Datathon 1- Statistical Analysis Plan

Minoo Matbouriahi, Jifan Wang, Yazid Zalai

25 June 2025

Title: Oxaliplatin Added to Fluorouracil-Based Chemoradiotherapy and Postoperative Chemotherapy for Locally Advanced Rectal Cancer

1. Introduction, Study Design, and Objectives

This statistical analysis plan is established to reproduce the tables and figures from Rodel et al. 2015.

Study Outline

The CAO/ARO/AIO-04 study is a **multicentre, open-label, randomised, phase 3 trial**. It was conducted in **Germany**. The study compared the standard regimen established by the German CAO/ARO/AIO-94 trial (preoperative chemoradiotherapy with infusional fluorouracil, total mesorectal excision surgery, and postoperative chemotherapy with fluorouracil) with an investigational regimen in which **oxaliplatin was added to both preoperative chemoradiotherapy and postoperative chemotherapy**. The study included patients with rectal adenocarcinoma, clinically staged as **cT3–4 or any node-positive disease**. Randomisation was done with computer-generated block-randomisation codes stratified by centre, clinical T category (cT1–3 vs cT4), and clinical N category (cN0 vs cN1–2) without masking.

The trial was conducted in 88 centers across Germany. The details of the centres are listed in the study registration on the clinicaltrials.gov with number **NCT00349076** and on EU clinical trials resiter with the number **2006-002385-20**.

Objectives

Primary Objective

The primary trial objective is to evaluate whether the addition of oxaliplatin to standard fluorouracil-based pre- and postoperative therapy improves disease-free survival (DFS). We hypothesis that the DFS will improve from 75% in the control group to 82% in the treatment group.

Secondary Objective(s)

Secondary trial objectives are to evaluate the efficacy of additional oxaliplatin to the standard regimen on survival and tumour control. The analyses is exploratory for the secondary objectives.

2. Outcome Measures

Primary Endpoint

The primary endpoint is **disease-free survival**. The DFS is defined as the time from randomization to the first occurrence of non-radical surgery of the primary tumour (R2 resection), locoregional recurrence after R0/1 resection of the primary tumour, metastatic disease or progression, or death from any cause, whichever occurred first. Second non-colorectal malignancies were disregarded in the analyses of disease-free survival.

Secondary Endpoint(s)

The secondary objectives on survival and tumour control will be evaluated using the following endpoints:

1. **Overall survival (OS)** is defined as time from randomisation to death from any cause.
2. **Incidence of local and distant recurrence** is defined as the time between randomization and occurrence of any locoregional and distant recurrence, respectively, irrespective of whether this was a first event or not.

3. Sample Size

The sample size is determined based on the log-rank test between a 75% 3-year DFS in the control group and 82% 3-year DFS in the treatment group (i.e. hazard ratio 0.81). A total sample size of 1200 patients is required to detect a clinically relevant improvement in 3-year DFS as hypothesized, with 80% power at a two-sided alpha of 0.05.

4. Randomisation and Blinding

Patients are randomized to the treatment and control groups on a 1:1 ratio using a web interface hosted by the Department of Medical Informatics, Biometry, and Epidemiology, University of Erlangen (Erlangen, Germany). Patient assignments are faxed to each centre and ensured that the next assignment in the sequence is masked. Stratification factors include clinical T stage (cT1–3 vs cT4), clinical nodal status (cN0 vs cN1-2), and treatment centres. This study is open-label because of the different administration and schedules between arms. Below shows the study scheme.

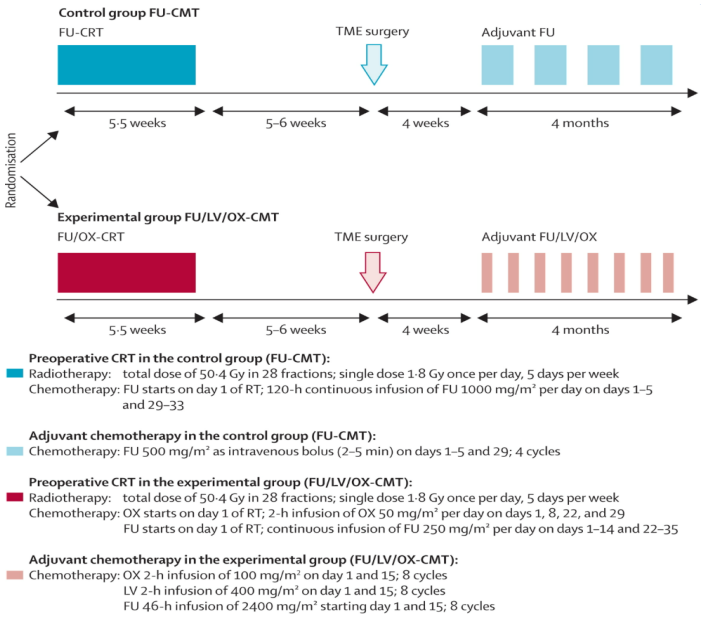

Figure 1. Study Scheme

5. Trial Management

Eighty-eight centres across Germany collaborates together throughout the study. A quality assurance programme, headed by reference institutions for surgery, radiotherapy, and chemotherapy, review the information submitted on case report forms. Participating surgeons, pathologists, radiation oncologists, and medical oncologists have training classes at study meetings twice per year. Compliance with protocol-defined standard operating procedures is reviewed centrally for two arbitrarily selected participants from each centre. The Randomisation code is generated centrally and distributed to each centre upon each enrolment.

A data monitoring committee monitors the safety of the trial on regular basis. Participants are monitored weekly during chemoradiotherapy and before each adjuvant treatment cycle, which

regard to vital signs and haematological and biochemical analyses. Doses were modified in response to toxicities according to predefined guidelines. On-site monitoring, source data verification, and database management were provided by WISP Research Institute

6. Data Sources and Analysis Populations

Data Sources

Patients were recruited between July 25, 2006, and February 6, 2010. A total of 1265 patients were initially enrolled from 88 centres in Germany.

Analysis Populations

The intention-to-treat (ITT) analysis is used for primary and secondary survival and cumulative incidence of recurrence endpoints. All randomly assigned patients who fulfilled the inclusion criteria are included in the efficacy analysis.

7. Statistical Analyses

General Considerations

The statistical uncertainty is reported using 95% confidence interval for all primary and secondary outcomes when applicable. The analyses are performed in R v3.1.2.

Baseline Characteristics

Baseline characteristics of participants (age, sex, clinical stage, tumour location, histology) are summarized within each arm. Continuous characteristics are reported in mean, standard deviation (SD), median, and interquartile range (IQR). Categorical characteristics are reported in counts and percentages including the number of missing data. The baseline characteristics are not tested between arms following the recommendation in the CONSORT statement [1].

Primary Endpoint Analysis

The DFS time is reported using Kaplan-Meier curve with a risk table per group. The primary hypothesis is tested using a log-rank test stratified by centre and clinical cN category on DFS time between groups on the ITT basis. Hazard ratios and corresponding 95% confidence intervals is estimated using a mixed-effects Cox proportional hazards model with centre and clinical cN category-specific random intercepts to account for the balancing variables in the randomization. The proportional hazard assumption is assessed graphically with the log-log plot. Lost-to-follow-ups are assumed to be non-informative and treated as right-censoring. The R package coxme v.2.2-3 is used for the mixed-effects Cox model.

Secondary Endpoint Analysis

Similar to the primary endpoint, overall survival, cumulative incidence of locoregional recurrences, and cumulative incidence of distant recurrences are reported using Kaplan-Meier curve with a risk table per group. Other secondary endpoints are reported by counts and percentages per group. Secondary endpoints are considered exploratory and thus not tested statistically since the trial is powered on the primary endpoint. Death is considered as the competing risk to incidences of local recurrence after R0/1 resection and distant recurrence. The cumulative incidences of local recurrence and distant recurrence are calculated using the cmprsk package v.2.2-7 in R.

Subgroup Analyses

Exploratory subgroup analysis of the primary endpoint, DFS, is conducted using a similar approach as in the primary analysis by characteristics including age, sex, Eastern Cooperative Oncology Group (ECOG) performance status, tumour distance from the anal verge, clinical T

category, clinical nodal status, ypT category, ypN category, completeness of local tumour resection, pathological TNM stage, and type of surgery. Baseline age is classified as ≤60, >60-70, and >70. ECOG performance status is further classified to 0 and 1-2. Clinical T category is condensed to cT2-3 and cT4. Clinical nodal status is condensed to cN0 and cN+. The subgroup analyses will be done on the available cases.

Additional Analysis

A similar mixed-effect Cox proportional hazard model is used for 3-year overall survival in additional analysis.

8. Handling of Missing Data

Participants missing from each intervention procedure are recorded and reported in the CONSORT diagram. They are included in the efficacy analyses. Lost to follow up causes the missing data in the endpoints. Those participants are treated as right-censoring in the primary and secondary analysis. Missing data in the baseline characteristics are reported. The missing pattern is checked through the descriptive statistics by missingness. No imputation is performed, and the missing is treated as a separate category for nominal variables when used in the primary analysis.

9. Tables and Figures

Tables:

1. Baseline Characteristics and Demographics
    - Patient demographics and disease characteristics by treatment arm
    - Stratification factors
    - Prior treatments and medical history
2. Efficacy Tables
    - Intention-to-treat analysis of first events for primary endpoint disease-free survival
    - Intention-to-treat analysis of all-cause deaths

Figures:

1. CONSORT diagram showing the number of patients through out the study
2. Survival Curves
    - Kaplan-Meier curves for disease-free survival
    - Kaplan-Meier curves for overall survival
    - Cumulative incidence curves for locoregional recurrence
    - Cumulative incidence curves for distant metastases
3. Forest Plots
    - Hazard ratios for DFS in predefined subgroups

11. Funders

This study was funded by a grant from the German Cancer Aid (Deutsche Krebshilfe). The funding source has no role in the study design, data collection, data analysis, data interpretation, report writing, and the decision to submit for publications,

12. References

1. Hopewell S, Chan A, Collins G S, HrÃ³bjartsson A, Moher D, Schulz K F et al. CONSORT 2025 explanation and elaboration: updated guideline for reporting randomised trials BMJ 2025; 389 :e081124 doi:10.1136/bmj-2024-081124
2. Fine, J. P., & Gray, R. J. (1999). A Proportional Hazards Model for the Subdistribution of a Competing Risk. Journal of the American Statistical Association, 94(446), 496–509. https://doi.org/10.1080/01621459.1999.10474144