
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

15.S41.VER.01.P32

Prompt 32

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 Research Generation: 20 Pages, July 9, 2025

“Trial 3”

Prompt 32:

Your task is to generate a direct, head-to-head comparison that quantifies the correlation and consistency between the summary report tables and the attached log file csv. Show 3 human verifiable sample calculations below each new table, along with data sources: ie. Patient 000042, Table 5-1, etc..

Present your findings exclusively in the following 6 tables. Each table must have the specified dimensions, row names (R1, R2...), and column names (C1, C2...). The "Calculated" columns must be derived by analyzing the full attached log file csv, while the "Reported" columns must extract data directly from the clinical study report text and its tables. The final column in each table should provide a quantitative critique of the alignment between the two sources.

Table 1: Overall Cohort Distribution Verification (6R x 4C)

- **R1:** Arm A
- **R2:** Arm B
- **R3:** Arm C
- **R4:** Arm D
- **R5:** Arm E
- **R6:** Total
- **C1:** Arm/Group
- **C2:** Patient Count (per CSR Section 4)
- **C3:** Patient Count (Calculated from Log)
- **C4:** Discrepancy (C3 - C2)
- **+3** Sample Calculations, verifiable with sources

Table 2: Baseline Characteristics Correlation Check (Focus on Arm A) (5R x 4C)

- **R1:** Mean Age (years)
- **R2:** Stage IV (%)
- **R3:** ECOG 1 (%)
- **R4:** KRAS-mutant (%)
- **R5:** gBRCA-mutant (%)
- **C1:** Characteristic
- **C2:** Reported Value (Table 5-1)
- **C3:** Calculated Value (from Log)
- **C4:** Deviation (Absolute Difference)
- **+3** Sample Calculations, verifiable with sources

Table 3: Median Overall Survival (OS) Correlation (5R x 4C)

- **R1:** Arm A
- **R2:** Arm B
- **R3:** Arm C
- **R4:** Arm D
- **R5:** Arm E
- **C1:** Treatment Arm
- **C2:** Reported Median OS (months, Table 6-1)
- **C3:** Calculated Median OS (months, from Log time_to_death)
- **C4:** Difference (months)
- **+3** Sample Calculations, verifiable with sources

Table 4: Median Progression-Free Survival (PFS) Correlation (5R x 4C)

- **R1:** Arm A
- **R2:** Arm B
- **R3:** Arm C
- **R4:** Arm D
- **R5:** Arm E
- **C1:** Treatment Arm
- **C2:** Reported Median PFS (months, Table 6-1)
- **C3:** Calculated Median PFS (months, from Log time_to_progression_or_death)
- **C4:** Difference (months)
- **+3** Sample Calculations, verifiable with sources

Table 5: 12-Month Overall Survival Rate Verification (5R x 4C)

- **R1:** Arm A
- **R2:** Arm B
- **R3:** Arm C
- **R4:** Arm D
- **R5:** Arm E
- **C1:** Treatment Arm
- **C2:** Reported 12-Month OS Rate (% , Table 6-1)
- **C3:** Calculated 12-Month OS Rate (% , from Log time_to_death > 12)
- **C4:** Difference (%)
- **+3** Sample Calculations, verifiable with sources

Table 6: Grade ≥3 Adverse Event Incidence Verification (5R x 4C)

- **R1:** Arm A
- **R2:** Arm B
- **R3:** Arm C
- **R4:** Arm D
- **R5:** Arm E
- **C1:** Treatment Arm
- **C2:** Reported ≥G3 AE Rate (% , Table 7-1)
- **C3:** Calculated ≥G3 AE Rate (% , from Log time_to_first_G3_AE ≤ 24)
- **C4:** Difference (%)

- **+3** Sample Calculations, verifiable with sources

“Start Tables”

Table 5-1: Baseline Characteristics by Arm. Each value is a mean or percentage for the given arm (N=20,000 per arm). All arms were well balanced with respect to demographics and disease features.

| Arm | Age (years, mean) | Stage IV (%) | ECOG 0 (%) | ECOG 1 (%) | ECOG 2 (%) | KRAS-mu tant (%) | gBRCA-mu tant (%) | CA19-9 (U/mL, mean) |
|-------|-------------------------|-----------------|---------------|---------------|---------------|---------------------|----------------------|---------------------------|
| Arm A | 66.4 | 88.1 | 20.4 | 73.5 | 6.1 | 90.9 | 5.0 | 5830.8 |
| Arm B | 66.3 | 87.8 | 19.8 | 74.1 | 6.1 | 90.9 | 5.0 | 5774.0 |
| Arm C | 66.3 | 88.0 | 20.2 | 73.9 | 6.0 | 90.8 | 5.0 | 5842.1 |
| Arm D | 66.2 | 88.1 | 19.6 | 74.5 | 5.9 | 91.0 | 5.0 | 5816.4 |
| Arm E | 66.2 | 87.9 | 20.0 | 74.1 | 5.9 | 90.9 | 5.0 | 5882.9 |

Table 6-1: Primary Efficacy Outcomes by Arm. Efficacy endpoints were analyzed by Kaplan-Meier method (24-month maximum follow-up). Hazard ratios (HR) are computed for each experimental arm vs. control (Arm E).

| Arm | Median PFS (mo) | Median OS (mo) | 12-month OS Rate (%) | PFS HR vs Control | OS HR vs Control |
|-------|--------------------|-------------------|-------------------------|----------------------|---------------------|
| Arm A | 4.6 | 8.8 | 38.8 | 0.68 | 0.70 |
| Arm B | 3.2 | 6.8 | 29.7 | 0.97 | 0.90 |
| Arm C | 3.9 | 7.2 | 31.5 | 0.80 | 0.85 |

| | | | | | |
|-------|-----|-----|------|------|------|
| Arm D | 4.2 | 8.1 | 34.9 | 0.75 | 0.76 |
| Arm E | 3.1 | 6.1 | 25.1 | 1.00 | 1.00 |

Table 7-1: Global Safety Summary by Arm. Percentage of patients who experienced any Grade ≥ 3 adverse event within 24 months of follow-up (cumulative incidence).

| Arm | Any ≥ 3 AE (%) |
|-----|---------------------|
|-----|---------------------|

| | |
|-------|------|
| Arm A | 94.6 |
|-------|------|

| | |
|-------|------|
| Arm B | 88.4 |
|-------|------|

| | |
|-------|------|
| Arm C | 86.1 |
|-------|------|

| | |
|-------|------|
| Arm D | 81.6 |
|-------|------|

| | |
|-------|------|
| Arm E | 76.5 |
|-------|------|

“End Tables”

[Above: Tables 5-1, 6-1, 7-1]

[S40.TRL.15.P30.LOG.csv]

Output:

Thanks for uploading the log file. Before I begin, could you confirm:

1. Is the full clinical study report available beyond Tables 5-1, 6-1, and 7-1 (e.g., Section 4 for cohort counts)? If yes, please upload or specify.
2. Should calculations be based on all patients listed in the log, or filtered by inclusion criteria (e.g., time_to_death not null, Arm assignment known, etc.)?
3. Are there any known patient exclusions or censoring rules (e.g., censoring at 24 months for survival metrics)?

Once I have this, I'll begin generating the six verification tables you requested.

Input:
Use provided prompt.

Generation 41 (Formatted)
ChatGPT o3-pro Research
1390 words, 19 minutes 35.26 seconds

Understood. I will begin a comprehensive head-to-head verification comparing reported clinical study data (from Tables 5-1, 6-1, and 7-1) to the patient-level log file. This includes cross-validating cohort sizes, baseline characteristics, efficacy outcomes (OS and PFS), 12-month survival rates, and serious adverse events. For each table, I will also show 3 human-verifiable sample calculations tied to specific patients and sources.

I'll let you know when the analysis is complete.

Table 1: Overall Cohort Distribution Verification (6R × 4C)

| Arm/Group | Patient Count (per CSR Section 4) | Patient Count (Calculated from Log) | Discrepancy (C3 – C2) |
|-----------|-----------------------------------|-------------------------------------|-----------------------|
| Arm A | 20,000 | 20,000 | 0 |
| Arm B | 20,000 | 20,000 | 0 |
| Arm C | 20,000 | 20,000 | 0 |
| Arm D | 20,000 | 20,000 | 0 |
| Arm E | 20,000 | 20,000 | 0 |
| Total | 100,000 | 100,000 | 0 |

- **Arm A Count:** The CSR reports Arm A had 20,000 patients. The log file confirms 20,000 entries labeled Arm A, yielding no discrepancy.

- **Arm E Count:** Similarly, Arm E is listed with 20,000 patients in the CSR and 20,000 in the log, again a perfect match (difference 0).
- **Total Cohort:** Each of the five arms contains 20,000 patients, so the total population is $5 \times 20,000 = 100,000$ in both sources. The log's grand total of 100,000 patients equals the CSR's total (no missing or extra patients).

Table 2: Baseline Characteristics Correlation Check (Focus on Arm A) (5R × 4C)
(Arm A: N=20,000 patients)

| Characteristic | Reported Value (CSR Table 5-1) | Calculated Value (Log Data) | Deviation (Absolute Difference) |
|------------------|-----------------------------------|--------------------------------|------------------------------------|
| Mean Age (years) | 66.4 | 66.3 | 0.1 |
| Stage IV (%) | 88.1% | 88.2% | 0.1% |
| ECOG 1 (%) | 73.5% | 73.9% | 0.4% |
| KRAS-mutant (%) | 90.9% | 5.1% | 85.8% |
| gBRCA-mutant (%) | 5.0% | 4.9% | 0.1% |

- **Stage IV Disease:** The CSR shows 88.1% of Arm A patients were Stage IV. The log calculation finds 88.2% (17,647/20,000) Stage IV – virtually identical (0.1% difference). For example, Patient **015237** in Arm A is Stage IV (stage_iv = 1) per the log.
- **ECOG Performance Status 1:** The report lists 73.5% ECOG 1 in Arm A, and the log yields 73.9%. Patient **015237** has ECOG 1 in the log (ecog = 1), exemplifying this category. The 0.4% deviation is negligible.
- **KRAS Mutation Discrepancy:** The CSR notes ~90.9% of Arm A had KRAS mutations, whereas the log indicates only 5.1% with a **KRAS G12C** mutation. This 85.8% gap suggests the report counted all KRAS mutations, while the log tracked only the G12C subset. (For instance, Patient **015237** has **kras_g12c = 0**, i.e. no G12C mutation.)

Table 3: Median Overall Survival (OS) Correlation (5R × 4C)
(Overall Survival measured up to 24 months)

| Treatment Arm | Reported Median OS (months, CSR Table 6-1) | Calculated Median OS (months, Log time_to_death) | Difference (months) (C3 – C2) |
|---------------|--|--|-------------------------------|
| Arm A | 8.8 | 9.0 | +0.2 |
| Arm B | 6.8 | 6.7 | –0.1 |
| Arm C | 7.2 | 7.2 | 0.0 |
| Arm D | 8.1 | 8.0 | –0.1 |
| Arm E | 6.1 | 6.1 | 0.0 |

- Arm A Example:** The reported median OS for Arm A is 8.8 months. The log data give a median of ~8.97 months (very close) for Arm A. Notably, Patient **015237** in Arm A died at 8.8 months, essentially matching the median OS time from the CSR.
- Minor Differences:** All arms show excellent alignment. For example, Arm B's median OS is 6.7 months by log vs 6.8 months reported – a mere 0.1 month difference. Arm D is similarly close (8.0 vs 8.1 months). These tiny discrepancies (~0.1–0.2 mo) are negligible in practical terms.
- Control Arm E:** The control arm's median OS in the log is 6.085 months (~6.1), exactly matching the reported 6.1 months. This perfect agreement (0.0 month difference) further demonstrates the consistency between the CSR and log data.

Table 4: Median Progression-Free Survival (PFS) Correlation (5R × 4C)
(Progression-Free Survival measured up to 24 months)

| Treatment Arm | Reported Median PFS (months, CSR Table 6-1) | Calculated Median PFS (months, Log time_to_progression_or_death) | Difference (months) (C3 – C2) |
|---------------|---|--|-------------------------------|
| Arm A | 4.6 | 4.6 | 0.0 |

| | | | |
|-------|-----|-----|------|
| Arm B | 3.2 | 3.3 | +0.1 |
| Arm C | 3.9 | 3.9 | 0.0 |
| Arm D | 4.2 | 4.1 | −0.1 |
| Arm E | 3.1 | 3.1 | 0.0 |

- Arm A Consistency:** Arm A's median PFS is 4.6 months in the CSR and 4.56 months by log calculation – effectively identical (difference 0.0 mo after rounding). This indicates the progression timing in the data matches the reported value closely.
- Typical Variance:** The other arms likewise show only trivial differences. For instance, Arm D has a 4.06 month median PFS in the log vs 4.2 months reported, a difference of −0.14 months (rounded to −0.1). Such a small gap is within rounding error.
- Overall Pattern:** All calculated median PFS values are within ~0.1 month of the CSR figures, demonstrating strong concordance. Even the shortest median (Arm E, ~3.08 mo log vs 3.1 mo reported) matches up to one decimal place.

Table 5: 12-Month Overall Survival Rate Verification (5R × 4C)

(Percentage of patients surviving ≥12 months)

| Treatment Arm | Reported 12-Month OS Rate (% , CSR Table 6-1) | Calculated 12-Month OS Rate (% , from Log <code>time_to_death > 12</code>) | Difference (%) (C3 – C2) |
|---------------|---|--|--------------------------|
| Arm A | 38.8% | 39.8% | +1.0 |
| Arm B | 29.7% | 29.5% | −0.2 |
| Arm C | 31.5% | 31.3% | −0.2 |
| Arm D | 34.9% | 35.0% | +0.1 |

| | | | |
|-------|-------|-------|------|
| Arm E | 25.1% | 25.4% | +0.3 |
|-------|-------|-------|------|

- Arm A Survival:** The CSR reports 38.8% of Arm A patients were alive at 12 months. The log shows ~39.8% ($\approx 7,965/20,000$) surviving past 12 months, a difference of +1.0%. This small gap may be due to rounding or censoring method.
- Individual Cases:** For example, Patient **000005** in Arm A was alive at the 24-month follow-up (time_to_death = 24.0, no death by 12 mo), so they count toward the 12-month OS rate. In contrast, Patient **000006** died at 1.88 months and does not contribute to the 1-year survivors. These cases illustrate how the log data determine the 12-month survival metric.
- Across Arms:** All arms show very close agreement between reported and calculated 1-year OS rates. Even the control Arm E has 25.4% in log vs 25.1% reported, a difference of only 0.3%. Such tight alignment (differences within ~1% or less) indicates the survival status data in the log faithfully reproduces the CSR values.

Table 6: Grade ≥ 3 Adverse Event Incidence Verification (5R \times 4C)

(Patients with any Grade ≥ 3 adverse event within 24 months)

| Treatment Arm | Reported $\geq G3$ AE Rate (% , CSR Table 7-1) | Calculated $\geq G3$ AE Rate (% , from Log <code>time_to_first_G3_AE \leq 24</code>) | Difference (%) (C3 – C2) |
|---------------|--|---|-----------------------------|
| Arm A | 94.6% | 94.6% | 0.0 |
| Arm B | 88.4% | 87.6% | –0.8 |
| Arm C | 86.1% | 85.4% | –0.7 |
| Arm D | 81.6% | 81.1% | –0.5 |
| Arm E | 76.5% | 76.3% | –0.2 |

- Arm B Example:** The CSR indicates 88.4% of Arm B had \geq Grade 3 AEs. The log shows 87.6% (17,522/20,000 patients) with $\geq G3$ AEs in Arm B, about 0.8% fewer. This corresponds to roughly 158 fewer patients with severe

AEs in the log than reported, a minor discrepancy.

- **Early vs. No AE Cases:** The log captures timing of severe AEs. For instance, Patient **000002** (Arm D) experienced a Grade ≥ 3 AE at 0.75 months (very early toxicity), whereas Patient **000020** (Arm E) had **no** Grade ≥ 3 AE by 24 months (**time_to_first_G3_AE = 24.0**). Such entries verify how patients contribute to the calculated rates – one is counted in the incidence, the other is not.
- **Consistency:** Overall, the calculated Grade ≥ 3 AE incidence per arm is within $<1\%$ of the CSR values across all arms. For example, Arm E's rate is 76.3% by log vs 76.5% reported. In Arm E, the log shows 4,742 out of 20,000 patients had no ≥ 3 event (thus 15,258 had events, 76.3%), which aligns closely with the reported 76.5%. The minimal differences suggest a high correlation between the safety data in the CSR and the raw log.

Generation 41 (Plain Text)

ChatGPT o3-pro Research

1393 words, 19 minutes 35.26 seconds

Understood. I will begin a comprehensive head-to-head verification comparing reported clinical study data (from Tables 5-1, 6-1, and 7-1) to the patient-level log file. This includes cross-validating cohort sizes, baseline characteristics, efficacy outcomes (OS and PFS), 12-month survival rates, and serious adverse events. For each table, I will also show 3 human-verifiable sample calculations tied to specific patients and sources.

I'll let you know when the analysis is complete.

****Table 1: Overall Cohort Distribution Verification (6R \times 4C)****

| Arm/Group Patient Count (per CSR Section 4) Patient Count (Calculated from Log) Discrepancy (C3 – C2) | | | | | |
|---|-------------|-------------|-------|--|--|
| ----- | ----- | ----- | ----- | | |
| Arm A | 20,000 | 20,000 | 0 | | |
| Arm B | 20,000 | 20,000 | 0 | | |
| Arm C | 20,000 | 20,000 | 0 | | |
| Arm D | 20,000 | 20,000 | 0 | | |
| Arm E | 20,000 | 20,000 | 0 | | |
| **Total** | **100,000** | **100,000** | **0** | | |

*****Arm A Count:**** The CSR reports Arm A had 20,000 patients. The log file confirms 20,000 entries labeled Arm A, yielding no discrepancy.

*****Arm E Count:**** Similarly, Arm E is listed with 20,000 patients in the CSR and 20,000 in the log, again a perfect match (difference 0).

*****Total Cohort:**** Each of the five arms contains 20,000 patients, so the total population is $5 \times 20,000 = 100,000$ in both sources. The log's grand total of 100,000 patients equals the CSR's total (no missing or extra patients).

****Table 2: Baseline Characteristics Correlation Check (Focus on Arm A) (5R × 4C)****

(Arm A: N=20,000 patients)

| Characteristic | Reported Value (CSR Table 5-1) | Calculated Value (Log Data) | Deviation (Absolute Difference) | |
|------------------|--------------------------------|-----------------------------|---------------------------------|--|
| ----- | ----- | ----- | ----- | |
| Mean Age (years) | 66.4 | 66.3 | 0.1 | |
| Stage IV (%) | 88.1% | 88.2% | 0.1% | |
| ECOG 1 (%) | 73.5% | 73.9% | 0.4% | |
| KRAS-mutant (%) | 90.9% | 5.1% | 85.8% | |
| gBRCA-mutant (%) | 5.0% | 4.9% | 0.1% | |

*****Stage IV Disease:**** The CSR shows 88.1% of Arm A patients were Stage IV. The log calculation finds 88.2% (17,647/20,000) Stage IV – virtually identical (0.1% difference). For example, Patient ****015237**** in Arm A is Stage IV (stage_iv = 1) per the log.

*****ECOG Performance Status 1:**** The report lists 73.5% ECOG 1 in Arm A, and the log yields 73.9%. Patient ****015237**** has ECOG 1 in the log (ecog = 1), exemplifying this category. The 0.4% deviation is negligible.

*****KRAS Mutation Discrepancy:**** The CSR notes ~90.9% of Arm A had KRAS mutations, whereas the log indicates only 5.1% with a ****KRAS G12C**** mutation. This 85.8% gap suggests the report counted all KRAS mutations, while the log tracked only the G12C subset. (For instance, Patient ****015237**** has kras_g12c = 0, i.e. no G12C mutation.)

****Table 3: Median Overall Survival (OS) Correlation (5R × 4C)****

(Overall Survival measured up to 24 months)

| Treatment Arm | Reported Median OS (months, CSR Table 6-1) | Calculated Median OS (months, Log `time_to_death`) | Difference (months) (C3 – C2) | |
|---------------|--|--|-------------------------------|--|
| ----- | ----- | ----- | ----- | |
| Arm A | 8.8 | 9.0 | +0.2 | |
| Arm B | 6.8 | 6.7 | -0.1 | |
| Arm C | 7.2 | 7.2 | 0.0 | |
| Arm D | 8.1 | 8.0 | -0.1 | |
| Arm E | 6.1 | 6.1 | 0.0 | |

*****Arm A Example:**** The reported median OS for Arm A is 8.8 months. The log data give a median of ~8.97 months (very close) for Arm A. Notably, Patient ****015237**** in Arm A died at 8.8 months, essentially matching the median OS time from the CSR.

*****Minor Differences:**** All arms show excellent alignment. For example, Arm B's median OS is 6.7 months by log vs 6.8 months reported – a mere 0.1 month difference. Arm D is similarly close (8.0 vs 8.1 months). These tiny discrepancies (~ 0.1 – 0.2 mo) are negligible in practical terms.

*****Control Arm E:**** The control arm's median OS in the log is 6.085 months (~ 6.1), exactly matching the reported 6.1 months. This perfect agreement (0.0 month difference) further demonstrates the consistency between the CSR and log data.

****Table 4: Median Progression-Free Survival (PFS) Correlation (5R \times 4C)****

(Progression-Free Survival measured up to 24 months)

| Treatment Arm | Reported Median PFS (months, CSR Table 6-1) | Calculated Median PFS (months, Log
`time_to_progression_or_death`) | Difference (months) (C3 – C2) |

| ----- | ----- | ----- | ----- |
|-------|-------|-------|-------|
| Arm A | 4.6 | 4.6 | 0.0 |
| Arm B | 3.2 | 3.3 | +0.1 |
| Arm C | 3.9 | 3.9 | 0.0 |
| Arm D | 4.2 | 4.1 | -0.1 |
| Arm E | 3.1 | 3.1 | 0.0 |

*****Arm A Consistency:**** Arm A's median PFS is 4.6 months in the CSR and 4.56 months by log calculation – effectively identical (difference 0.0 mo after rounding). This indicates the progression timing in the data matches the reported value closely.

*****Typical Variance:**** The other arms likewise show only trivial differences. For instance, Arm D has a 4.06 month median PFS in the log vs 4.2 months reported, a difference of -0.14 months (rounded to -0.1). Such a small gap is within rounding error.

*****Overall Pattern:**** All calculated median PFS values are within ~ 0.1 month of the CSR figures, demonstrating strong concordance. Even the shortest median (Arm E, ~ 3.08 mo log vs 3.1 mo reported) matches up to one decimal place.

****Table 5: 12-Month Overall Survival Rate Verification (5R \times 4C)****

(Percentage of patients surviving ≥ 12 months)

| Treatment Arm | Reported 12-Month OS Rate (% , CSR Table 6-1) | Calculated 12-Month OS Rate (% , from Log
`time_to_death > 12`) | Difference (%) (C3 – C2) |

| ----- | ----- | ----- | ----- |
|-------|-------|-------|-------|
| Arm A | 38.8% | 39.8% | +1.0 |
| Arm B | 29.7% | 29.5% | -0.2 |
| Arm C | 31.5% | 31.3% | -0.2 |
| Arm D | 34.9% | 35.0% | +0.1 |
| Arm E | 25.1% | 25.4% | +0.3 |

*****Arm A Survival:**** The CSR reports 38.8% of Arm A patients were alive at 12 months. The log shows ~39.8% (≈7,965/20,000) surviving past 12 months, a difference of +1.0%. This small gap may be due to rounding or censoring method.

*****Individual Cases:**** For example, Patient ****000005**** in Arm A was alive at the 24-month follow-up (time_to_death = 24.0, no death by 12 mo), so they count toward the 12-month OS rate. In contrast, Patient ****000006**** died at 1.88 months and does not contribute to the 1-year survivors. These cases illustrate how the log data determine the 12-month survival metric.

*****Across Arms:**** All arms show very close agreement between reported and calculated 1-year OS rates. Even the control Arm E has 25.4% in log vs 25.1% reported, a difference of only 0.3%. Such tight alignment (differences within ~1% or less) indicates the survival status data in the log faithfully reproduces the CSR values.

****Table 6: Grade ≥3 Adverse Event Incidence Verification (5R × 4C)****
 (Patients with any Grade ≥3 adverse event within 24 months)

| Treatment Arm | Reported ≥G3 AE Rate (% , CSR Table 7-1) | Calculated ≥G3 AE Rate (% , from Log 'time_to_first_G3_AE ≤ 24') | Difference (%) (C3 – C2) |
|---------------|--|---|--------------------------|
| Arm A | 94.6% | 94.6% | 0.0 |
| Arm B | 88.4% | 87.6% | –0.8 |
| Arm C | 86.1% | 85.4% | –0.7 |
| Arm D | 81.6% | 81.1% | –0.5 |
| Arm E | 76.5% | 76.3% | –0.2 |

*****Arm B Example:**** The CSR indicates 88.4% of Arm B had ≥Grade 3 AEs. The log shows 87.6% (17,522/20,000 patients) with ≥G3 AEs in Arm B, about 0.8% fewer. This corresponds to roughly 158 fewer patients with severe AEs in the log than reported, a minor discrepancy.

*****Early vs. No AE Cases:**** The log captures timing of severe AEs. For instance, Patient ****000002**** (Arm D) experienced a Grade ≥3 AE at 0.75 months (very early toxicity), whereas Patient ****000020**** (Arm E) had ****no**** Grade ≥3 AE by 24 months ('time_to_first_G3_AE = 24.0'). Such entries verify how patients contribute to the calculated rates – one is counted in the incidence, the other is not.

*****Consistency:**** Overall, the calculated Grade ≥3 AE incidence per arm is within <1% of the CSR values across all arms. For example, Arm E's rate is 76.3% by log vs 76.5% reported. In Arm E, the log shows 4,742 out of 20,000 patients had no ≥G3 event (thus 15,258 had events, 76.3%), which aligns closely with the reported 76.5%. The minimal differences suggest a high correlation between the safety data in the CSR and the raw log.

Research Summary Generation 41: 1846 Words

Clinical Report vs Log File Concordance Verification

ChatGPT

I'm working through the log file and report tables to ensure consistency across all specified tables and characteristics. This involves verifying patient counts, baseline characteristics, survival metrics, and adverse event rates, with detailed sample calculations for each table.

Searched for S40.TRL.15.P30.LOG.csv file

ChatGPT

OK, let me see. The file is at /mnt/data/S40.TRL.15.P30.LOG.csv. Using Python seems like the way to go.

Analyzed with python

ChatGPT

I'm figuring out column meanings for a CSV with 100,000 rows, exploring patient demographics, tumor stage, genetic mutations, and progression times.

Analyzed with python

ChatGPT

OK, let me see. The labels 'A' to 'E' for the arms map neatly to "Arm A" to "Arm E" in the table text. This helps clarify any prior confusion.

Analyzed with python

ChatGPT

I'm cross-checking the baseline metrics for Arm A, such as mean age, stage IV percentage, and specific KRAS and gBRCA mutations, with data from Table 5-1 and the log file.

Analyzed with python

ChatGPT

I noticed a significant deviation in the KRAS mutation percentage, with the log showing 5.095% while the baseline suggests 90.9%. This discrepancy implies a potential misunderstanding or mislabeling of mutation data.

Analyzed with python

ChatGPT

I'm actually pulling together that 5,000 out of 100,000 individuals have the 'kras_g12c' mutation, which is a solid 5% of the total population.

Analyzed with python

ChatGPT

I'm looking at the 5% figure aligning with the baseline table, suggesting 'gBRCA' is approximately 5%.

Analyzed with python

ChatGPT

I'm noting the discrepancy between the KRAS-mutant (%) reported and the specific KRAS G12C mutation tracked in the log, aiming to highlight statistical differences for each row.

Analyzed with python

ChatGPT

Evaluating if "24" in 'time_to_death' indicates death or censoring, especially for median OS calculation. Hm, that's interesting!

I'm comparing calculated and reported outcomes, noting small differences, likely due to rounding or minimal variances. Clarification on the difference sign might further refine accuracy.

Analyzed with python

ChatGPT

It's interesting to see the median overall PFS at 3.71, slightly above the control arm. The improved PFS in other arms might be a contributing factor.

Analyzed with python

ChatGPT

I'm piecing together the 12-month overall survival rates for arms A through E. Ensuring precise interpretation, I include only those surviving strictly beyond 12 months to maintain accurate survival estimates.

Analyzed with python

ChatGPT

I'm comparing the tabulated data to the reported figures, noticing minor discrepancies within 1% range, except for a 1.0% higher figure in category A.

Analyzed with python

ChatGPT

Trying to figure out if time_to_first_G3_AE=24 means an event happened or if it's censored. Clarifying event status at 24 months.

Analyzed with python

ChatGPT

I'm observing a minor gap between calculated and reported values, around 0.2 to 0.8 points, with a notable difference for B (~0.79).

Analyzed with python

ChatGPT

I'm thinking about using ≤ 23.999 instead of < 24 to reduce rounding differences when calculating time to the first Grade 3 adverse event, as exact 24 might not be no events.

Analyzed with python

ChatGPT

I'm curious about the arm times and their calculations.

Analyzed with python

ChatGPT

I'm verifying the calculated and reported percentages, ensuring alignment to one decimal place, and preparing sample calculations for accuracy.

Analyzed with python

ChatGPT

I'm working through if patient IDs might use padding or prefixes, as the log file shows no padding. This raises questions about how these IDs are formatted.

I'm exploring ways to quote specific data points from the log, compare them with reported results, and ensure the citations are substantiated and accurate.

Analyzed with python

ChatGPT

OK, let me see. Both the 10,000th and 10,001st values are 8.97. This defines the median, and I can start thinking about this data point more deeply.

Analyzed with python

ChatGPT

Hmm, let me see. Eleven patients died at 8.97 months, a curious but precise figure, probably the median. This precision is intriguing, though the exact cause is unclear.

Analyzed with python

ChatGPT

I'm mapping out the survival timeline of patient ID 801, noting a median overall survival of approximately 8.97 months for Arm A.

Analyzed with python

Analyzed with python

ChatGPT

I'm piecing together the median OS for Arm A, noting that Patient 15237's death at 8.8 months aligns perfectly. To cite this, I need to pinpoint the exact CSV line.

Searched for 15237 8.8

Searched for patient_id 15237

ChatGPT

Wondering about a potential issue with CSV indexing. Considering opening it via browse.

Searched for S40 TRL 15 P30 LOG

ChatGPT

I'm piecing together why the CSV isn't recognized for textual search in the environment. Direct line reading or converting it to a smaller text might work better.

Analyzed with python

ChatGPT

Trying a new approach by searching within the user_files for the term "15237," which I remember being in the CSV snippet, to see if it includes any content.

Searched for 15237, A, ARCH-01

ChatGPT

I'm thinking through how to cite content accurately from the CSV file by using Python to read it line by line and extract relevant lines with citations.

Analyzed with python

ChatGPT

OK, let me see. Checking if 'patient_id' is orderly or if a comma or quote might be missing. I suspect the file is comma-separated, possibly random or sorted by another index.

Analyzed with python

ChatGPT

Considering how pandas reads CSVs, leading zeros may disappear. The printed row suggests numbers are numeric or converted to integers.

Analyzed with python

ChatGPT

OK, let me see. Tracking sequential IDs 1 to 100,000 and noting the file's likely sort method by archetype or arms. Progress is steady in identifying consistent patterns of association.

Analyzed with python

ChatGPT

I'm verifying the reasons behind sorting by patient_id and confirming the random arm distribution.

Analyzed with python

ChatGPT

I'm thinking through the patient's clinical details, including age, stage, treatments, and outcomes. The progression or death timing seems inconsistent, indicating a potential logic discrepancy.

Analyzed with python

ChatGPT

I'm examining 4.56 as the median.

Analyzed with python

ChatGPT

Mapping out how early death influences progression-free survival (PFS), despite the odd progression timing.

Analyzed with python

ChatGPT

I'm noting the progression and survival metrics for patient 000489 in Arm A, aligning it with median PFS and accurate OS rate calculations.

Analyzed with python

ChatGPT

I'm working out if 29.535% of 20000 is an integer. This calculation helps in understanding the numerical details.

Analyzed with python

ChatGPT

I'm checking the exact counts behind each percentage to ensure they fit perfectly into these tiers:

- 7965/20000
- 5907/20000
- 6256/20000
- 7000/20000
- 5028/20000

Analyzed with python

ChatGPT

Grouping 20,000 patients, 7,965 from Arm A survived beyond 12 months, closely aligning with the reported 38.8%.

Analyzed with python

ChatGPT

Patient 000005, from Arm A, shows a 24-month survival and a PFS of 6.4, striking a balance between living and disease progression.

Analyzed with python

ChatGPT

I'm compiling information on sample calculations for AE incidence by arm, highlighting variations in reporting and specific patient cases to illustrate these differences effectively.

Analyzed with python

ChatGPT

Verifying earlier records shows 000005 received the initial G3 AE 6.73 vaccine, but this doesn't meet our requirement. Progress is being made towards the needed confirmation.

Analyzed with python

ChatGPT

I'm compiling specific patient examples to illustrate different clinical events, ensuring proper citations from identified tables enhance the clarity and reliability of the data.

Analyzed with python

Searched for Table 5-1 Baseline Characteristics

ChatGPT

I'm rethinking the file path and considering `browser.open` with "source" for reading. The guidelines suggest using "user_files" with query "*" to list uploaded files.

I'm thinking about the CSV file I created in /mnt/data. It might not be counted as a user-uploaded file. Consider using `browser.open(cursor=some number)` to attach content.

Analyzed with python

ChatGPT

I'm noting the content lines that are subtly highlighted for easier citation.

I'm considering whether to use `browser.open` with a file path in `user_files` and how to cite the table properly.

Searched for csr_tables.txt file

ChatGPT

I'm looking at treating [48] as an open link. Copying output text into an answer seems unwise. Using `browse.open(cursor=some, id?)` is under consideration. For known table content, a footnote approach might be suitable.

Searched for Arm A 66.4 88.1 90.9 5830.8

ChatGPT

I'm noting the need to cite log file references and reported table values. The guidelines emphasize preserving citations in the specified format. Patient lines from the log and provided tables are to be referenced appropriately.

Analyzed with python

Searched for Arm B data

Analyzed with python

ChatGPT

I'm gathering information on the best way to cite tables and rows from the CSR to enhance the clarity and support the accuracy of the data distribution, considering the significance of specific lines and footnotes.

Analyzed with python

ChatGPT

I'm making sure the counts are the main data, leaving out other details like "Name: arm, dtype: int64" at the bottom. This keeps the data concise and to the point.

Analyzed with python

Searched for A B C 20000

ChatGPT

I'm working through placing citations for specific values in Tables 1 and 2, considering clarity and avoiding clutter, by using the same source and highlighting key data points.

Analyzed with python

ChatGPT

I noticed the data cut off at line 89, likely skipping the Arm E row. This suggests a need for review or adjustment.

Analyzed with python

ChatGPT

Noticing the omission of the previous OS HR for E, likely due to a blank line 96.

Analyzed with python

ChatGPT

I'm working through various citation methods, considering individual citations, row citations, and complete block citations for Tables 3-6 to ensure consistency and avoid confusion.

Analyzed with python

ChatGPT

Wondering if discrepancies in median OS/PFS calculations should be highlighted in tables, possibly using separate citations to contrast values.

Analyzed with python

ChatGPT

I'm adding detailed medians by arm and citing reference [61†L1-L5] to cover all median OS values.

Analyzed with python

ChatGPT

I'm starting to see a notable trend in the medians, currently at 62. This could indicate significant insights for further analysis.

Analyzed with python

ChatGPT

I'm examining something labeled [63], trying to retrieve info.

Analyzed with python

Analyzed with python

ChatGPT

Seven patients tragically succumbed on 8.8. Analyzing the pattern's significance highlights its unsettling nature.

Searched for subset15230-40

ChatGPT

I'm comparing discrepancies in median OS and PFS across experimental arms, focusing on absolute differences to highlight minimal or significant deviations. This helps in assessing the consistency of data between two sources.

I'm verifying sample calculations for median OS and PFS, using patient entries to confirm accuracy. This approach highlights deviations between CSR and log data, aiding in discrepancies identification.

Analyzed with python

ChatGPT

I'm piecing together citations and formatting to maintain consistency, separating tables from sample calculations, and avoiding extra commentary. Progress is steady, confirming adherence to guidelines and verifying line numbers.

Research completed in 17m · 0 sources

Citations Generation 41

All Sources