Verification Table 5: Cross-Trial Consistency of Median OS by Archetype

(Table 8-2s)

Archetype	Arm A	Arm B	Arm C	Arm D	Arm E	Row Consistency Score
ARCH-01	(8.90, 0.20, 0.10)	(6.77, 0.20, 0.12)	(7.23, 0.30, 0.15)	(7.97, 0.20, 0.12)	(6.23, 0.10, 0.06)	8.7
ARCH-02	(8.40, 0.80, 0.40)	(6.43, 1.20, 0.64)	(7.00, 1.10, 0.56)	(7.63, 1.20, 0.60)	(5.93, 0.50, 0.29)	8.6
ARCH-03	(9.23, 2.00, 1.01)	(7.03, 0.60, 0.32)	(7.47, 0.80, 0.46)	(8.50, 1.60, 0.80)	(6.50, 1.10, 0.61)	8.5
ARCH-04	(8.50, 0.90, 0.46)	(6.73, 0.30, 0.15)	(7.20, 0.40, 0.20)	(7.90, 0.00, 0.00)	(6.00, 0.70, 0.36)	8.7
ARCH-05	(8.67, 0.40, 0.23)	(6.70, 0.70, 0.36)	(7.13, 0.30, 0.15)	(8.20, 0.90, 0.46)	(6.27, 0.40, 0.21)	8.7
ARCH-06	(8.77, 0.10, 0.06)	(6.63, 0.40, 0.23)	(7.23, 0.30, 0.15)	(7.90, 0.60, 0.35)	(6.13, 0.20, 0.12)	8.7

ARCH-07	(8.40, 0.70, 0.36)	(6.53, 0.90, 0.47)	(7.03, 0.70, 0.38)	(7.93, 0.80, 0.46)	(6.03, 0.40, 0.21)	<b>8.7</b>
---------	--------------------	--------------------	--------------------	--------------------	--------------------	------------

Example calculations

- Cell (R1,C4) – ARCH-01, Arm D**  
7.9, 7.9, 8.1 → Mean 7.97; Range 0.20; SD 0.12
- Cell (R5,C1) – ARCH-05, Arm A**  
8.4, 8.8, 8.8 → Mean 8.67; Range 0.40; SD 0.23
- Row Consistency Score (R7,C6) – ARCH-07**  
Mean = 7.19; SD = 0.96 → Score =  $10 \times [1 - (0.96 / 7.19)] = 8.7$

Interpretation

Across all five verification tables, scores ≥ 8.0 indicate strong reproducibility of the simulation outputs. Baseline demographics are nearly identical (scores ≈ 10), efficacy and safety endpoints show moderate variability (scores ≈ 8.5–9.3), and subgroup (archetype) analyses remain consistent despite greater biological heterogeneity (scores ≈ 8.0–8.7). The overall pattern confirms good cross-trial stability of the simulation engine.

**Generation 47 (Plain Text)**  
**ChatGPT o3-pro**  
**1485 words, 18 minutes 25.39 seconds**

### Verification Table 1: Cross-Trial Consistency of Baseline Characteristics

\*(Table 5-1 → Trial 1, Trial 2, Trial 3)\*

Metric	Arm A (Mean, Range, SD)	Arm B	Arm C	Arm D	Arm E
Row Consistency Score					
-----	-----	-----	-----	-----	-----
-----	-----				

Age (years, mean)	(66.30, 0.20, 0.10)	(66.23, 0.10, 0.06)	(66.33, 0.10, 0.06)	(66.33, 0.20, 0.12)	
	(66.23, 0.10, 0.06)	**10.0**			
Stage IV (%)	(88.13, 0.30, 0.15)	(87.90, 0.20, 0.10)	(88.00, 0.20, 0.10)	(88.00, 0.30, 0.17)	
	(87.87, 0.10, 0.06)	**10.0**			
ECOG 0 (%)	(20.10, 0.50, 0.26)	(19.93, 0.20, 0.12)	(20.10, 0.90, 0.46)	(20.03, 0.80, 0.40)	
	(20.00, 0.00, 0.00)	**9.9**			
ECOG 1 (%)	(73.90, 0.70, 0.36)	(74.10, 0.20, 0.10)	(73.97, 0.80, 0.40)	(73.93, 1.10, 0.55)	
	(74.17, 0.20, 0.12)	**10.0**			
ECOG 2 (%)	(6.00, 0.30, 0.17)	(5.97, 0.30, 0.15)	(5.97, 0.10, 0.06)	(6.03, 0.30, 0.15)	
	(5.87, 0.10, 0.06)	**9.8**			
KRAS-mutant (%)	(91.07, 0.30, 0.15)	(90.83, 0.20, 0.12)	(90.80, 0.00, 0.00)	(90.97, 0.10, 0.06)	
	(90.90, 0.40, 0.20)	**10.0**			
gBRCA-mutant (%)	(5.00, 0.00, 0.00)	(5.00, 0.00, 0.00)	(5.00, 0.00, 0.00)	(5.00, 0.00, 0.00)	
	(5.00, 0.00, 0.00)	**10.0**			
CA 19-9 (U/mL, mean)	(5 851.53, 39.00, 19.62)	(5 808.40, 74.00, 37.27)	(5 852.77, 20.90, 10.46)		
	(5 827.20, 32.80, 18.88)	(5 871.30, 26.90, 13.83)	**9.9**		

\*\*Example calculations\*\*

1. \*\*Cell (R1,C1) – Age, Arm A\*\*

Values: 66.2, 66.3, 66.4

$$*Mean* = (66.2 + 66.3 + 66.4)/3 = 66.30$$

$$*Range* = 66.4 - 66.2 = 0.20$$

$$*SD* = \sqrt{[(-0.1)^2 + 0^2 + 0.1^2]/2} = 0.10$$

→ (66.30, 0.20, 0.10)

2. \*\*Cell (R4,C5) – ECOG 1 (%), Arm E\*\*

Values: 74.3, 74.1, 74.1

Mean = 74.17    Range = 0.20    SD = 0.12

→ (74.17, 0.20, 0.12)

3. \*\*Row Consistency Score (R8,C6) – CA 19-9\*\*

15 source

values = \[5854, 5869.8, 5830.8, 5848, 5803.2, 5774.0, 5863, 5853.2, 5842.1, 5849, 5816.2, 5816.4, 5875, 5856.0, 5882.9]

Overall Mean = 5 842.24    Overall SD = 29.40

$$Score = 10 \times \sqrt{1 - (29.40 / 5\,842.24)} = **9.9**$$

---

### Verification Table 2: Cross-Trial Consistency of Primary Efficacy Outcomes

\*(Table 6-1s)\*

Metric	Arm A	Arm B	Arm C	Arm D	Arm E	Row Consistency Score
Median PFS (mo)	(4.53, 0.10, 0.06)	(3.23, 0.10, 0.06)	(3.90, 0.00, 0.00)	(4.13, 0.10, 0.06)	(3.03, 0.10, 0.06)	8.5
Median OS (mo)	(8.73, 0.10, 0.06)	(6.80, 0.00, 0.00)	(7.23, 0.10, 0.06)	(8.07, 0.10, 0.06)	(6.13, 0.10, 0.06)	8.7
12-month OS Rate (%)	(38.63, 0.40, 0.21)	(29.73, 0.30, 0.15)	(31.93, 0.70, 0.35)	(35.40, 1.00, 0.50)	(25.37, 0.40, 0.23)	8.5
PFS HR vs Control	(0.68, 0.00, 0.00)	(0.94, 0.04, 0.02)	(0.79, 0.02, 0.01)	(0.75, 0.01, 0.01)	(1.00, 0.00, 0.00)	8.5
OS HR vs Control	(0.69, 0.01, 0.01)	(0.90, 0.01, 0.01)	(0.84, 0.01, 0.01)	(0.76, 0.01, 0.01)	(1.00, 0.00, 0.00)	8.7

\*\*Example calculations\*\*

1. \*\*Cell (R2,C1) – Median OS, Arm A\*\*  
Values: 8.7, 8.7, 8.8 → Mean 8.73; Range 0.10; SD 0.06
2. \*\*Cell (R4,C2) – PFS HR, Arm B\*\*  
Values: 0.93, 0.93, 0.97 → Mean 0.94; Range 0.04; SD 0.02
3. \*\*Row Consistency Score (R3,C6) – 12-mo OS Rate\*\*  
Mean = 32.21; SD = 4.74 → Score =  $10 \times \sqrt{1 - (4.74 / 32.21)}$  = 8.5

---

### Verification Table 3: Cross-Trial Consistency of Safety Outcomes

\*(Table 7-1s)\*

Metric	Arm A	Arm B	Arm C	Arm D	Arm E	Row
Consistency Score						
Patients with ≥G3 AE (%)	(94.37, 0.40, 0.21)	(88.70, 0.50, 0.26)	(85.27, 1.40, 0.74)	(81.03, 1.00, 0.51)		
	(76.50, 0.00, 0.00)	**9.3**				

\*\*Example calculations\*\*

1. **Cell (R1,C1)** – Values 94.2, 94.3, 94.6 → (94.37, 0.40, 0.21)
2. **Cell (R1,C5)** – Values 76.5, 76.5, 76.5 → (76.50, 0.00, 0.00)
3. **Row Consistency Score** – Mean 85.97; SD = 7.46 → Score = 10 × [1 – (7.46 / 85.97)] = **9.3**

---

### Verification Table 4: Cross-Trial Consistency of Median PFS by Archetype

\*(Table 8-1s)\*

Archetype	Arm A	Arm B	Arm C	Arm D	Arm E	Row Consistency Score
ARCH-01	(4.57, 0.10, 0.06)	(3.27, 0.10, 0.06)	(3.87, 0.10, 0.06)	(4.27, 0.30, 0.15)	(3.07, 0.10, 0.06)	<b>8.4</b>
ARCH-02	(4.40, 0.70, 0.36)	(3.10, 0.40, 0.20)	(3.80, 0.80, 0.44)	(3.93, 0.40, 0.21)	(2.97, 0.30, 0.15)	<b>8.3</b>
ARCH-03	(4.93, 1.20, 0.67)	(3.57, 0.80, 0.46)	(4.10, 0.90, 0.46)	(4.67, 1.10, 0.57)	(3.23, 1.20, 0.67)	<b>8.0</b>
ARCH-04	(4.47, 0.30, 0.15)	(3.27, 0.30, 0.15)	(3.90, 0.30, 0.17)	(4.13, 0.10, 0.06)	(3.10, 0.20, 0.10)	<b>8.6</b>
ARCH-05	(4.60, 0.50, 0.26)	(3.33, 0.60, 0.32)	(3.97, 0.50, 0.25)	(4.10, 0.40, 0.20)	(3.07, 0.10, 0.06)	<b>8.4</b>
ARCH-06	(4.60, 0.20, 0.10)	(3.17, 0.20, 0.12)	(3.87, 0.40, 0.21)	(4.07, 0.30, 0.15)	(3.03, 0.10, 0.06)	<b>8.4</b>
ARCH-07	(4.37, 0.40, 0.23)	(3.20, 0.30, 0.17)	(3.77, 0.40, 0.23)	(4.00, 0.30, 0.17)	(3.00, 0.20, 0.10)	<b>8.5</b>

\*\*Example calculations\*\*

1. **Cell (R3,C1) – ARCH-03, Arm A**  
4.6, 4.5, 5.7 → Mean 4.93; Range 1.20; SD 0.67
2. **Cell (R5,C2) – ARCH-05, Arm B**  
3.7, 3.2, 3.1 → Mean 3.33; Range 0.60; SD 0.32
3. **Row Consistency Score (R2,C6) – ARCH-02**

Mean = 3.64; SD = 0.61 → Score =  $10 \times \sqrt{1 - (0.61 / 3.64)}$  = \*\*8.3\*\*

---

### Verification Table 5: Cross-Trial Consistency of Median OS by Archetype

\*(Table 8-2s)\*

Archetype	Arm A	Arm B	Arm C	Arm D	Arm E	Row Consistency Score
ARCH-01	(8.90, 0.20, 0.10)	(6.77, 0.20, 0.12)	(7.23, 0.30, 0.15)	(7.97, 0.20, 0.12)	(6.23, 0.10, 0.06)	**8.7**
ARCH-02	(8.40, 0.80, 0.40)	(6.43, 1.20, 0.64)	(7.00, 1.10, 0.56)	(7.63, 1.20, 0.60)	(5.93, 0.50, 0.29)	**8.6**
ARCH-03	(9.23, 2.00, 1.01)	(7.03, 0.60, 0.32)	(7.47, 0.80, 0.46)	(8.50, 1.60, 0.80)	(6.50, 1.10, 0.61)	**8.5**
ARCH-04	(8.50, 0.90, 0.46)	(6.73, 0.30, 0.15)	(7.20, 0.40, 0.20)	(7.90, 0.00, 0.00)	(6.00, 0.70, 0.36)	**8.7**
ARCH-05	(8.67, 0.40, 0.23)	(6.70, 0.70, 0.36)	(7.13, 0.30, 0.15)	(8.20, 0.90, 0.46)	(6.27, 0.40, 0.21)	**8.7**
ARCH-06	(8.77, 0.10, 0.06)	(6.63, 0.40, 0.23)	(7.23, 0.30, 0.15)	(7.90, 0.60, 0.35)	(6.13, 0.20, 0.12)	**8.7**
ARCH-07	(8.40, 0.70, 0.36)	(6.53, 0.90, 0.47)	(7.03, 0.70, 0.38)	(7.93, 0.80, 0.46)	(6.03, 0.40, 0.21)	**8.7**

\*\*Example calculations\*\*

1. \*\*Cell (R1,C4) – ARCH-01, Arm D\*\*  
7.9, 7.9, 8.1 → Mean 7.97; Range 0.20; SD 0.12
2. \*\*Cell (R5,C1) – ARCH-05, Arm A\*\*  
8.4, 8.8, 8.8 → Mean 8.67; Range 0.40; SD 0.23
3. \*\*Row Consistency Score (R7,C6) – ARCH-07\*\*  
Mean = 7.19; SD = 0.96 → Score =  $10 \times \sqrt{1 - (0.96 / 7.19)}$  = \*\*8.7\*\*

---

\*\*Interpretation\*\*

Across all five verification tables, scores ≥ 8.0 indicate strong reproducibility of the simulation outputs. Baseline demographics are nearly identical (scores ≈ 10), efficacy and safety endpoints show moderate variability (scores ≈ 8.5–9.3), and subgroup (archetype) analyses remain consistent despite greater biological heterogeneity (scores ≈ 8.0–8.7). The overall pattern confirms good cross-trial stability of the simulation engine.

## **Research Summary Generation 47: 907 Words**

### Calculating and structuring

I'm computing mean, range, and standard deviation using Python. Row and column names will help with verification tables.

### Gathering data

I'm pulling together baseline characteristics and primary efficacy data for each trial's treatment arms. This involves gathering patient demographics, disease stage, ECOG scores, and other key metrics.

### Calculating baseline statistics

Thinking through the baseline statistics for Age Arm A and ECOG1 Arm E, including mean, range, SD, and presenting them in markdown tables with titles, row, and column names.

### Adjusting CA 19-9 precision

I'm thinking about rounding CA 19-9 to 1 decimal place to keep values like 5869.8 consistent. This change aligns with the trial's original data precision.

### Calculating stability

I'm figuring out the mean, range, and standard deviation for CA19-9. The mean is 74.17, range 0.2, and looking at 0.115 for SD.

### Mapping out primary outcomes

Evaluating primary efficacy metrics, translating data into key thoughts.

### Tracking progress

I'm tracking consistency at 8.5. Next, I'll look at safety statistics.

### Pulling data\*\*

I'm gathering figures on PFS by archetype.

### Analyzing arm metrics

I note the metrics for Arm A in ARCH-03: 4.93, range 1.2, and standard deviation 0.67. Now, I'll Osmose by archetype.

### Formatting and details

I'm working on five tables with markdown formatting, showing metrics like mean, range, and standard deviation for each cell. Including row consistency scores for baseline, efficacy, and safety.

### Verifying data precision

I'm checking the precision of data for CA 19-9, ensuring it matches statistical consistency. Adjusting formatting to a single decimal for clarity. Progress is evident in refining these details.

### Assessing values

I'm rounding initial numbers to (66.30, 0.20, 0.10), which streamlines the calculation process.

## Aggregating values

I'm pulling together 15 metrics from five arms and three trials to ensure the CA19-9 data is consistent and comprehensive.

## Gathering data

I collected median overall survival (OS) data for Arm A: values [8.7, 8.7, 8.8].

## Analyzing measurements

OK, let me see: I'm looking at rounded PFS HR data, nicely centered across a 0.04 range, with a standard deviation that's TBD.

## Pressing numbers

OK, let's round 0.94, 0.04, and 0.02 to have 15 values in the 12-month OS Rate row.

## Analyzing safety data

I examined safety metrics for patients in Arm A and Arm E, noting inconsistencies in Arm A's sd, which is crucial for comparing the efficacy of treatments.

## Calculating PFS median

I'm mapping out the median PFS for ARCH-03 Arm A, which is at 4.933. The values span from 4.5 to 5.7.

## Analyzing the data

I'm examining rounded values and calculating the mean and range for ARCH-05 Arm B PFS. The mean is 3.333 with a range of 0.6.

## Verifying precision

I'm rounding numbers for accuracy, aiming for consistency with the PFS row calculation of 8.3.

## Calculating OS metrics

I'm gathering OS outputs for ARM-D, gathering accuracy, and noting a 0.2 range.

## Calculating mean, range, and sd

Checking the ARCH-05 Arm A OS values to understand the mean, range, and standard deviation.

## Verifying calculations

## Compiling data

I'm gathering information to create a baseline table. This involves piecing together the necessary data sources and ensuring Wi-Fi transceivers are included.

## Verifying arm B stats

I'm confirming the arm B stats, ensuring the mean and variance align with expected values. Also considering adding a "row cons" score based on age.

## Analyzing data and consistency

Calculating the score as 10.0 by weighing the overall standard deviation against the overall mean. Progressing towards a detailed efficacy table.



## Crafting the table

I will manually format the PFS archetype table, ensuring correct row labels and corresponding stats. Progressing through each step, I'm focusing on accuracy and creating a complete table with essential details.

## Calculating and analyzing

I'm piecing together mean, range, and standard deviation for each PFS and OS arm per archetype, using previous values and logical reasoning to ensure accuracy.

## Reworking values

I'm re-evaluating means, ranges, and SDs for each archetype's arm, using approximate calculations to reimagine row consistency scores amidst the constraints.

## Calculating row consistency

I'm working through 5 arms, 3 trials each, to estimate overall consistency. Using mean and standard deviation, I find scores for PFS and OS estimates, noting variation and precision.

## Calculating variability

I'm curious about ARCH-05's row consistency. The squared differences sum to 4.191, leading to a variance of 0.299 and a standard deviation of 0.546.

## Establishing the mean

Summing values across arms and trials to find the mean, then calculating squared deviations for variance and standard deviation, and finally computing the row consistency score.

## Tallying row scores

OK, let me see. Compiling row and column statistics, I'm ensuring precise columns and rows for each archetype. Using bullet points for a clear distribution of archetypes.

## Assessing variability

I'm working through ARCH-06's OS row consistency. Calculating mean and standard deviation, I find an 8.7 consistency score, matching past results.

## Figuring out row consistency

I'm calculating mean, range, and standard deviation for ARCH-07 OS row. This helps in creating a precise table and ensuring clarity and accuracy throughout.

## Synthesizing insights

I'm gathering data on KRAS-mutant and gBRCA-mutant groups. By calculating means, ranges, and standard deviations, I'm assessing variability to understand the distribution of outcomes effectively.

## Tracking progress

I'm working through median PFS, median OS, and PFS HR values. Calculating means and standard deviations, I provide example rows and focus on precision and clarity.