

# **Longitudinal Study of DMARD+NSAID Combination Versus NSAID in Rheumatoid Arthritis Study**

Yunshen Bai yb2584

## **Abstract**

This report presents a comprehensive six-month longitudinal study comparing the efficacy of a Disease-Modifying Antirheumatic Drugs (DMARDs) and Nonsteroidal Anti-Inflammatory Drugs (NSAIDs) combination therapy (Treatment 2) versus NSAID monotherapy (Treatment 1). Employing the ACR20 as the primary endpoint, along with seven secondary endpoints where lower scores indicate improved conditions. The CMH tests show a significantly higher proportion of ACR20 responders in the treatment 2 group, both overall and when controlling for center effects. A series of Analysis of Covariance (ANCOVA) was performed to examine the effect of treatment, center, and baseline, revealing significant treatment effects for secondary endpoints and some interaction effects for treatment by center and treatment by baseline. However, time-to-event analysis did not show a significant difference in the time to discontinuation for lack of efficacy between the two treatments.

# Contents

Abstract.....	1
Introduction .....	3
Data Summary .....	3
Methods .....	4
Results .....	4
Categorical Data Analysis .....	4
ANCOVA .....	5
Time-to-event Analysis .....	7
Discussion and Conclusions .....	8
Reference .....	9
Appendix .....	9

## Introduction

Rheumatoid arthritis (RA) is a chronic inflammatory disorder that not only affects joints but also may involve various organs in the body. Understanding the effectiveness of different treatments is crucial to improving patient outcomes. This report aimed at comparing the efficacy of two treatment for rheumatoid arthritis over a six-month period. The treatments compared were a combination of Disease-Modifying Antirheumatic Drugs (DMARDs) with Nonsteroidal Anti-Inflammatory Drugs (NSAIDs) versus NSAIDs alone.

In this report, we perform longitudinal analysis on the rheumatoid arthritis study. Longitudinal analysis is a statistical method used to analyze data collected from the same subjects repeatedly over a period. This type of analysis is particularly relevant in studies where we want to assess how outcomes change over time or understand the time-dependent effects of treatments or interventions.

## Data Summary

The primary measure of success in this study was the proportion of patients who achieved the American College of Rheumatology 20% improvement criteria (ACR20). This primary endpoint was supplemented by seven secondary endpoints, including physician assessment of disease activity (PHYASMT), patient assessment of disease activity (PATASMT), number of painful joints (PAINJT), number of swollen joints (SWELLJT), visual analog pain scale (VAPS), C-reactive protein (CRP), and health assessment questionnaire (HAQ). These components are integral to a comprehensive evaluation of treatment impact, where a lower score typically indicates an improvement in condition.

ACR20				
Treatment	1	0	Total	Proportion (%)
1	24	64	88	27.27
2	45	48	93	48.39
Total	69	112	181	38.12

Table 1. Distribution of ACR20 and Proportions by Treatment

The study involves two treatment groups: one receiving DMARD+NSAID (treatment 2) and the other receiving NSAID alone (treatment 1). A total of 181 patients participated in the study, with 88 receiving treatment 1 and 93 receiving treatment 2. The analysis of the ACR20 response shows a higher proportion of responders in the treatment group receiving treatment 2 compared to the group receiving treatment 1.

The dataset exhibits instances of missing data, specifically in C-reactive protein and health assessment questionnaire. Handling missing data is a critical aspect of statistical analysis and can be approached through various methodologies. In this report, we use LOCF to handle missing data. Last Observation Carried Forward (LOCF) is a method used in longitudinal data analysis to handle missing data, particularly in clinical trials. It involves substituting missing values for a participant with the most recent non-missing value available for that participant. While it's a simple and commonly used method, it's important to be

aware that LOCF can introduce bias and should be applied with caution.

## Methods

In this report, endpoints will be analyzed including using method for categorical data analysis. We use CMH test to assess the association between two treatments. The Cochran-Mantel-Haenszel (CMH) test is a statistical analysis method used primarily in epidemiological studies to evaluate the association between an exposure and an outcome, while controlling for one or more confounding variables. It provides a way to test for an overall association between the exposure and the outcome across all strata. We will also perform an exploratory analysis using ACR20 with adjustment of center.

Analysis of Covariance (ANCOVA) will be employed to dissect the influence of treatment, center, and baseline measurements on the endpoints in this report. ANCOVA is a blend of analysis of variance (ANOVA) and regression that is used in statistical analysis. It allows you to compare one or more means while statistically controlling for variation associated with one or more covariate variables that can influence the dependent variable. ANCOVA assumes that the residuals are normally distributed and that the variances are equal across groups and homogeneity of regression slopes. The form of ANCOVA can be expressed as

$$Y_{ijk} = \mu + \tau_i + \gamma_j + \beta X_{ijk} + \varepsilon_{ijk}$$

where  $\mu$  is the overall mean,  $\tau_i$  is the  $i$ th treatment effect,  $\gamma_j$  is the  $j$ th center effect,  $X_{ijk}$  is the corresponding baseline and  $\varepsilon_{ijk}$  is the residual.

In this report, we will compare time to discontinuation for lack of efficacy using time-to-event analysis. Time-to-event analysis is a set of statistical methods used to analyze the expected duration of time until one or more events of interest occur. This type of analysis is particularly useful when subjects in a study are followed over time and the outcomes are not only whether an event occurred, but also when it occurred. The outcome of interest in this project is discontinuation for lack of efficacy. We use Log-Rank test to compare time to discontinuation for lack of efficacy of two treatments. The Log-Rank test is a non-parametric statistical test that is widely used in time-to-event analysis or survival analysis. The test is used to compare survival distributions of two or more independent groups.

## Results

### Categorical Data Analysis

In this report, we use CMH test to assess the association between two treatments. The hypotheses for this test are

- Null hypothesis ( $H_0$ ): There is no association between treatments and proportion of ACR20 responders
- Alternative Hypothesis ( $H_1$ ): There is an association between treatments and proportion of ACR20 responders

Test Group	$\chi^2$	DF	p value	OR
Treatment 1 vs Treatment 2	8.4982	1	0.0036	2.5000

Table 2. Result of CMH test on the association between two treatments and proportion of ACR20 responders

The p-value of 0.0036 is less than the significance level of 0.05, leading us to reject the null hypothesis. This suggests that there is a statistically significant association between the two treatments with respect to the proportion of ACR20 responders. The Odds Ratio (OR) of 2.500 indicates that the odds of being an ACR20 responder is 2.5 times higher in Treatment 2 compared to Treatment 1.

However, medical conditions in different centers may affect the effectiveness of treatment. Therefore, we will perform an CMH test with adjustment of center. The hypotheses for this test are

- Null hypothesis ( $H_0$ ): There is no association between treatments and proportion of ACR20 responders across the centers
- Alternative Hypothesis ( $H_1$ ): There is an association between treatments and proportion of ACR20 responders across the centers

Test Group	$\chi^2$	DF	p value	OR
Treatment 1 vs Treatment 2 (controlling for centers)	8.2713	1	0.0040	2.5009

Table 3. Result of CMH test on the association between two treatments and proportion of ACR20 responders across the centers

The p-value 0.0040 is less than the significance level of 0.05, leading us to reject the null hypothesis. This indicates that significant associations between treatment and the proportion of ACR20 responders when controlling for the center. The odds ratio of 2.5009 implies that the odds of being an ACR20 responder are 2.5009 times higher in treatment 2 compared to the treatment 1.

## ANCOVA

For these seven endpoints used to compute ACR20 response, we analyze them using ANCOVA – treatment, center, and baseline as covariates. Since PHYASMT and PATASMT variables, ANCOVA is not the appropriate analysis. Therefore, we use ANCOVA on other five secondary endpoints.

Endpoint	SV	DF	SS	MS	F value	p value
PAINJT	Treatment	1	289.3	289.33	1.8191	0.1792
	Center	26	14202.9	546.27	3.4346	<0.0001
	Baseline	1	2431.8	2431.82	15.2897	0.0001
	Residuals	174	27674.6	159.05		
SWELLJT	Treatment	1	875.4	875.44	10.9970	0.0011
	Center	26	17926.8	8.6612	8.6612	<0.0001
	Baseline	1	2778.97	2778.97	34.9085	<0.0001
	Residuals	174	13851.6	79.61		
VAPS	Treatment	1	243.4	243.4	4.0014	0.0467
	Center	26	3566.0	137.15	2.2607	0.0010
	Baseline	1	1766.1	1766.1	29.1100	<0.0001
	Residuals	174	10556.5	60.67		

<b>CRP</b>	<b>Treatment</b>	1	31.20	31.20	12.3079	0.0005
	<b>Center</b>	26	88.50	3.404	1.3426	0.1364
	<b>Baseline</b>	1	264.55	264.55	104.3461	<0.0001
	<b>Residuals</b>	174	441.14	2.535		
<b>HAQ</b>	<b>Treatment</b>	1	0.498	0.498	1.0591	0.3049
	<b>Center</b>	26	15.420	0.5931	1.2619	0.1907
	<b>Baseline</b>	1	0.061	0.061	0.1296	0.7192
	<b>Residuals</b>	174	81.777	0.4699		

Table 4. Result of ANCOVAs on secondary endpoints

The treatment effect on SWELLJT, VAPS and CRP are significant and not for others, indicating that the effect of the treatment on the endpoints is not consistent. Consequently, we aim to explore the interaction effects, specifically how the treatment interacts with the center and baseline values. This exploration could reveal potential effects that have not yet been identified. First, we use ANOVA F-test to check whether ANCOVAs with interaction have better fit.

- Null Hypothesis ( $H_0$ ): The model with interaction terms don't provides a significantly better fit to the data compared to the simpler model.
- Alternative Hypothesis ( $H_1$ ): The model with interaction terms provides a significantly better fit to the data compared to the simpler model.

<b>Endpoint</b>	<b>DF</b>	<b>F value</b>	<b>p value</b>
<b>PAINJT</b>	27	1.6677	0.0294
<b>SWELLJT</b>	27	2.4059	0.0005
<b>VAPS</b>	27	1.2632	0.1907
<b>CRP</b>	27	1.7209	0.0223
<b>HAQ</b>	27	0.5134	0.9776

Table 5. Result of ANOVA F test on ANCOVAs with and without interaction terms

ANCOVAs with interaction have better fit on PAINJT, SWELLJT and CRP. Therefore, we need to check the interaction terms on these three ANCOVAs.

<b>Endpoint</b>	<b>SV</b>	<b>DF</b>	<b>SS</b>	<b>MS</b>	<b>F value</b>	<b>p value</b>
<b>PAINJT</b>	<b>Treatment</b>	1	289.3	289.33	2.0076	0.1586
	<b>Center</b>	26	14202.9	546.27	3.7904	<0.0001
	<b>Baseline</b>	1	2431.8	2431.8	16.8739	<0.0001
	<b>Treatment:Center</b>	26	6466.3	248.70	1.7257	0.0232
	<b>Treatment:Baseline</b>	1	23.0	23.04	0.1599	0.6899
	<b>Residuals</b>	174	21185.3	144.12		
<b>SWELLJT</b>	<b>Treatment</b>	1	875.4	875.44	13.3960	0.0004
	<b>Center</b>	26	17926.8	689.49	10.5507	<0.0001
	<b>Baseline</b>	1	2779.0	2779.0	42.5241	<0.0001
	<b>Treatment:Center</b>	26	3560.9	136.96	2.0957	0.0032
	<b>Treatment:Baseline</b>	1	684.3	684.25	10.4705	0.0015
	<b>Residuals</b>	174	9606.5	65.35		
<b>CRP</b>	<b>Treatment</b>	1	31.20	31.20	13.6848	0.0003
	<b>Center</b>	26	88.50	3.404	1.4927	0.0723
	<b>Baseline</b>	1	264.55	264.55	116.0189	<0.0001
	<b>Treatment:Center</b>	26	63.46	2.441	1.0705	0.3830

<b>Treatment:Baseline</b>	1	42.49	42.29	18.6323	<0.0001
<b>Residuals</b>	174	335.19	2.280		

Table 6. Result of ANCOVAs on secondary endpoints with interaction terms

For PAINJT and SWELLJT, the p value of treatment by center interaction is less than significant level 0.05, which indicate that treatment by center interaction in these two ANCOVAs is significant. For SWELLJT and CRP, the p value of treatment by baseline interaction is less than significant level 0.05, which indicate that treatment by baseline interaction in these two ANCOVAs is significant.

## Time-to-event Analysis

In this report, we use time-to-event analysis to compare time to discontinuation for lack of efficacy between two treatments. We will employ survival curves to illustrate the variation in survival probabilities over days across both treatment groups.

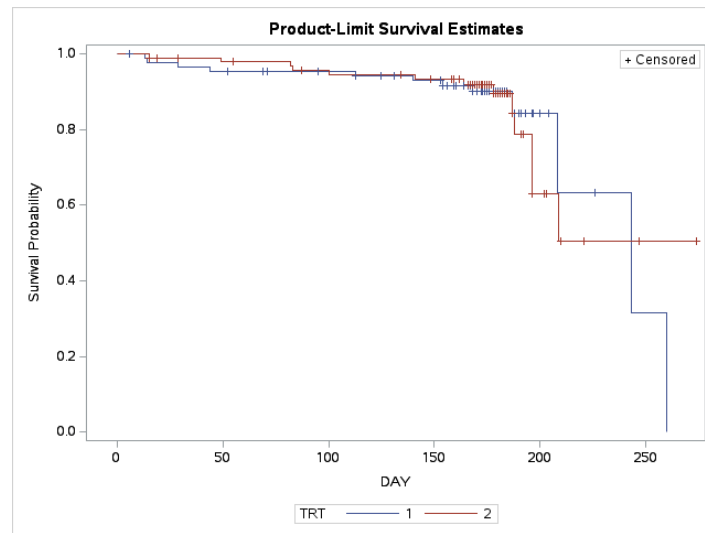


Figure 1. survival curves of survival probabilities over days across both treatment groups.

Both treatments start with a survival probability of 1, meaning all subjects were alive or event-free at the start of the observation period. The presence of censored data points ('+') suggests that for some subjects, the outcome or event was not observed. There is a noticeable decline in survival probability for Treatment 2 after around 200 days, indicating a higher number or a faster rate of events in this group compared to Treatment 1.

To test whether two curves are significantly different, we will use Log-Rank test to compare the survival distributions of two treatment groups.

Test Group	$\chi^2$	DF	p value
Treatment 1 vs Treatment 2	0.0008	1	0.9774

Table 7. Result of Log-Rank test

The p value for the log-rank test is 0.9774. This p-value is much higher than the conventional significance level of 0.05, which indicates that there is no significant difference in the survival distributions between the

two treatment groups when comparing the time to discontinuation for lack of efficacy.

## **Discussion and Conclusions**

The comprehensive statistical analysis presented in this report underscores the effectiveness of Treatment 2 (DMARD+NSAID) over Treatment 1 (NSAID) in managing rheumatoid arthritis over a six-month period. Our findings from the CMH test demonstrated a statistically significant association between the treatment types and the proportion of ACR20 responders, both overall and when controlling for center effects.

The ANCOVA results further elucidated the effect of treatment, center, and baseline on the secondary endpoints. We performed a closer examination of the interaction effects, particularly treatment by center interaction and treatment by baseline interaction. For PAINJT, there is a significant interaction between treatment and center. This suggests that the effect of the treatment on PAINJT may depend on the center where the treatment is administered. SWELLJT and CRP both show significant main effects of treatment. Additionally, for SWELLJT, both interactions are significant, which could mean that individual patient physical conditions or center medical conditions can affect how patients respond to the treatment regarding SWELLJT. For CRP, there is a significant main effect of treatment and baseline, but the interactions with the center are not significant. This suggests that effects do not seem to depend on the center or interact with each other in a significant way.

Time-to-event analysis using the Log-Rank test did not indicate a significant difference in the time to discontinuation for lack of efficacy between the two treatment groups. This may indicate that while the treatment 2 offers a higher probability of ACR20, it does not necessarily correlate with longer lasting efficacy before discontinuation.



## Reference

- [1] Cochran, W. G. (1954). Some Methods for Strengthening the Common  $\chi^2$  Tests. *Biometrics*, 10(4), 417–451. <https://doi.org/10.2307/3001616>
- [2] MANTEL N, HAENSZEL W (1959). Statistical Aspects of the Analysis of Data From Retrospective Studies of Disease, *JNCI: Journal of the National Cancer Institute*, Volume 22, Issue 4, April 1959, Pages 719–748, <https://doi.org/10.1093/jnci/22.4.719>
- [3] Khammar A, Yarahmadi M, Madadzadeh F. What Is Analysis of Covariance (ANCOVA) and How to Correctly Report Its Results in Medical Research? *Iran J Public Health*. 2020 May;49(5):1016-1017. PMID: 32953697; PMCID: PMC7475615.
- [4] Rutherford A (2001). Introducing ANOVA and ANCOVA: a GLM approach. 2nd ed. Sage, London, pp.: 105.
- [5] Jennifer L., Xiaofei W. (2021), Time-To-Event Data: An Overview and Analysis Considerations, *Journal of Thoracic Oncology*, Volume 16, Issue 7, 2021, Pages 1067-1074, ISSN 1556-0864, <https://doi.org/10.1016/j.jtho.2021.04.004>.

## Appendix

The R code used for this report is

```
1. library(car)
2. library(zoo)
3. library(tidyverse)
4. data=read.csv("DTA01.csv")
5. data=na.locf(data)
6. colnames(data)[1]="CENTER"
7. data$CENTER=as.factor(data$CENTER)
8. data$TRT=as.factor(data$TRT)
9. data$acr=as.factor(data$acr)
10.
11. data2=data|>
12.   group_by(PATIENT)|>
13.   filter(DAY==min(DAY))
14. data3=data|>
15.   group_by(PATIENT)|>
16.   filter(DAY==max(DAY))
17. DATA=left_join(data2,data3,by="PATIENT",keep=NULL)
18.
19. model_PAINJT=lm(PAINJT.y~TRT.x+CENTER.x+PAINJT.x,data = DATA)
20. summary(model_PAINJT)
21. anova(model_PAINJT)
22.
23. model_SWELLJT=lm(SWELLJT.y~TRT.x+CENTER.x+SWELLJT.x,data = DATA)
```

```

24. summary(model_SWELLJT)
25. anova(model_SWELLJT)
26.
27. model_VAPS=lm(VAPS.y~TRT.x+CENTER.x+VAPS.x, data = DATA)
28. summary(model_VAPS)
29. anova(model_VAPS)
30.
31. model_CRP=lm(CRP.y~TRT.x+CENTER.x+CRP.x, data = DATA)
32. summary(model_CRP)
33. anova(model_CRP)
34.
35. model_HAQ=lm(HAQ.y~TRT.x+CENTER.x+HAQ.x, data = DATA)
36. summary(model_HAQ)
37. anova(model_HAQ)
38.
39. model_PAINJT2=lm(PAINJT.y~TRT.x+CENTER.x+PAINJT.x+TRT.x:CENTER.x+TRT.x:PAINJT.x, data = DATA)
40. summary(model_PAINJT2)
41. anova(model_PAINJT2)
42.
43. model_SWELLJT2=lm(SWELLJT.y~TRT.x+CENTER.x+SWELLJT.x+TRT.x:CENTER.x+TRT.x:SWELLJT.x, data = DATA)
44. summary(model_SWELLJT2)
45. anova(model_SWELLJT2)
46.
47. model_VAPS2=lm(VAPS.y~TRT.x+CENTER.x+VAPS.x+TRT.x:CENTER.x+TRT.x:VAPS.x, data = DATA)
48. summary(model_VAPS2)
49. anova(model_VAPS2)
50.
51. model_CRP2=lm(CRP.y~TRT.x+CENTER.x+CRP.x+TRT.x:CENTER.x+TRT.x:CRP.x, data = DATA)
52. summary(model_CRP2)
53. anova(model_CRP2)
54.
55. model_HAQ2=lm(HAQ.y~TRT.x+CENTER.x+HAQ.x+TRT.x:CENTER.x+TRT.x:HAQ.x, data = DATA)
56. summary(model_HAQ2)
57. anova(model_HAQ2)
58.
59. anova(model_PAINJT, model_PAINJT2)
60. anova(model_SWELLJT, model_SWELLJT2)
61. anova(model_VAPS, model_VAPS2)
62. anova(model_CRP, model_CRP2)
63. anova(model_HAQ, model_HAQ2)

```

the SAS code used for this project is

```
1. data dta01;
2. input CENTER PATIENT DAY VISIT VAPS CRP FINAL TRT PHYASMT PATASMT PAINJT SWELLJT HA
Q ;
3. cards ;
4.      1      55      -6      -1      22      7.10      1      1      5      4      38      31      .
5.      1      55      -1      0      27      7.16      1      1      5      4      38      29      2.31530
6. ;
7.
8. data dta11; set dta01; by center patient day ; retain b1-b7 e1-e7 ;
9.      array b b1-b7 ; array e e1-e7 ; array d d1-d7 ;
10.      array v phyasmt patasmt painjt swelljt vaps haq crp ;
11.      if first.patient then do i=1 to 7; b{i}=.; e{i}=.; d{i}=.; end ;
12.      if day<2 then do i=1 to 7; if v{i} ne . then b{i}=v{i}; end ;
13.      if day>1 then do i=1 to 7; if v{i} ne . then e{i}=v{i}; end ;
14.      do i=1 to 7; d{i}=e{i}-b{i}; end ;
15.      a1=0; a2=0; a3=0; a4=0; a5=0; a6=0; a7=0; acr=0 ;
16.      if .<d1<0 then a1=1; if .<d2<0 then a2=1 ;
17.      if .<d3/b3<=-.2 then a3=1; if .<d4/b4<=-.2 then a4=1 ;
18.      if .<d5/b5<=-.2 then a5=1; if .<d6/b6<=-.2 then a6=1; if .<d7/b7<=-
.2 then a7=1 ;
19.      a8=a1+a2+a5+a6+a7; if a3=1 and a4=1 and a8>2 then acr=1 ;
20.      if d1=. or d2=. or d3=. or d4=. or d5=. or d6=. or d7=. then delete ;
21.      if last.patient then output ;
22.
23.  proc means data=dta11 noprint; var b1-b7; output out=dta12 mean=bas1-bas7 ;
24.  data dta13; merge dta11 dta12 (in=g) ;
25.      retain bx1-bx7; array bas bas1-bas7; array bx bx1-bx7; array b b1-
b7; array nb nb1-nb7 ;
26.      if g then do i=1 to 7; bx{i}=bas{i}; end ;
27.      do i=1 to 7; nb{i}=b{i}-bx{i}; end ;
28.      dcloe=0; if final=2 then dcloe=1 ;
29.
30.  proc means data=dta13 n mean std; class trt; var b1-b7 d1-d7 ;
31.  run ;
32.
33.  title "Analysis of time to discontinuation for lack of efficacy" ;
34.  proc lifetest data=dta13 ;* outsurv=test notable ;
35.      time day*dcloe(0); id patient ; strata trt ;
36.
37.  run ;
38.
39.  /* CMH Test */
40.  proc freq data=dta11;
```

```
41.      tables TRT*acr / cmh;
42.  run;
43.
44.  proc freq data=dta11;
45.      tables CENTER*TRT*acr / cmh;
46.  run;
```