

# Biochemistry

Nathan Constantine-Cooke

2025-08-13

## Table of contents

Introduction . . . . .	3
C-reactive protein . . . . .	3
Crohn's disease . . . . .	3
Patient-reported flare . . . . .	3
Proportional hazards assumption test . . . . .	6
DF betas . . . . .	6
Hard flare . . . . .	8
Proportional hazards assumption test . . . . .	11
DF betas . . . . .	11
Ulcerative colitis . . . . .	13
Patient-reported flare . . . . .	13
Proportional hazards assumption test . . . . .	16
DF betas . . . . .	16
Hard flare . . . . .	18
Proportional hazards assumption test . . . . .	21
DF betas . . . . .	21
Haemoglobin . . . . .	23
Crohn's disease . . . . .	23
Patient-reported flare . . . . .	23
Proportional hazards assumption test . . . . .	26
DF betas . . . . .	26
Hard flare . . . . .	28
Proportional hazards assumption test . . . . .	31
DF betas . . . . .	31
Ulcerative colitis . . . . .	33
Patient-reported flare . . . . .	33
Proportional hazards assumption test . . . . .	36
DF betas . . . . .	36

Hard flare . . . . .	38
Proportional hazards assumption test . . . . .	41
DF betas . . . . .	41
White cell count . . . . .	43
Crohn's disease . . . . .	43
Patient-reported flare . . . . .	43
Proportional hazards assumption test . . . . .	46
DF betas . . . . .	46
Hard flare . . . . .	48
Proportional hazards assumption test . . . . .	51
DF betas . . . . .	51
Ulcerative colitis . . . . .	53
Patient-reported flare . . . . .	53
Proportional hazards assumption test . . . . .	56
DF betas . . . . .	56
Hard flare . . . . .	58
Proportional hazards assumption test . . . . .	61
DF betas . . . . .	61
Platelets . . . . .	63
Crohn's disease . . . . .	64
Patient-reported flare . . . . .	64
Proportional hazards assumption test . . . . .	66
DF betas . . . . .	66
Hard flare . . . . .	68
Proportional hazards assumption test . . . . .	71
DF betas . . . . .	71
Ulcerative colitis . . . . .	73
Patient-reported flare . . . . .	73
Proportional hazards assumption test . . . . .	76
DF betas . . . . .	76
Hard flare . . . . .	78
Proportional hazards assumption test . . . . .	81
DF betas . . . . .	81
Albumin . . . . .	83
Crohn's disease . . . . .	83
Patient-reported flare . . . . .	83
Proportional hazards assumption test . . . . .	86
DF betas . . . . .	86
Hard flare . . . . .	88
Proportional hazards assumption test . . . . .	91
DF betas . . . . .	91
Ulcerative colitis . . . . .	93
Patient-reported flare . . . . .	93

Proportional hazards assumption test	96
DF betas	96
Hard flare	98
Proportional hazards assumption test	101
DF betas	101
Reproduction and reproducibility	103

## Introduction

```

source("Survival/utils.R")

# Setup analysis environment
analysis_setup <- setup_analysis()
paths <- analysis_setup$paths
demo <- analysis_setup$demo

flare.df <- readRDS(paste0(paths$outdir, "flares-IBD.RDS"))
flare.cd.df <- readRDS(paste0(paths$outdir, "flares-IBD-cd.RDS"))
flare.uc.df <- readRDS(paste0(paths$outdir, "flares-IBD-uc.RDS"))

cd.clin.forest <- readRDS(paste0(paths$outdir, "cd-clin-IBD.RDS"))
cd.hard.forest <- readRDS(paste0(paths$outdir, "cd-hard-IBD.RDS"))
uc.clin.forest <- readRDS(paste0(paths$outdir, "uc-clin-IBD.RDS"))
uc.hard.forest <- readRDS(paste0(paths$outdir, "uc-hard-IBD.RDS"))

```

## C-reactive protein

### Crohn's disease

#### Patient-reported flare

```

# Categorize CRP
flare.cd.df <- categorize_variable(flare.cd.df, "CReactiveProtein",
                                      breaks = c(0, 5, 10.0000001, Inf),
                                      labels = c("CRP < 5", "5 < CRP < 10", "CRP > 10"))

# Run survival analysis using utility function
analysis_result <- run_survival_analysis(
  data = flare.cd.df,
  var_name = "CReactiveProtein",

```

```

    outcome_time = "softflare_time",
    outcome_event = "softflare",
    legend_title = "CRP",
    plot_base_path = "plots/cd/soft-flare/biochem/crp",
    break_time_by = 200
)

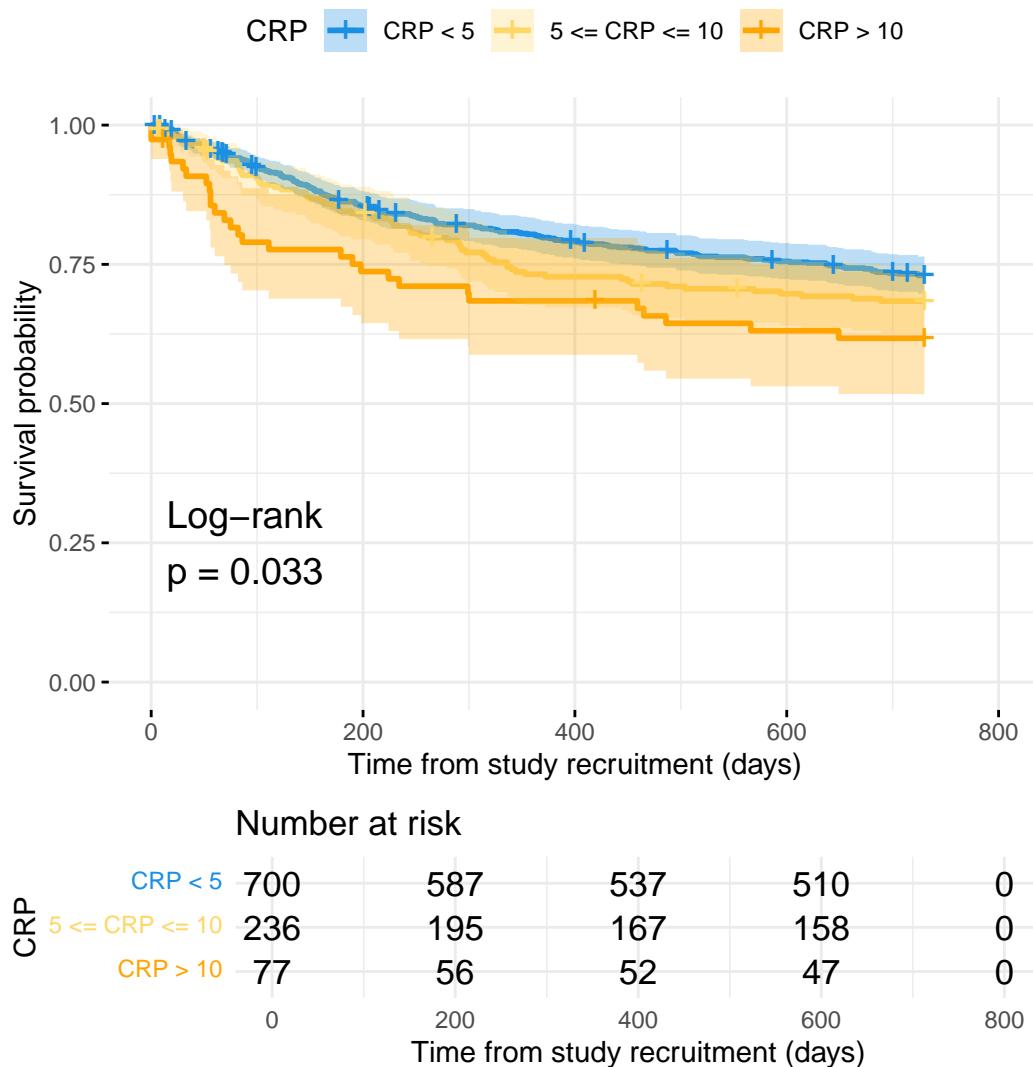
# Extract hazard ratio for continuous CRP variable
fit.me <- coxph(
  Surv(softflare_time, softflare) ~
    Sex + IMD + cat + CReactiveProtein + frailty(SiteNo),
  control = coxph.control(outer.max = 20),
  data = flare.cd.df
)

cd.clin.forest <- rbind(
  cd.clin.forest,
  get_HR(fit.me, "CReactiveProtein")
)

# Display plot and model summary
analysis_result$plot

```

### Time to patient-reported flare



```
invisible(cox_summary(fit.me))
```

Cox model summary:

Variable	HR	Lower 95%	Upper 95%	P-value
SexFemale	2.5114	1.9050	3.3108	0.0000
IMD2	1.2280	0.7064	2.1346	0.4666
IMD3	1.0639	0.6055	1.8696	0.8294

Variable	HR	Lower 95%	Upper 95%	P-value
IMD4	1.2974	0.7536	2.2335	0.3476
IMD5	1.1735	0.6914	1.9917	0.5535
catFC 50-250	1.5652	1.1720	2.0903	0.0024
catFC > 250	2.4876	1.8060	3.4265	0.0000
CReactiveProtein	1.0017	0.9894	1.0141	0.7891

Diagnostics:

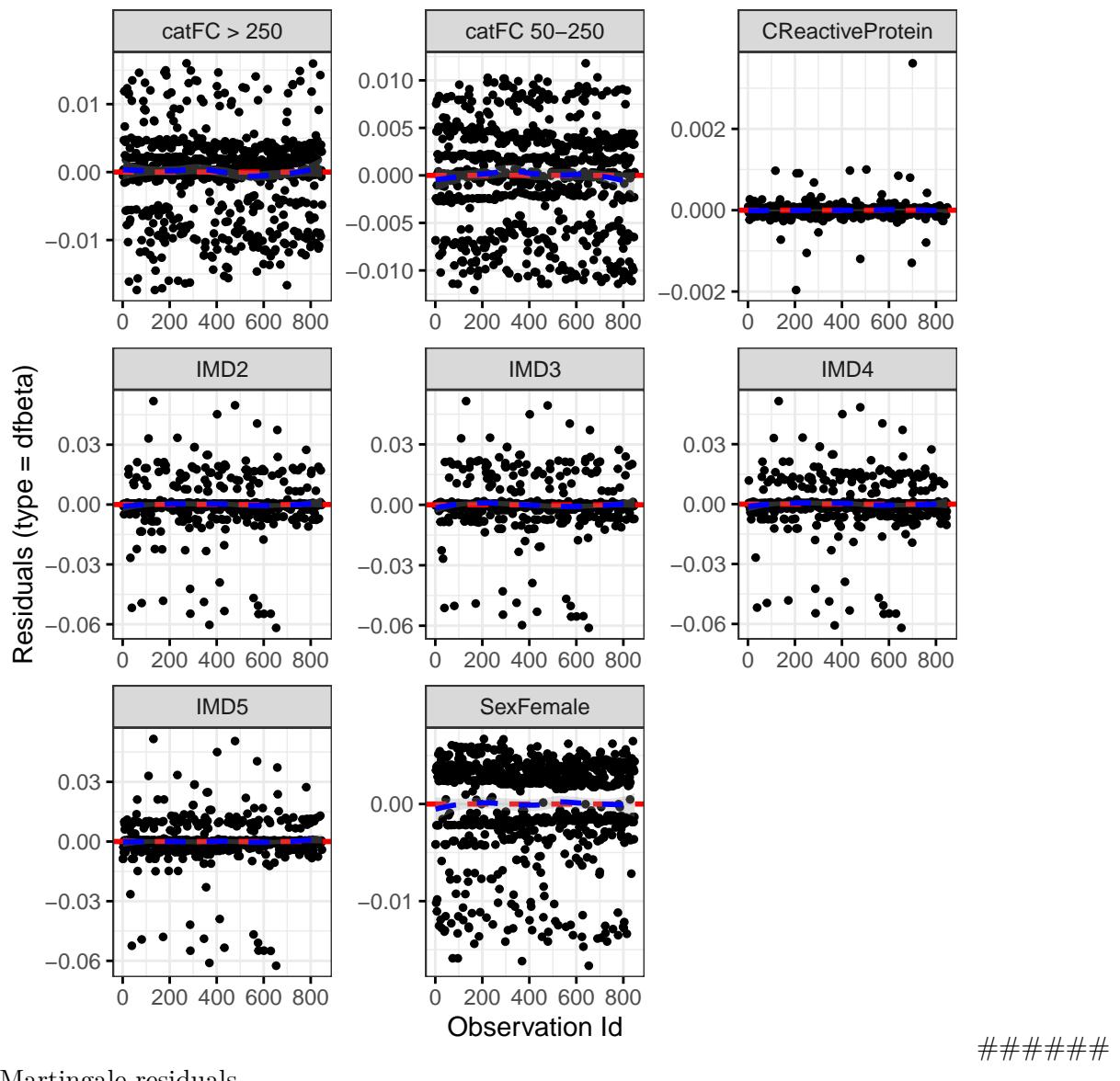
### Proportional hazards assumption test

	Chi-squared statistic	DF	P-value
Sex	0.1347	0.9918	0.7104
IMD	1.5655	3.9514	0.8093
cat	3.5876	1.9800	0.1637
CReactiveProtein	2.2628	0.9912	0.1310
GLOBAL	6.8886	14.0404	0.9400

### DF betas

```
Warning: `gather_()` was deprecated in tidyverse 1.2.0.
  i Please use `gather()` instead.
  i The deprecated feature was likely used in the survminer package.
    Please report the issue at <https://github.com/kassambara/survminer/issues>.
```

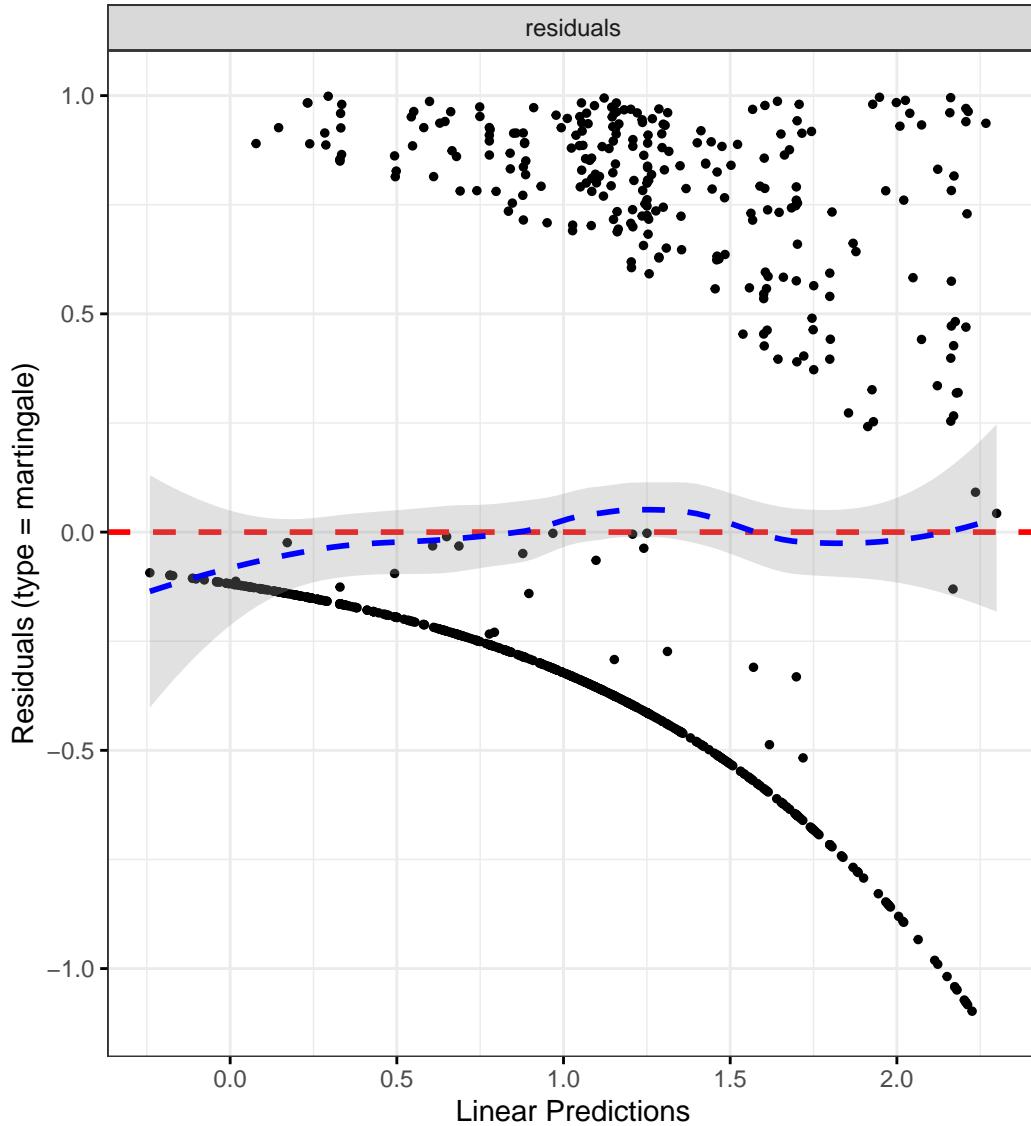
```
`geom_smooth()` using formula = 'y ~ x'
```



Martingale residuals

#####

```
`geom_smooth()` using formula = 'y ~ x'
```



## Hard flare

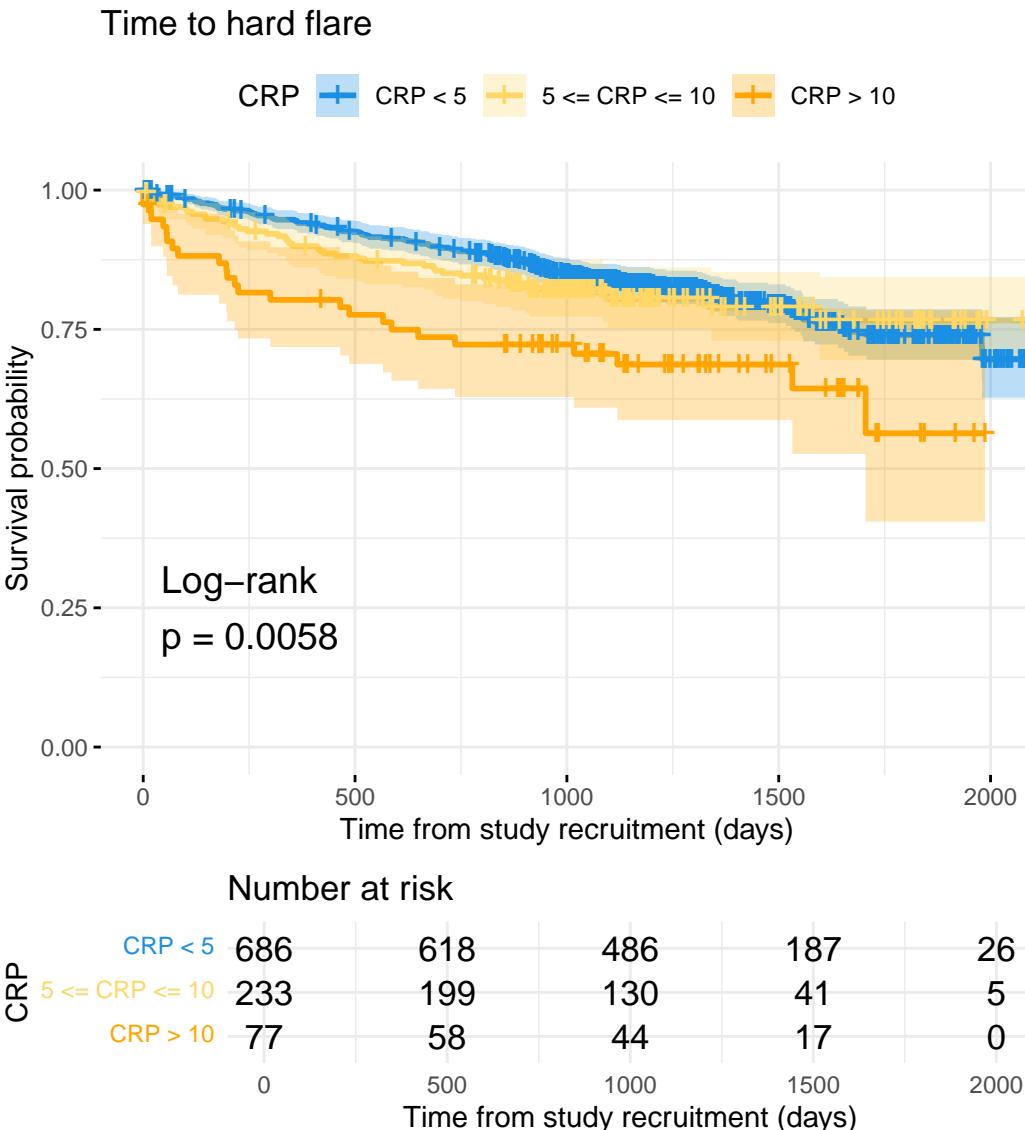
```
# Generate survival plot and run Cox model for hard flare using utility function
analysis_result <- run_survival_analysis(
  data = flare.cd.df,
  var_name = "CReactiveProtein",
  outcome_time = "hardflare_time",
  outcome_event = "hardflare",
  legend_title = "CRP",
```

```
plot_base_path = "plots/cd/hard-flare/biochem/crp",
break_time_by = 500
)

# Extract hazard ratio for continuous CRP variable
fit.me <- coxph(
  Surv(hardflare_time, hardflare) ~
    Sex + IMD + cat + CReactiveProtein + frailty(SiteNo),
  control = coxph.control(outer.max = 20),
  data = flare.cd.df
)

cd.hard.forest <- rbind(
  cd.hard.forest,
  get_HR(fit.me, "CReactiveProtein")
)

# Display plot and model summary
analysis_result$plot
```



```
invisible(cox_summary(fit.me))
```

Cox model summary:

Variable	HR	Lower 95%	Upper 95%	P-value
SexFemale	1.5403	1.1285	2.1022	0.0065
IMD2	1.1741	0.5904	2.3348	0.6472
IMD3	1.1512	0.5724	2.3152	0.6929

Variable	HR	Lower 95%	Upper 95%	P-value
IMD4	1.2896	0.6580	2.5274	0.4587
IMD5	1.2097	0.6295	2.3247	0.5678
catFC 50-250	2.0009	1.3908	2.8786	0.0002
catFC > 250	3.0710	2.0784	4.5375	0.0000
CReactiveProtein	1.0088	0.9995	1.0181	0.0635

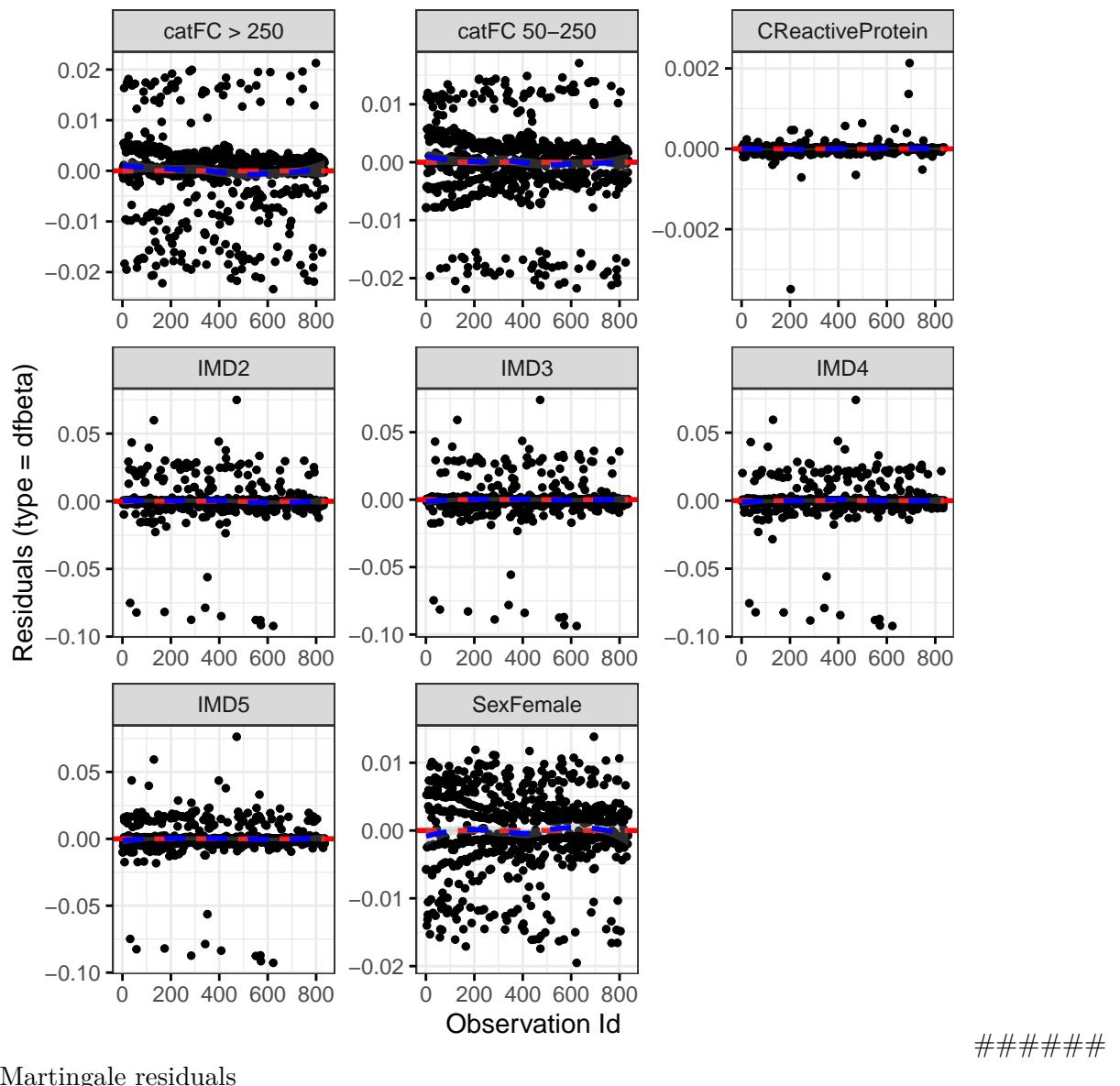
Diagnostics:

### Proportional hazards assumption test

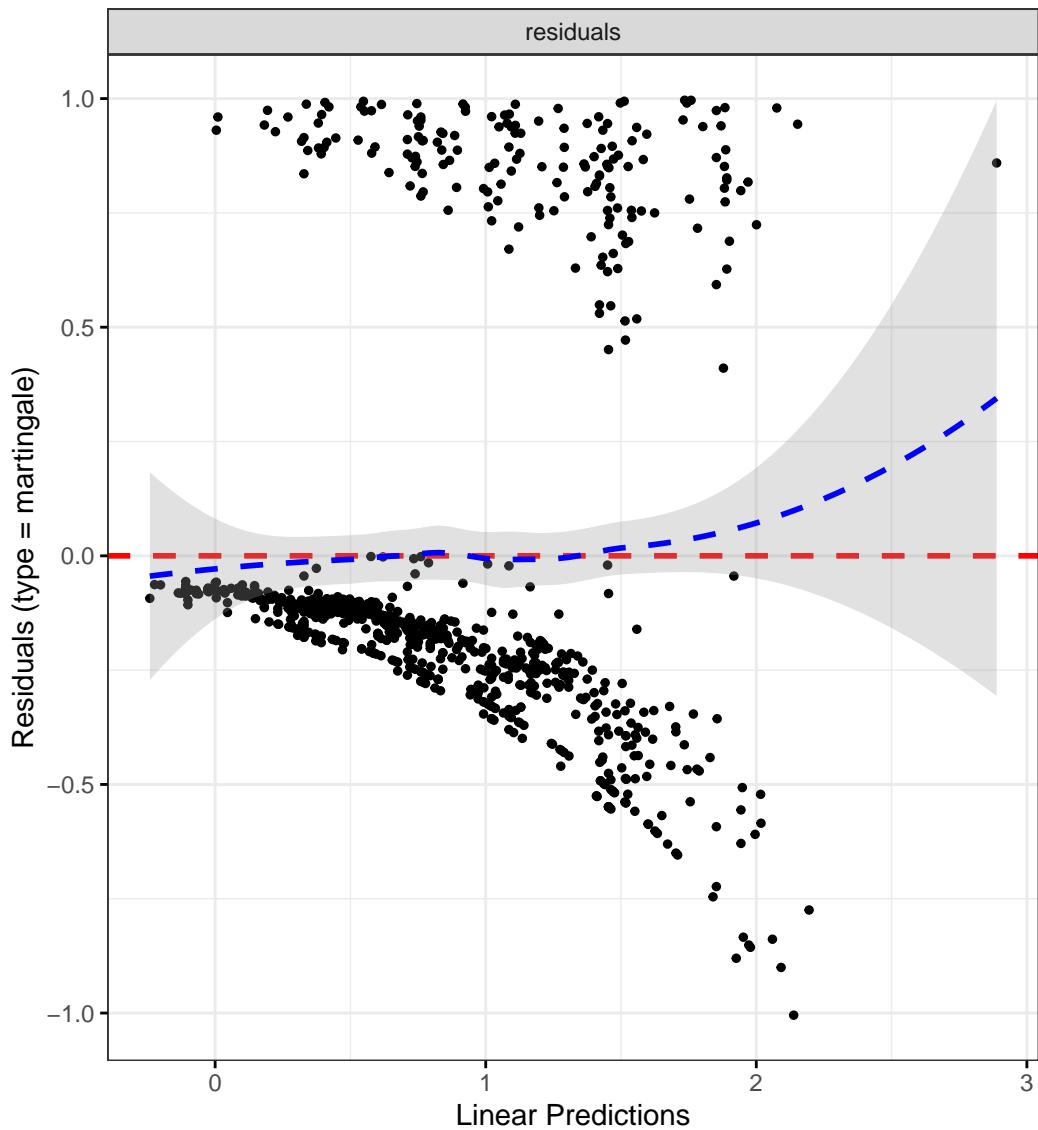
	Chi-squared statistic	DF	P-value
Sex	3.3489	0.9880	0.0661
IMD	2.8908	3.9532	0.5689
cat	6.8059	1.9875	0.0329
CReactiveProtein	2.7995	0.9941	0.0935
GLOBAL	15.7224	14.4155	0.3594

### DF betas

```
`geom_smooth()` using formula = 'y ~ x'
```



```
`geom_smooth()` using formula = 'y ~ x'
```



### Ulcerative colitis

#### Patient-reported flare

```
# Categorize CRP
flare.uc.df <- categorize_variable(flare.uc.df, "CReactiveProtein",
                                      breaks = c(0, 5, 10.000001, Inf),
                                      labels = c("CRP < 5", "5 < CRP < 10", "CRP > 10"))
```

```

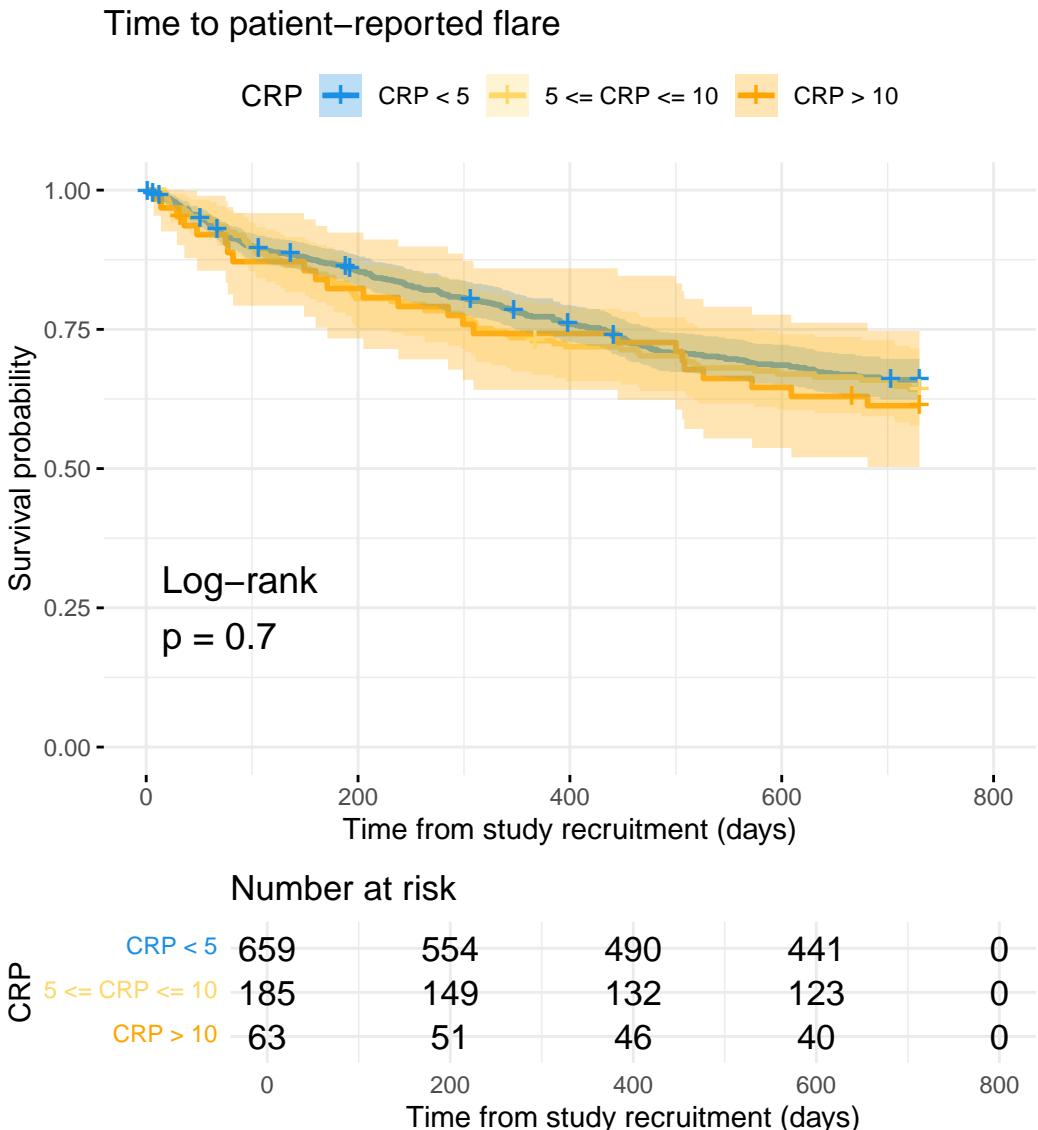
# Run survival analysis using utility function
analysis_result <- run_survival_analysis(
  data = flare.uc.df,
  var_name = "CReactiveProtein",
  outcome_time = "softflare_time",
  outcome_event = "softflare",
  legend_title = "CRP",
  plot_base_path = "plots/uc/soft-flare/biochem/crp",
  break_time_by = 200
)

# Extract hazard ratio for continuous CRP variable
fit.me <- coxph(
  Surv(softflare_time, softflare) ~
    Sex + IMD + cat + CReactiveProtein + frailty(SiteNo),
  control = coxph.control(outer.max = 20),
  data = flare.uc.df
)

uc.clin.forest <- rbind(
  uc.clin.forest,
  get_HR(fit.me, "CReactiveProtein")
)

# Display plot and model summary
analysis_result$plot

```



```
invisible(cox_summary(fit.me))
```

Cox model summary:

Variable	HR	Lower 95%	Upper 95%	P-value
SexFemale	1.6446	1.2924	2.0926	0.0001
IMD2	1.2409	0.7406	2.0790	0.4124
IMD3	1.0484	0.6261	1.7557	0.8573

Variable	HR	Lower 95%	Upper 95%	P-value
IMD4	1.4396	0.8833	2.3463	0.1437
IMD5	1.1606	0.7198	1.8712	0.5412
catFC 50-250	1.5937	1.2038	2.1098	0.0011
catFC > 250	2.2826	1.6943	3.0753	0.0000
CReactiveProtein	1.0018	0.9860	1.0179	0.8258

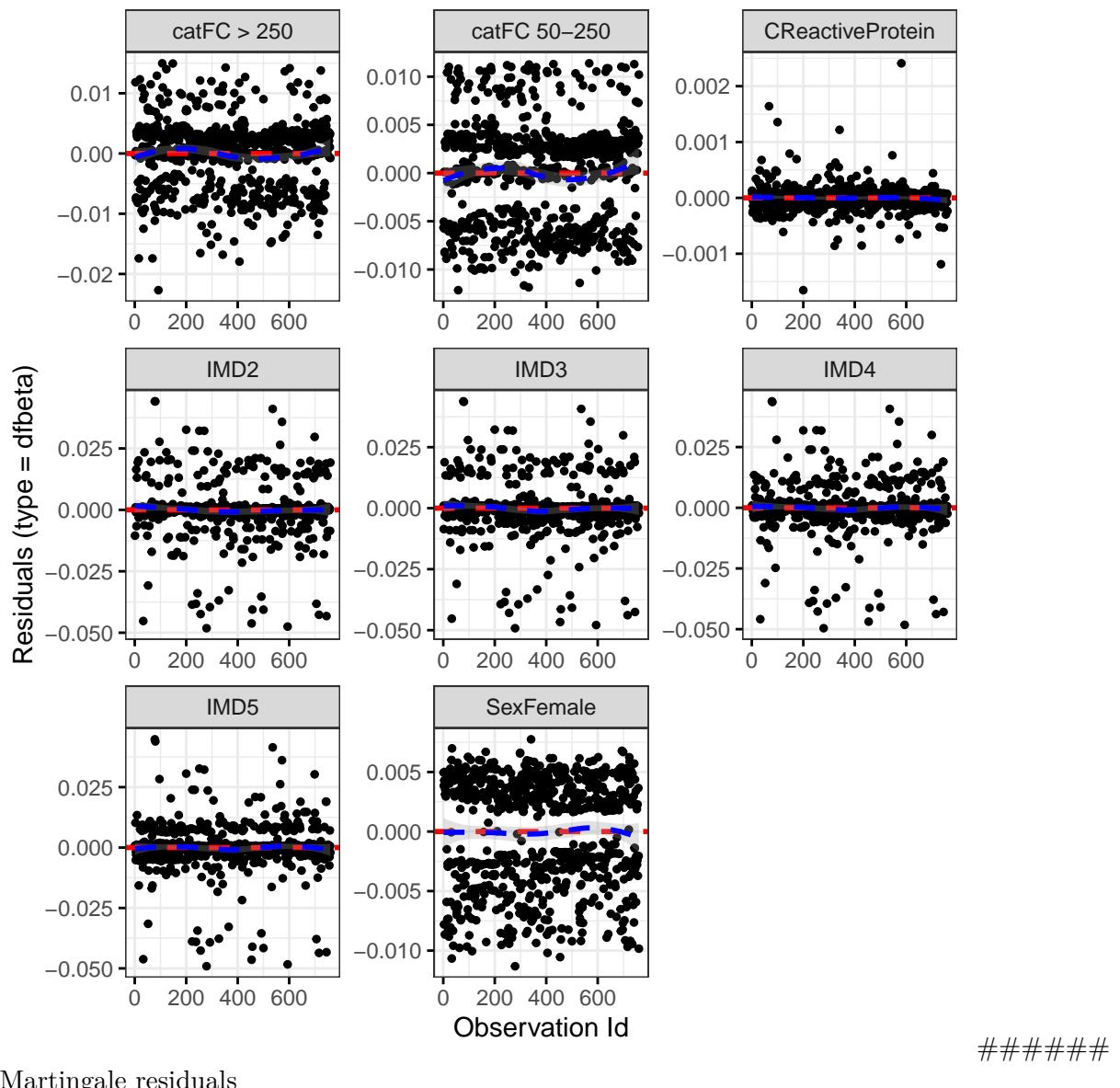
Diagnostics:

### Proportional hazards assumption test

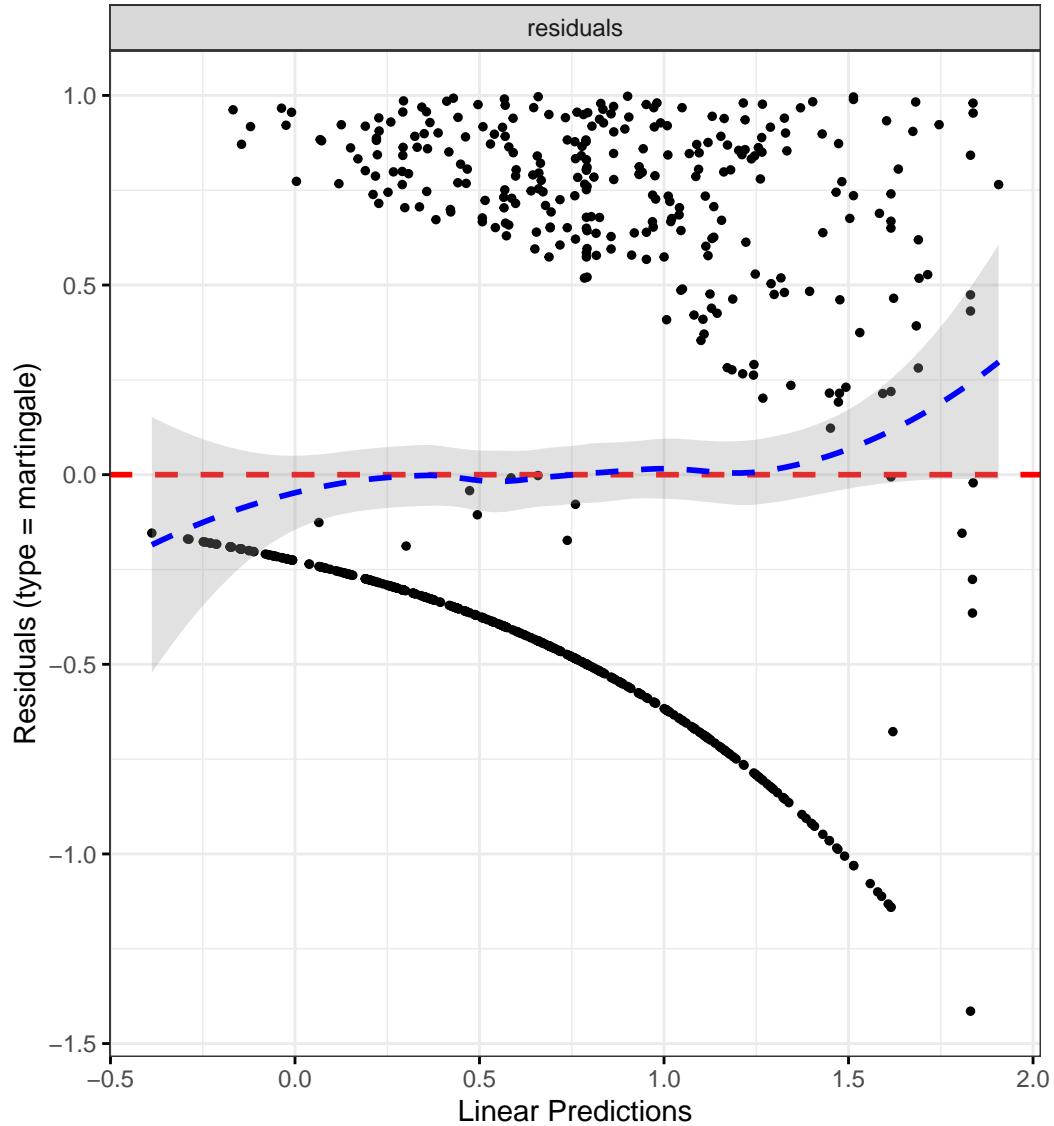
	Chi-squared statistic	DF	P-value
Sex	2.0669	0.9902	0.1487
IMD	2.5721	3.9299	0.6209
cat	4.4034	1.9631	0.1072
CReactiveProtein	4.3050	0.9837	0.0371
GLOBAL	12.9174	19.3120	0.8555

### DF betas

```
`geom_smooth()` using formula = 'y ~ x'
```



```
`geom_smooth()` using formula = 'y ~ x'
```



### Hard flare

```
# Generate survival plot and run Cox model for hard flare using utility function
analysis_result <- run_survival_analysis(
  data = flare.uc.df,
  var_name = "CReactiveProtein",
  outcome_time = "hardflare_time",
  outcome_event = "hardflare",
  legend_title = "CRP",
```

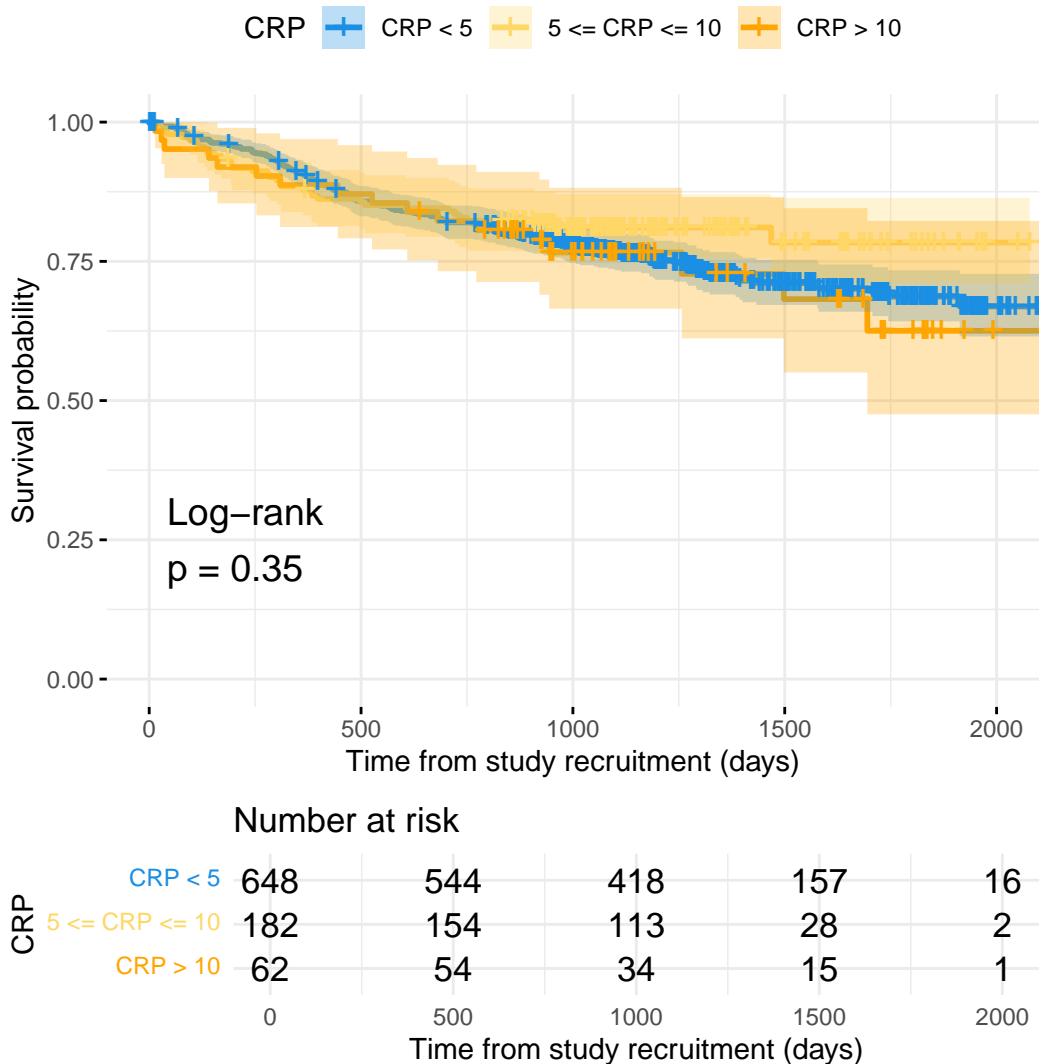
```
plot_base_path = "plots/uc/hard-flare/biochem/crp",
break_time_by = 500
)

# Extract hazard ratio for continuous CRP variable
fit.me <- coxph(
  Surv(hardflare_time, hardflare) ~
    Sex + IMD + cat + CReactiveProtein + frailty(SiteNo),
  control = coxph.control(outer.max = 20),
  data = flare.uc.df
)

uc.hard.forest <- rbind(
  uc.hard.forest,
  get_HR(fit.me, "CReactiveProtein")
)

# Display plot and model summary
analysis_result$plot
```

### Time to hard flare



```
invisible(cox_summary(fit.me))
```

Cox model summary:

Variable	HR	Lower 95%	Upper 95%	P-value
SexFemale	1.2811	0.9631	1.7040	0.0888
IMD2	1.6713	0.8545	3.2688	0.1335
IMD3	1.3587	0.6949	2.6563	0.3702

Variable	HR	Lower 95%	Upper 95%	P-value
IMD4	1.9141	1.0067	3.6396	0.0477
IMD5	1.5133	0.8053	2.8437	0.1981
catFC 50-250	2.0323	1.4449	2.8585	0.0000
catFC > 250	3.0334	2.1208	4.3386	0.0000
CReactiveProtein	0.9953	0.9739	1.0170	0.6666

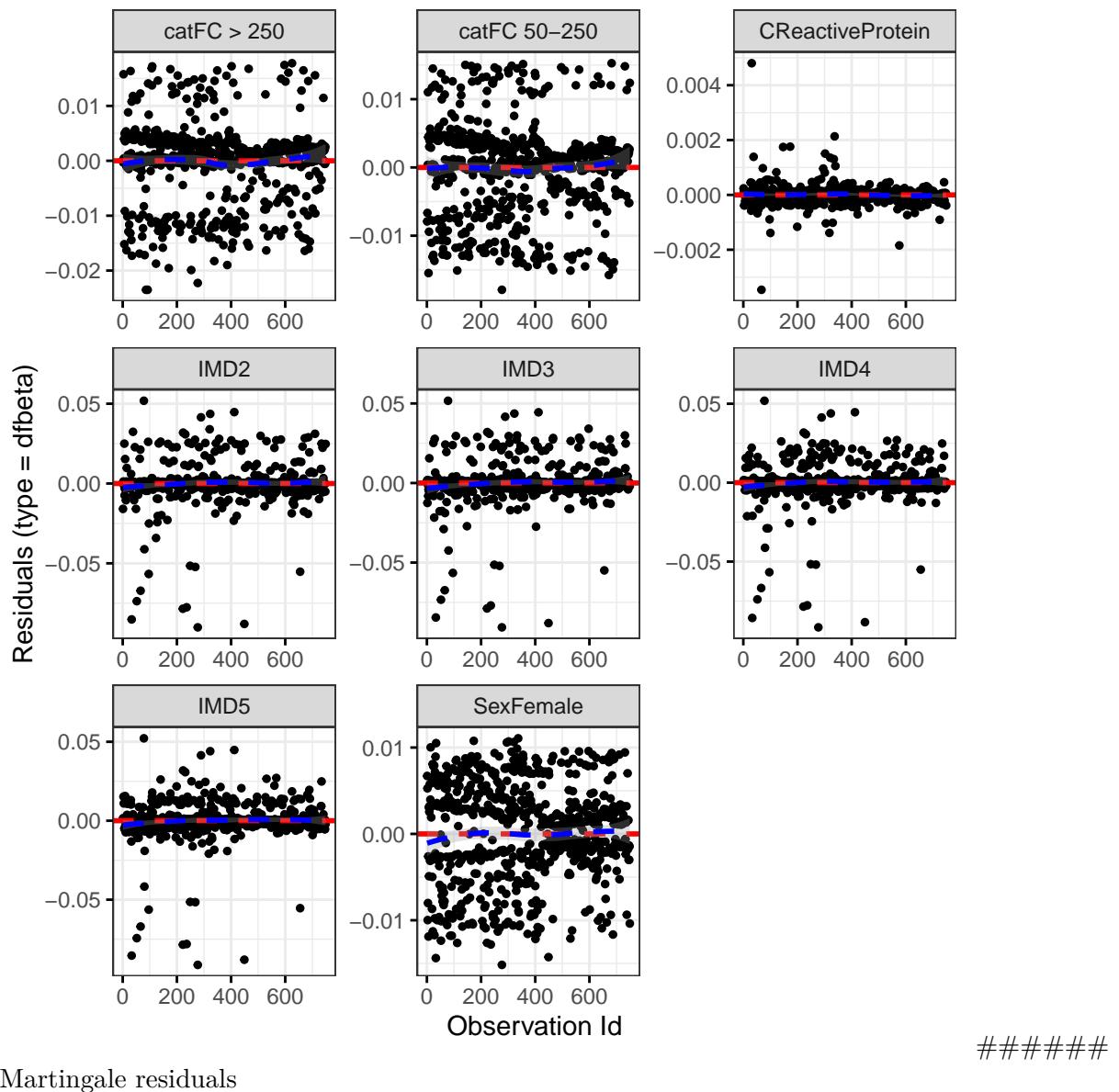
Diagnostics:

### Proportional hazards assumption test

	Chi-squared statistic	DF	P-value
Sex	0.4924	0.9867	0.4775
IMD	0.8766	3.9347	0.9236
cat	3.9346	1.9653	0.1359
CReactiveProtein	0.0770	0.9850	0.7758
GLOBAL	5.4018	21.3499	0.9998

### DF betas

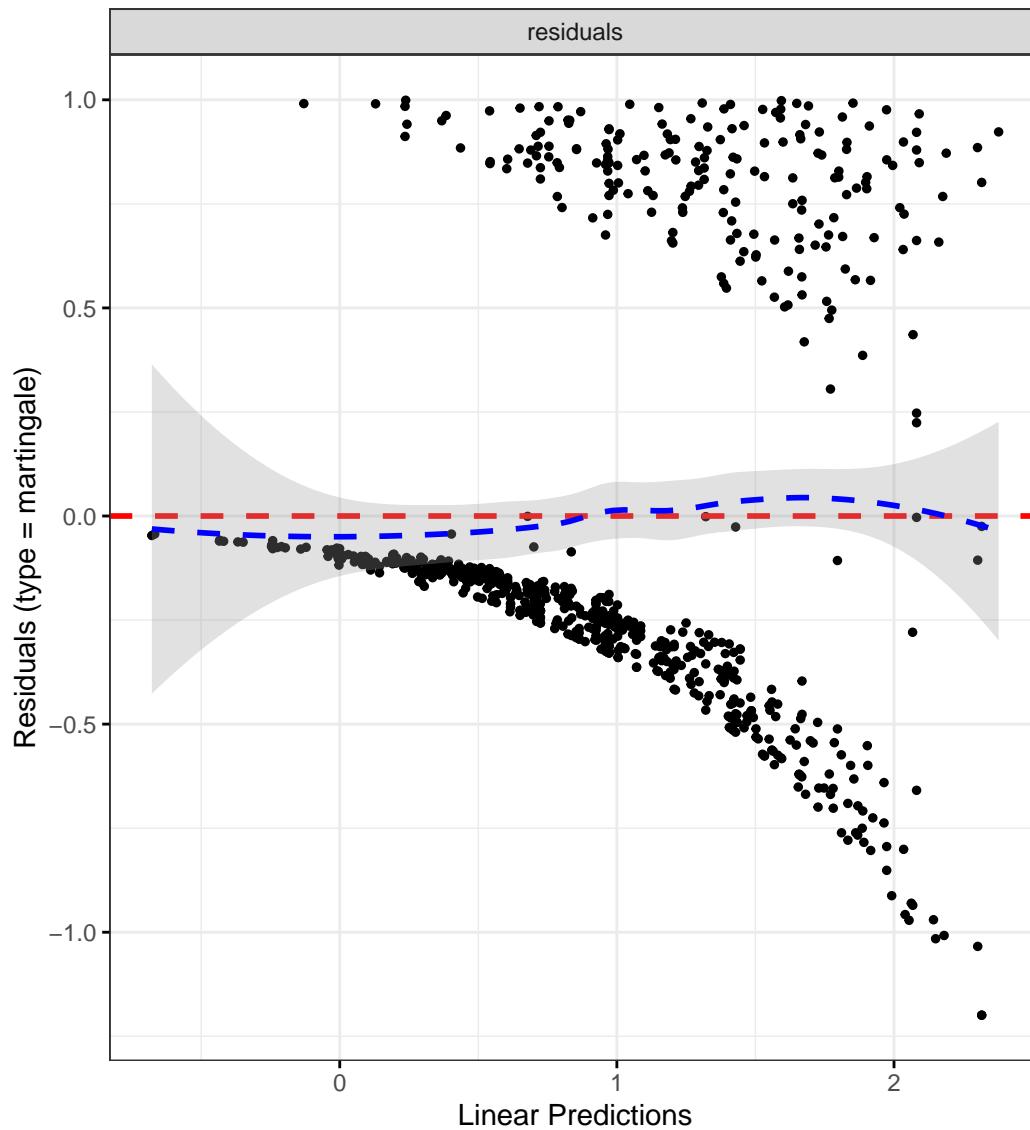
```
`geom_smooth()` using formula = 'y ~ x'
```



#####

Martingale residuals

```
`geom_smooth()` using formula = 'y ~ x'
```



**Haemoglobin**

**Crohn's disease**

**Patient-reported flare**

```
# Categorize Haemoglobin by quantiles
flare.cd.df <- categorize_by_quantiles(flare.cd.df, "Haemoglobin", reference_data = flare.df)
```

```

# Run survival analysis using utility function
analysis_result <- run_survival_analysis(
  data = flare.cd.df,
  var_name = "Haemoglobin",
  outcome_time = "softflare_time",
  outcome_event = "softflare",
  legend_title = "Haemoglobin",
  plot_base_path = "plots/cd/soft-flare/biochem/haemoglobin",
  break_time_by = 200
)

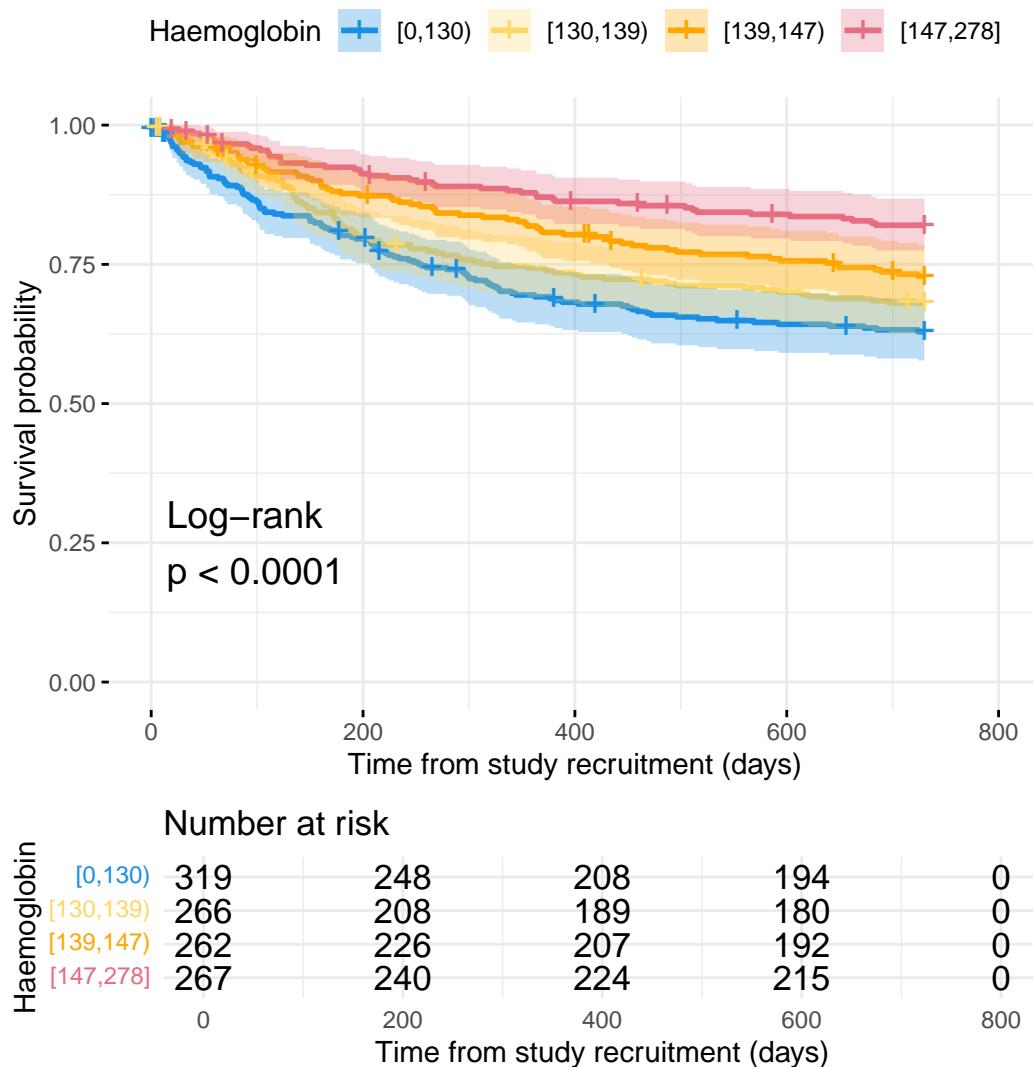
# Extract hazard ratio for continuous Haemoglobin variable
fit.me <- coxph(
  Surv(softflare_time, softflare) ~
    Sex + IMD + cat + Haemoglobin + frailty(SiteNo),
  control = coxph.control(outer.max = 20),
  data = flare.cd.df
)

cd.clin.forest <- rbind(
  cd.clin.forest,
  get_HR(fit.me, "Haemoglobin")
)

# Display plot and model summary
analysis_result$plot

```

### Time to patient-reported flare



```
invisible(cox_summary(fit.me))
```

Cox model summary:

Variable	HR	Lower 95%	Upper 95%	P-value
SexFemale	2.0755	1.5455	2.7872	0.0000
IMD2	1.1381	0.6728	1.9251	0.6296
IMD3	1.0670	0.6274	1.8145	0.8109

Variable	HR	Lower 95%	Upper 95%	P-value
IMD4	1.1227	0.6698	1.8817	0.6605
IMD5	1.0728	0.6506	1.7690	0.7830
catFC 50-250	1.5963	1.2125	2.1017	0.0009
catFC > 250	2.5652	1.8900	3.4818	0.0000
Haemoglobin	0.9925	0.9829	1.0023	0.1330

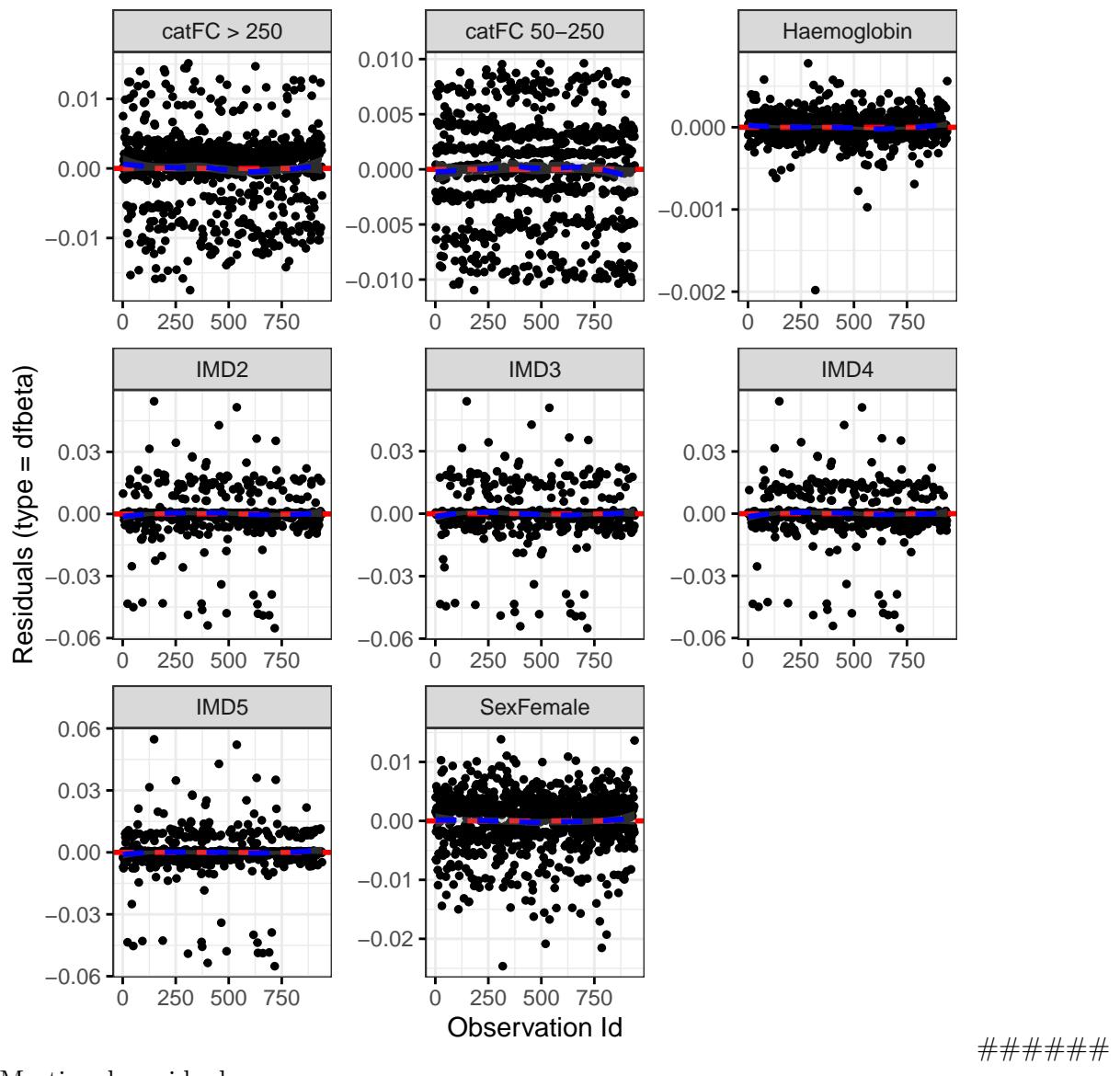
Diagnostics:

### Proportional hazards assumption test

	Chi-squared statistic	DF	P-value
Sex	0.3008	0.9904	0.5794
IMD	4.2401	3.9529	0.3678
cat	4.1442	1.9827	0.1241
Haemoglobin	1.2177	0.9929	0.2677
GLOBAL	9.2130	13.9190	0.8127

### DF betas

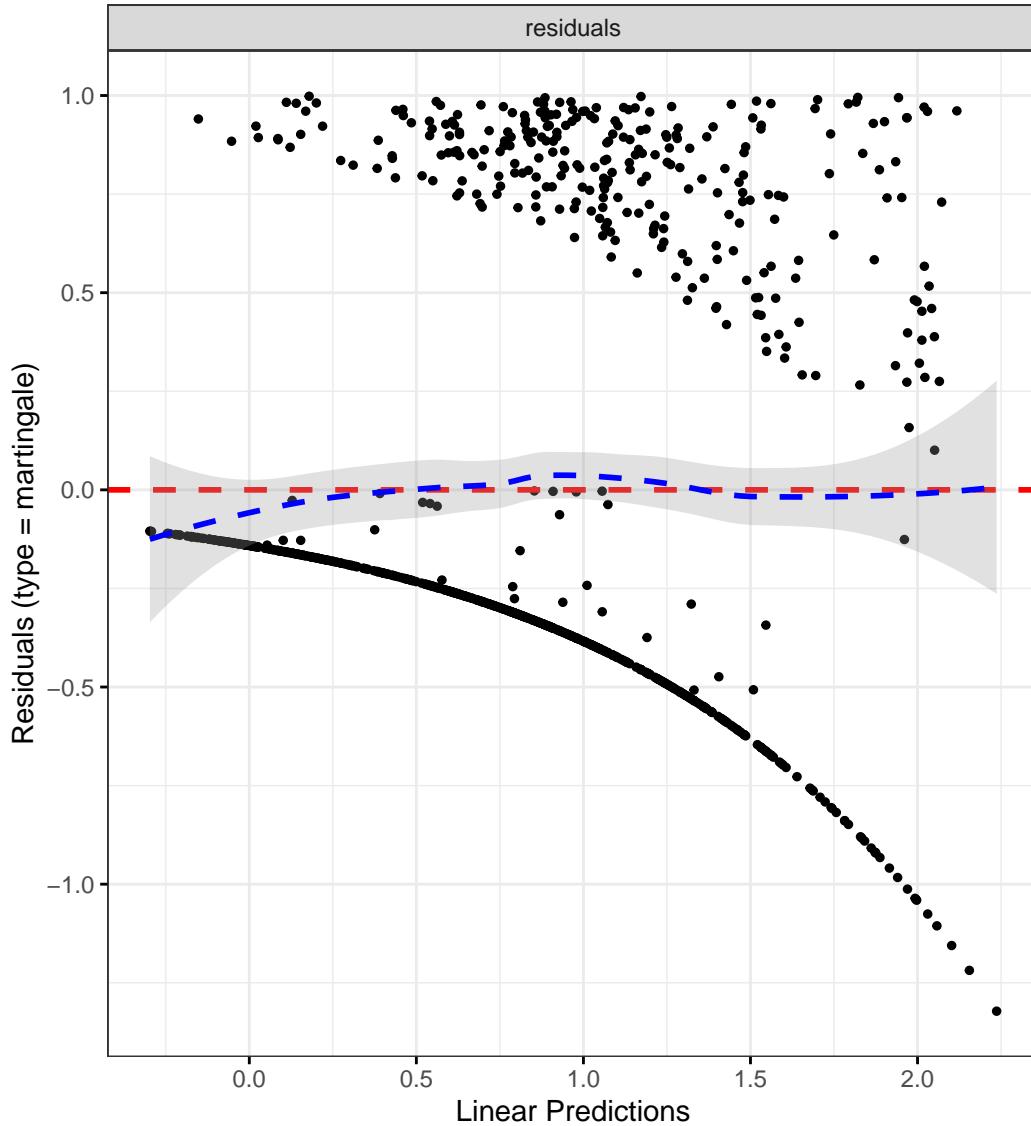
```
`geom_smooth()` using formula = 'y ~ x'
```



Martingale residuals

#####

```
`geom_smooth()` using formula = 'y ~ x'
```



## Hard flare

```
# Generate survival plot and run Cox model for hard flare using utility function
analysis_result <- run_survival_analysis(
  data = flare.cd.df,
  var_name = "Haemoglobin",
  outcome_time = "hardflare_time",
  outcome_event = "hardflare",
  legend_title = "Haemoglobin",
```

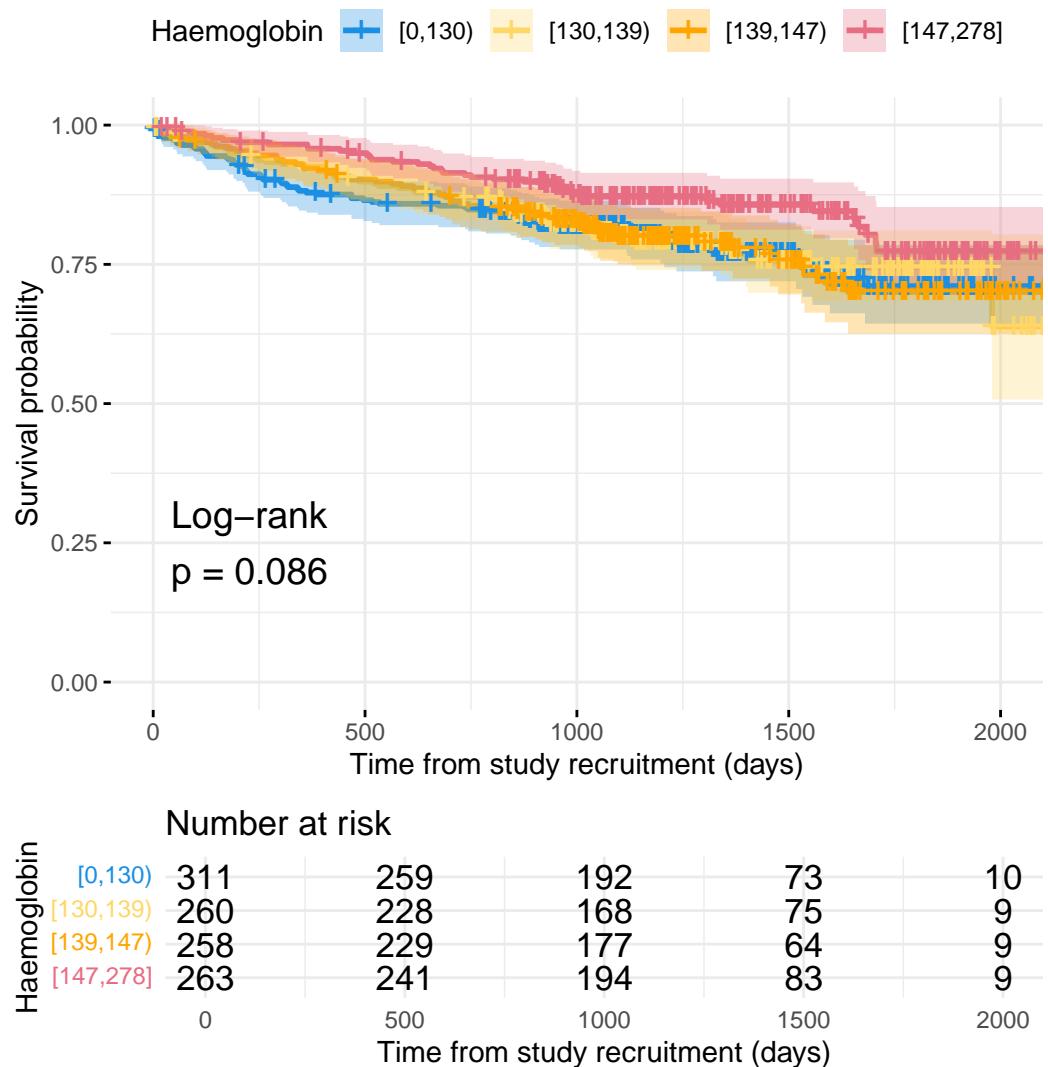
```
plot_base_path = "plots/cd/hard-flare/biochem/haemoglobin",
break_time_by = 500
)

# Extract hazard ratio for continuous Haemoglobin variable
fit.me <- coxph(
  Surv(hardflare_time, hardflare) ~
    Sex + IMD + cat + Haemoglobin + frailty(SiteNo),
  control = coxph.control(outer.max = 20),
  data = flare.cd.df
)

cd.hard.forest <- rbind(
  cd.hard.forest,
  get_HR(fit.me, "Haemoglobin")
)

# Display plot and model summary
analysis_result$plot
```

### Time to hard flare



```
invisible(cox_summary(fit.me))
```

Cox model summary:

Variable	HR	Lower 95%	Upper 95%	P-value
SexFemale	1.3499	0.9565	1.9051	0.0878
IMD2	1.0105	0.5316	1.9207	0.9745
IMD3	1.1379	0.5964	2.1711	0.6951

Variable	HR	Lower 95%	Upper 95%	P-value
IMD4	1.0484	0.5581	1.9694	0.8832
IMD5	0.9917	0.5397	1.8221	0.9785
catFC 50-250	2.0550	1.4486	2.9154	0.0001
catFC > 250	3.5176	2.4149	5.1238	0.0000
Haemoglobin	0.9956	0.9841	1.0072	0.4523

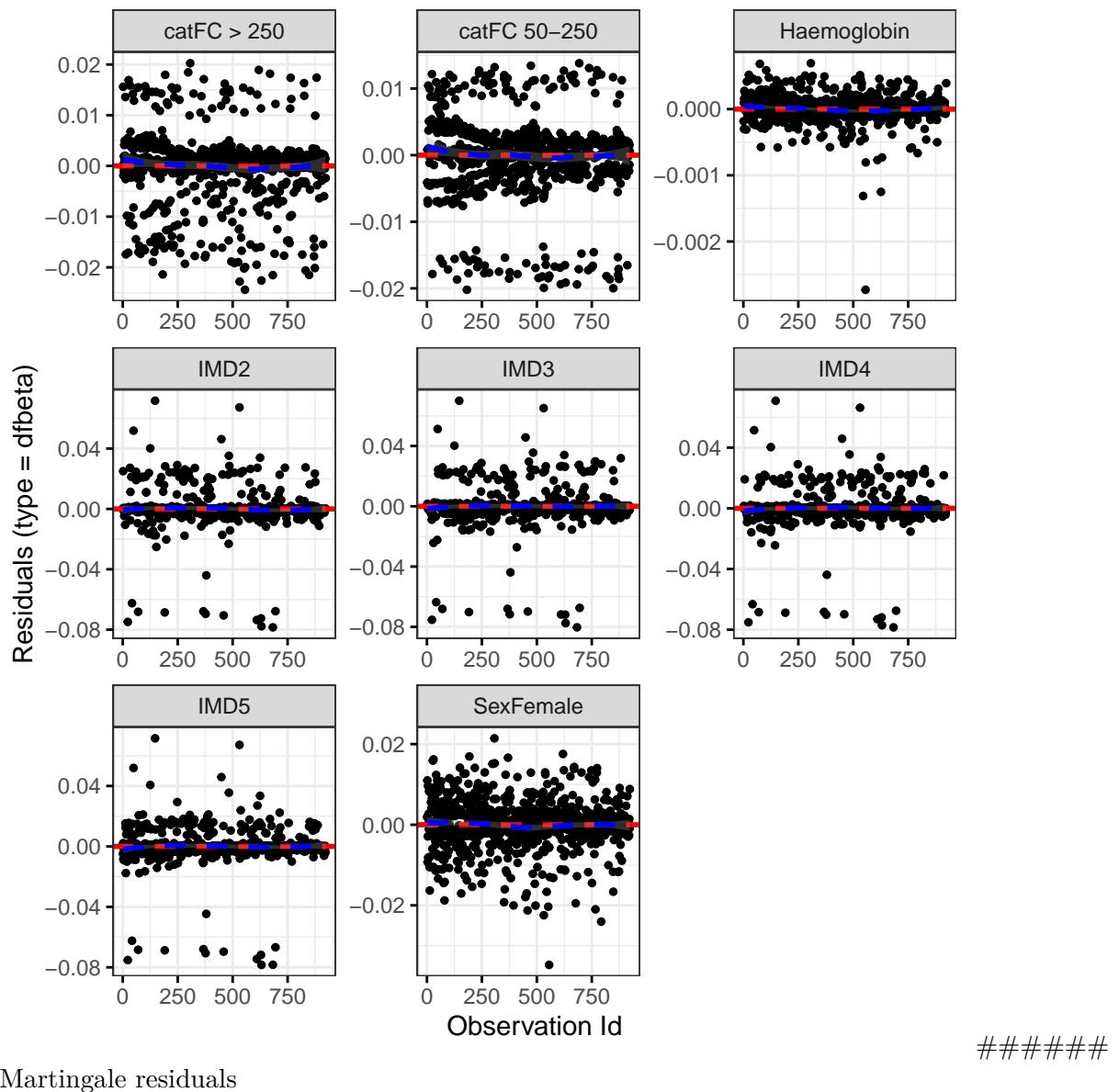
Diagnostics:

### Proportional hazards assumption test

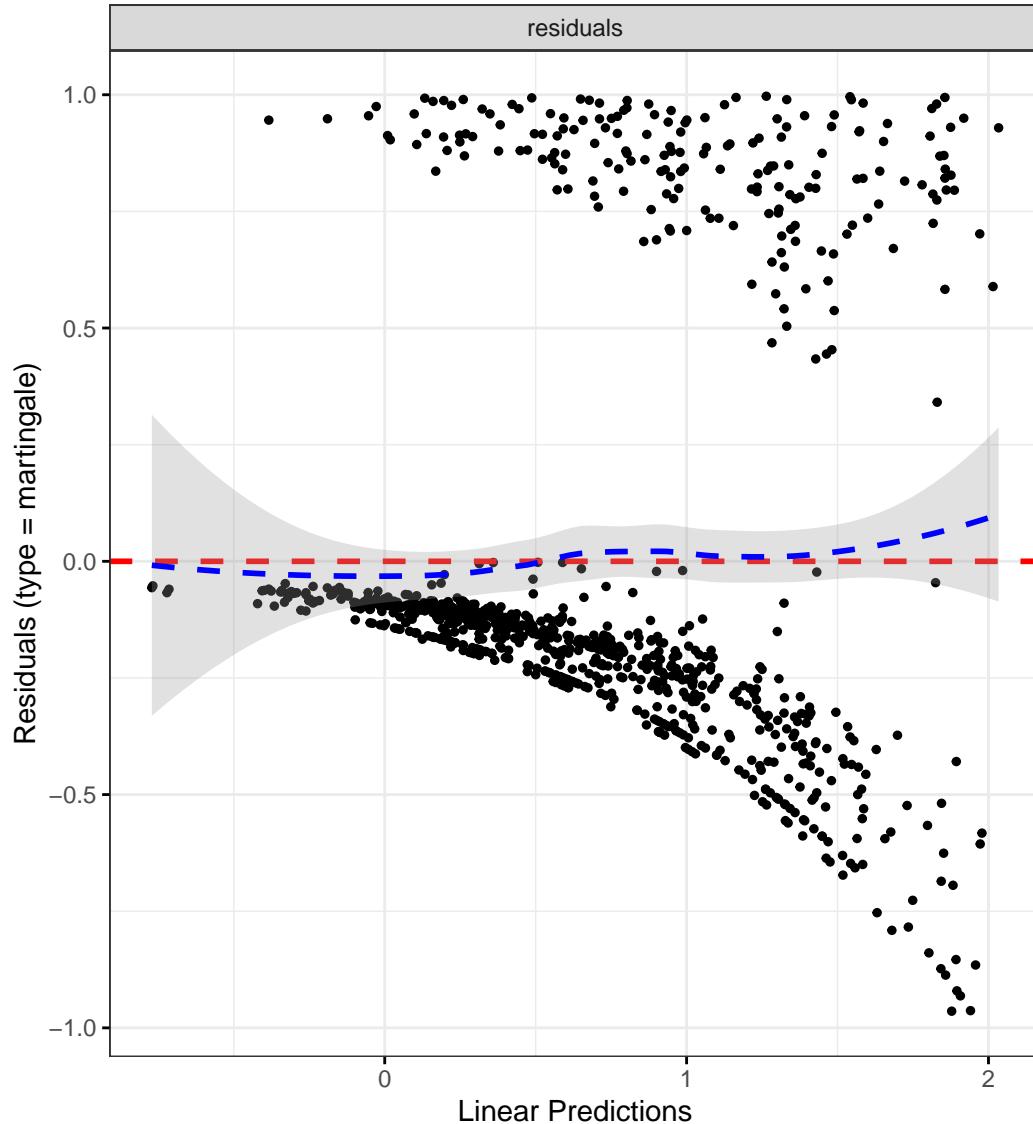
	Chi-squared statistic	DF	P-value
Sex	2.8560	0.9871	0.0894
IMD	3.7488	3.9454	0.4328
cat	9.7005	1.9849	0.0077
Haemoglobin	6.8981	0.9911	0.0085
GLOBAL	20.8165	18.9811	0.3459

### DF betas

```
`geom_smooth()` using formula = 'y ~ x'
```



```
`geom_smooth()` using formula = 'y ~ x'
```



## Ulcerative colitis

### Patient-reported flare

```
# Categorize Haemoglobin by quantiles
flare.uc.df <- categorize_by_quantiles(flare.uc.df, "Haemoglobin", reference_data = flare.df)

# Run survival analysis using utility function
analysis_result <- run_survival_analysis()
```

```

data = flare.uc.df,
var_name = "Haemoglobin",
outcome_time = "softflare_time",
outcome_event = "softflare",
legend_title = "Haemoglobin",
plot_base_path = "plots/uc/soft-flare/biochem/haemoglobin",
break_time_by = 200
)

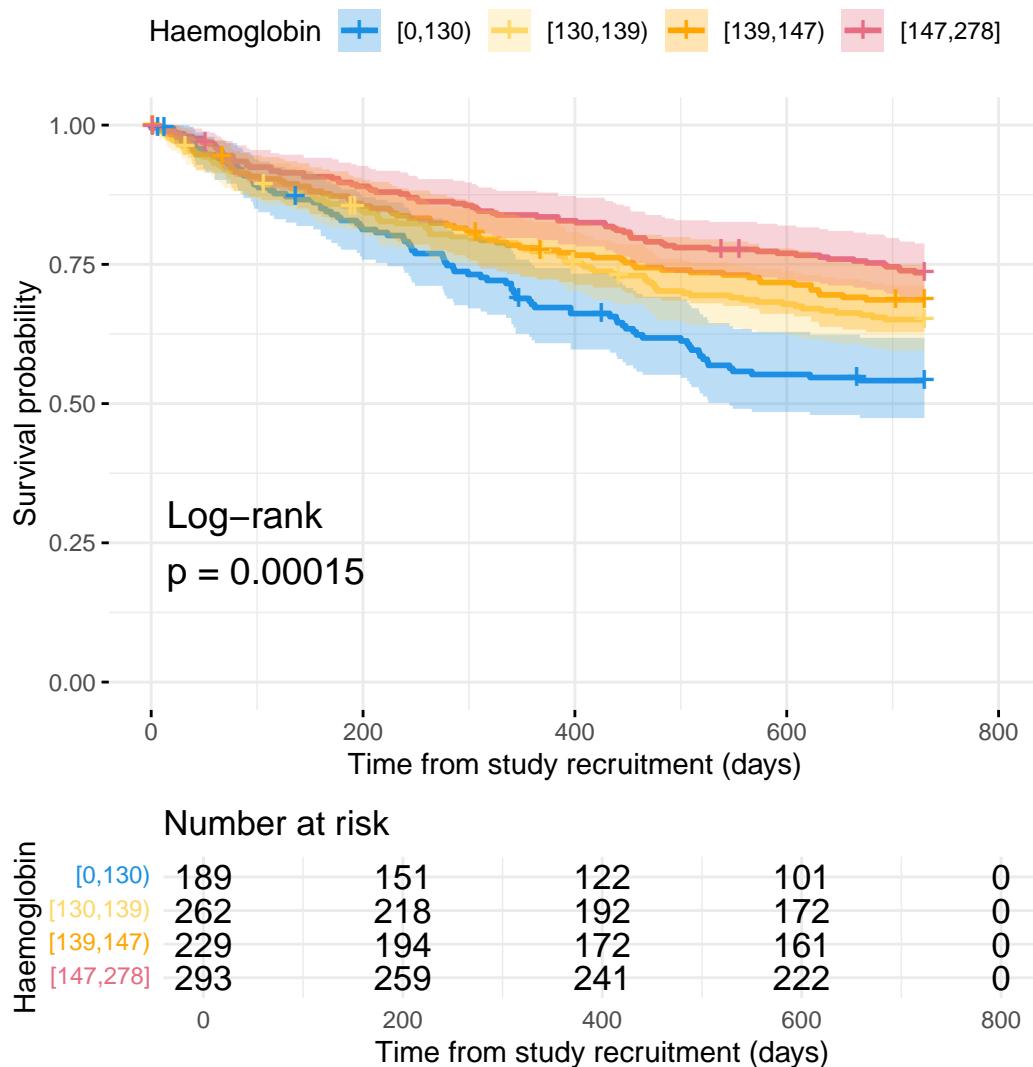
# Extract hazard ratio for continuous Haemoglobin variable
fit.me <- coxph(
  Surv(softflare_time, softflare) ~
    Sex + IMD + cat + Haemoglobin + frailty(SiteNo),
  control = coxph.control(outer.max = 20),
  data = flare.uc.df
)

uc.clin.forest <- rbind(
  uc.clin.forest,
  get_HR(fit.me, "Haemoglobin")
)

# Display plot and model summary
analysis_result$plot

```

### Time to patient-reported flare



```
invisible(cox_summary(fit.me))
```

Cox model summary:

Variable	HR	Lower 95%	Upper 95%	P-value
SexFemale	1.3056	0.9811	1.7374	0.0674
IMD2	1.2309	0.7538	2.0100	0.4064
IMD3	0.8915	0.5417	1.4672	0.6515

Variable	HR	Lower 95%	Upper 95%	P-value
IMD4	1.1646	0.7272	1.8651	0.5261
IMD5	1.1289	0.7162	1.7792	0.6016
catFC 50-250	1.6837	1.2762	2.2213	0.0002
catFC > 250	2.1975	1.6335	2.9564	0.0000
Haemoglobin	0.9857	0.9748	0.9967	0.0112

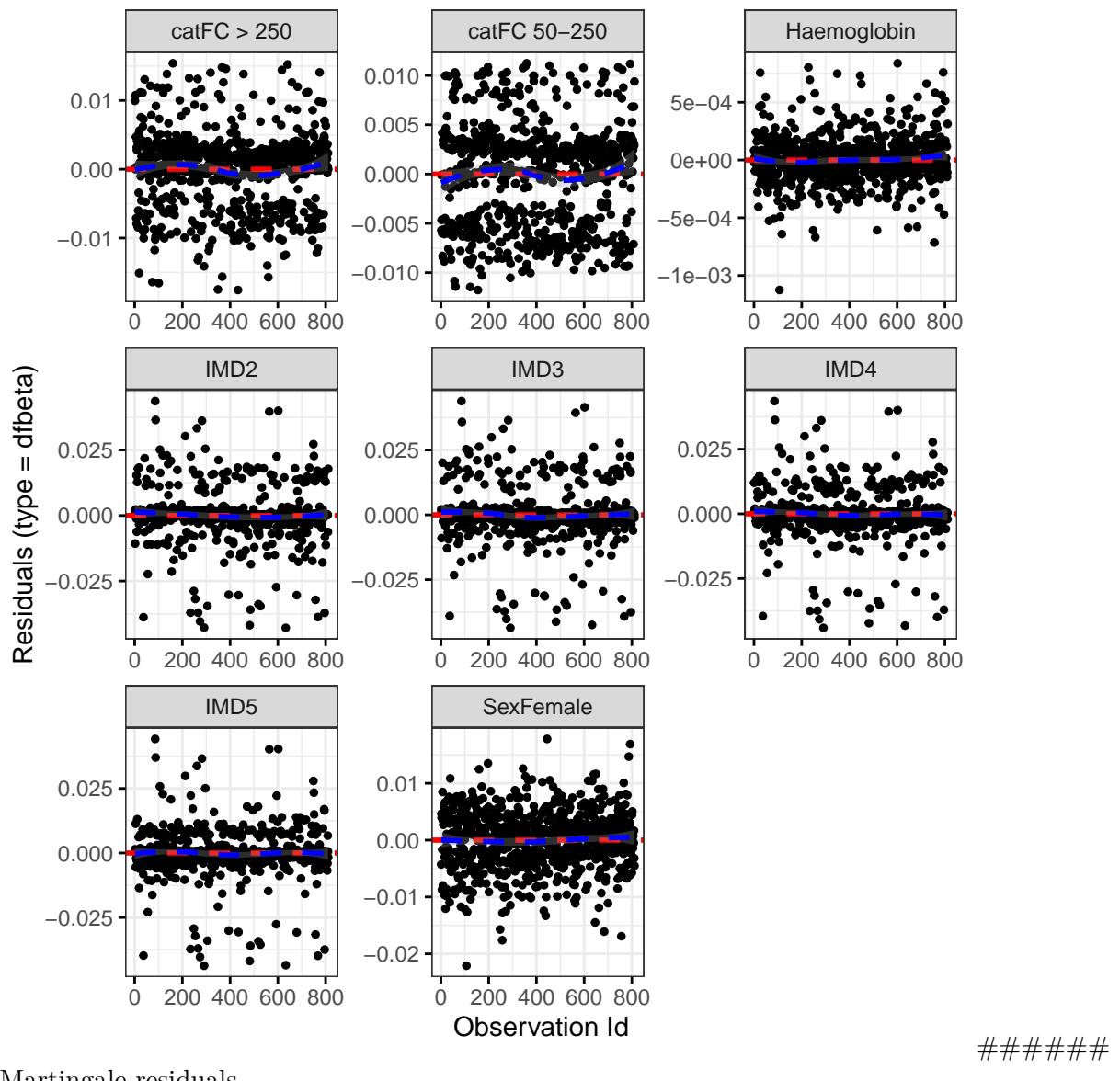
Diagnostics:

### Proportional hazards assumption test

	Chi-squared statistic	DF	P-value
Sex	0.9971	0.9836	0.3127
IMD	3.7353	3.9192	0.4308
cat	5.1028	1.9606	0.0752
Haemoglobin	0.0154	0.9753	0.8949
GLOBAL	10.3959	22.0376	0.9826

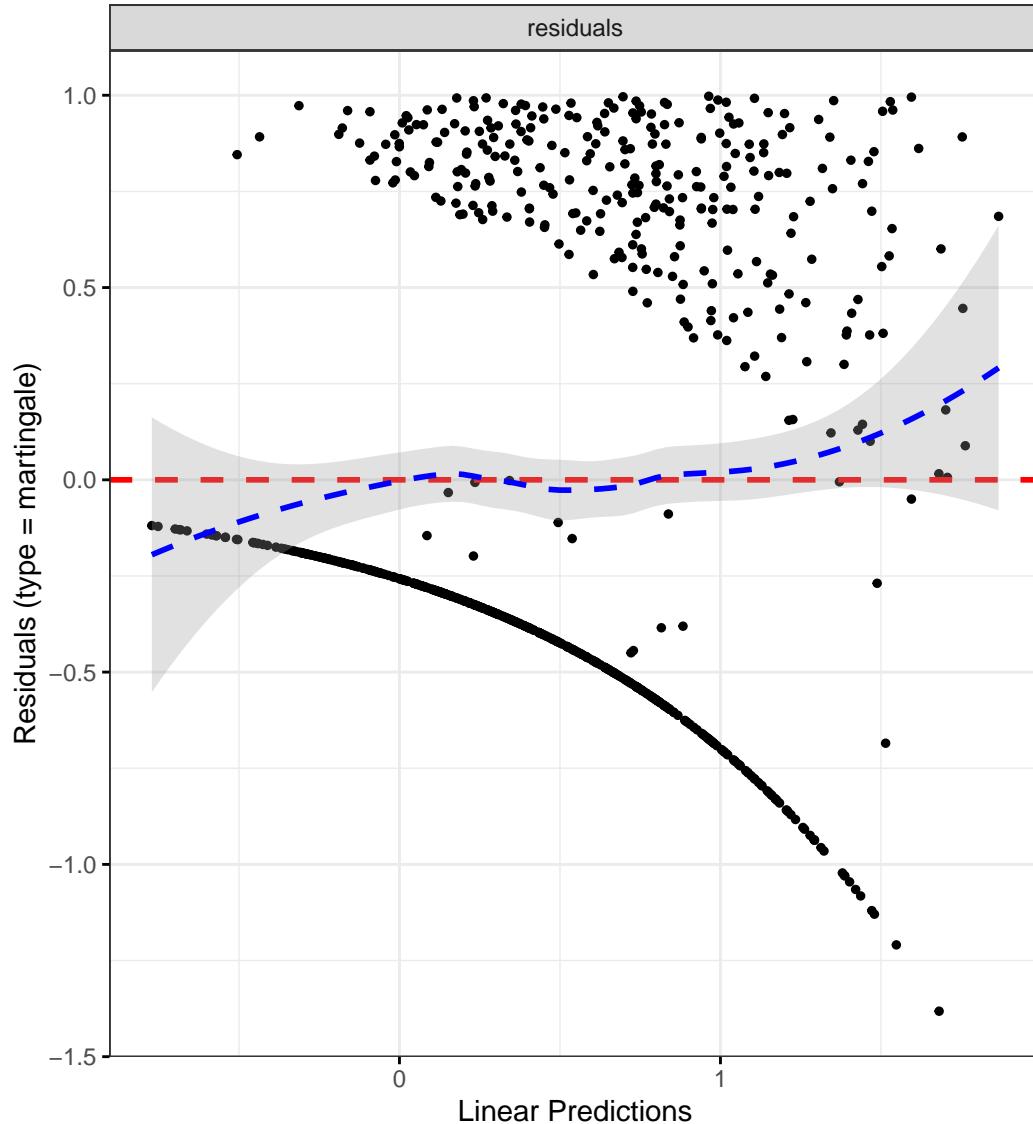
### DF betas

```
`geom_smooth()` using formula = 'y ~ x'
```



Martingale residuals

```
`geom_smooth()` using formula = 'y ~ x'
```



## Hard flare

```
# Generate survival plot and run Cox model for hard flare using utility function
analysis_result <- run_survival_analysis(
  data = flare.uc.df,
  var_name = "Haemoglobin",
  outcome_time = "hardflare_time",
  outcome_event = "hardflare",
  legend_title = "Haemoglobin",
```

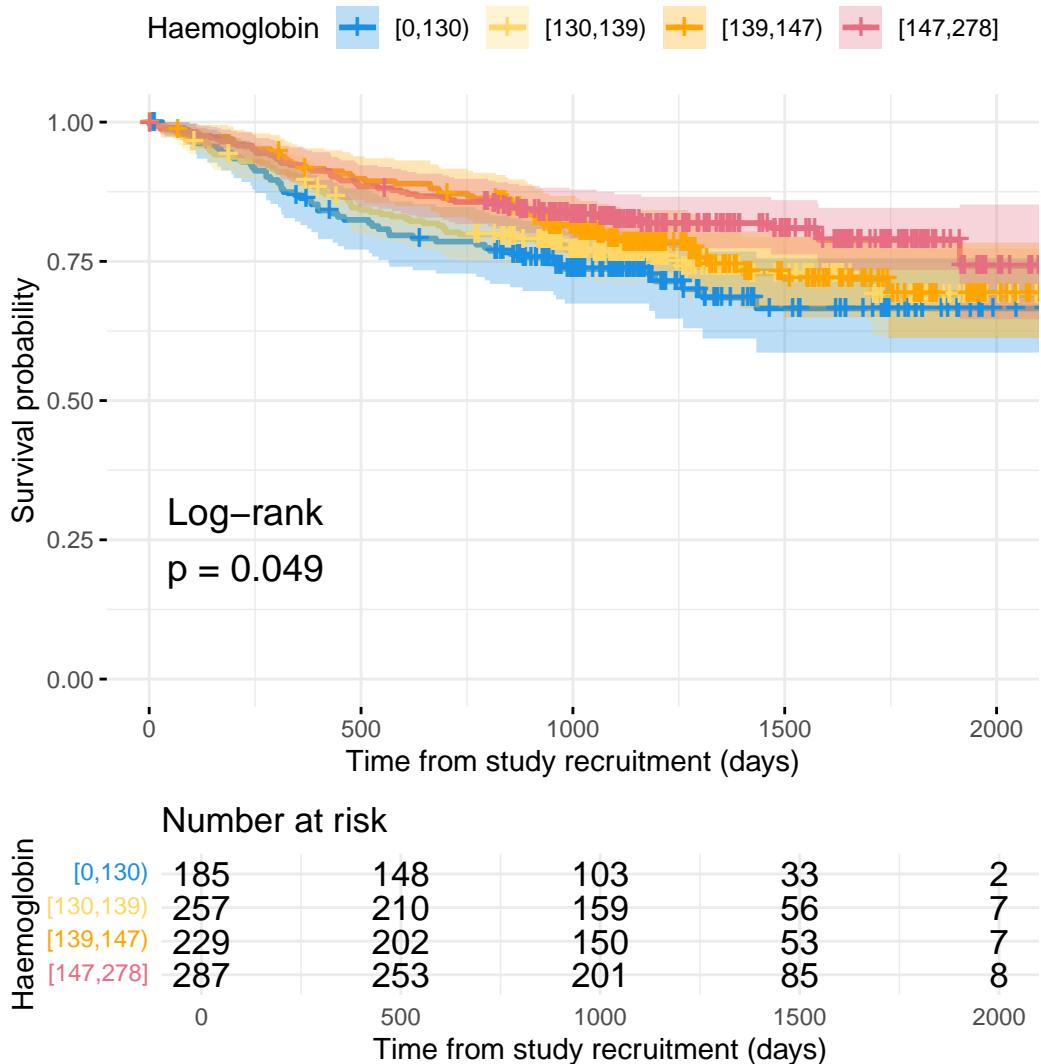
```
plot_base_path = "plots/uc/hard-flare/biochem/haemoglobin",
break_time_by = 500
)

# Extract hazard ratio for continuous Haemoglobin variable
fit.me <- coxph(
  Surv(hardflare_time, hardflare) ~
    Sex + IMD + cat + Haemoglobin + frailty(SiteNo),
  control = coxph.control(outer.max = 20),
  data = flare.uc.df
)

uc.hard.forest <- rbind(
  uc.hard.forest,
  get_HR(fit.me, "Haemoglobin")
)

# Display plot and model summary
analysis_result$plot
```

### Time to hard flare



```
invisible(cox_summary(fit.me))
```

Cox model summary:

Variable	HR	Lower 95%	Upper 95%	P-value
SexFemale	1.2574	0.8991	1.7585	0.1808
IMD2	1.3618	0.7341	2.5262	0.3274
IMD3	1.0655	0.5713	1.9873	0.8418

Variable	HR	Lower 95%	Upper 95%	P-value
IMD4	1.4946	0.8265	2.7028	0.1837
IMD5	1.2655	0.7096	2.2568	0.4250
catFC 50-250	1.9287	1.3727	2.7100	0.0002
catFC > 250	2.9167	2.0447	4.1607	0.0000
Haemoglobin	0.9932	0.9804	1.0062	0.3018

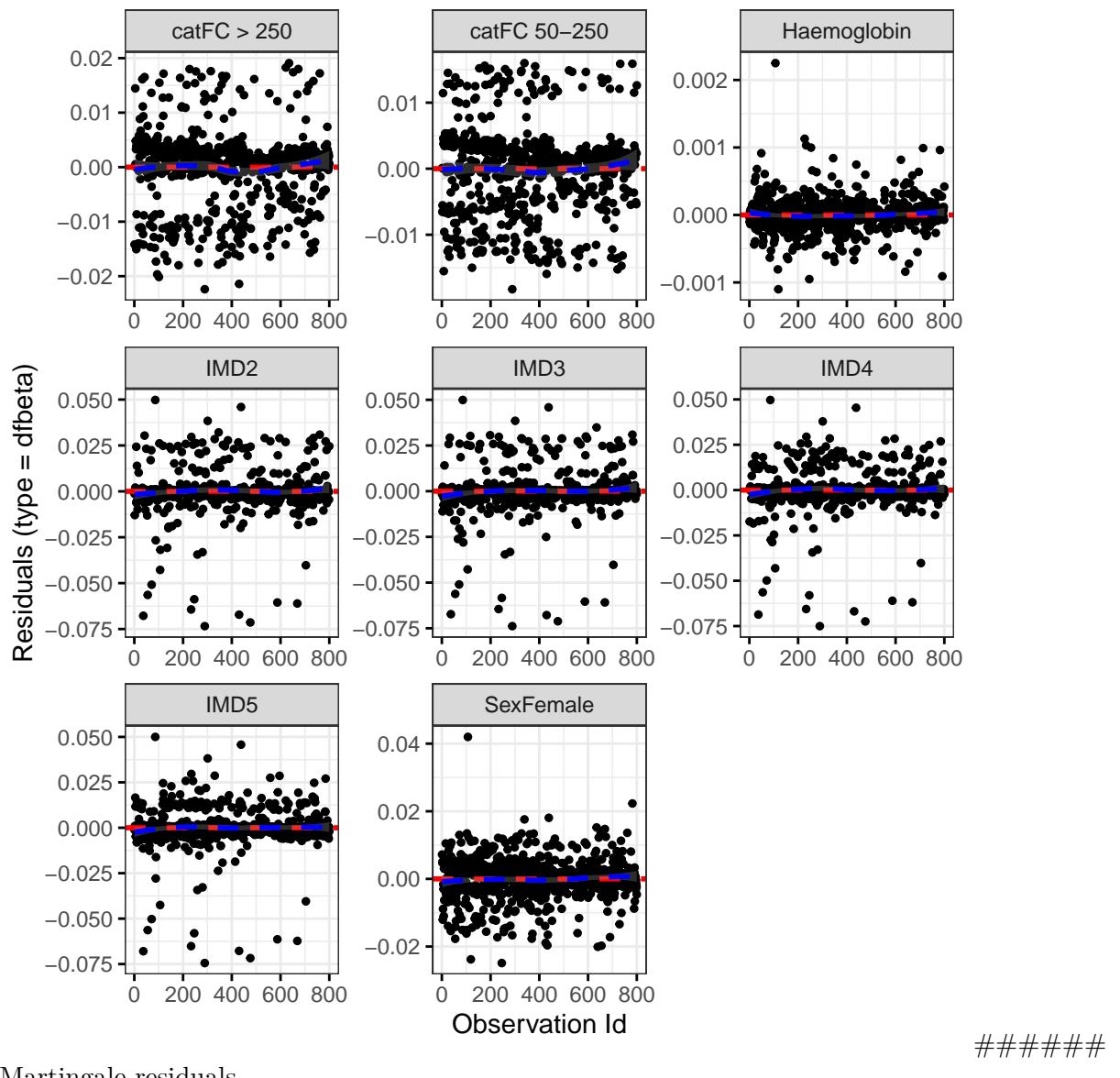
Diagnostics:

### Proportional hazards assumption test

	Chi-squared statistic	DF	P-value
Sex	0.0292	0.9843	0.8597
IMD	1.7907	3.9323	0.7653
cat	5.0265	1.9652	0.0785
Haemoglobin	0.8658	0.9821	0.3459
GLOBAL	7.2975	21.3166	0.9979

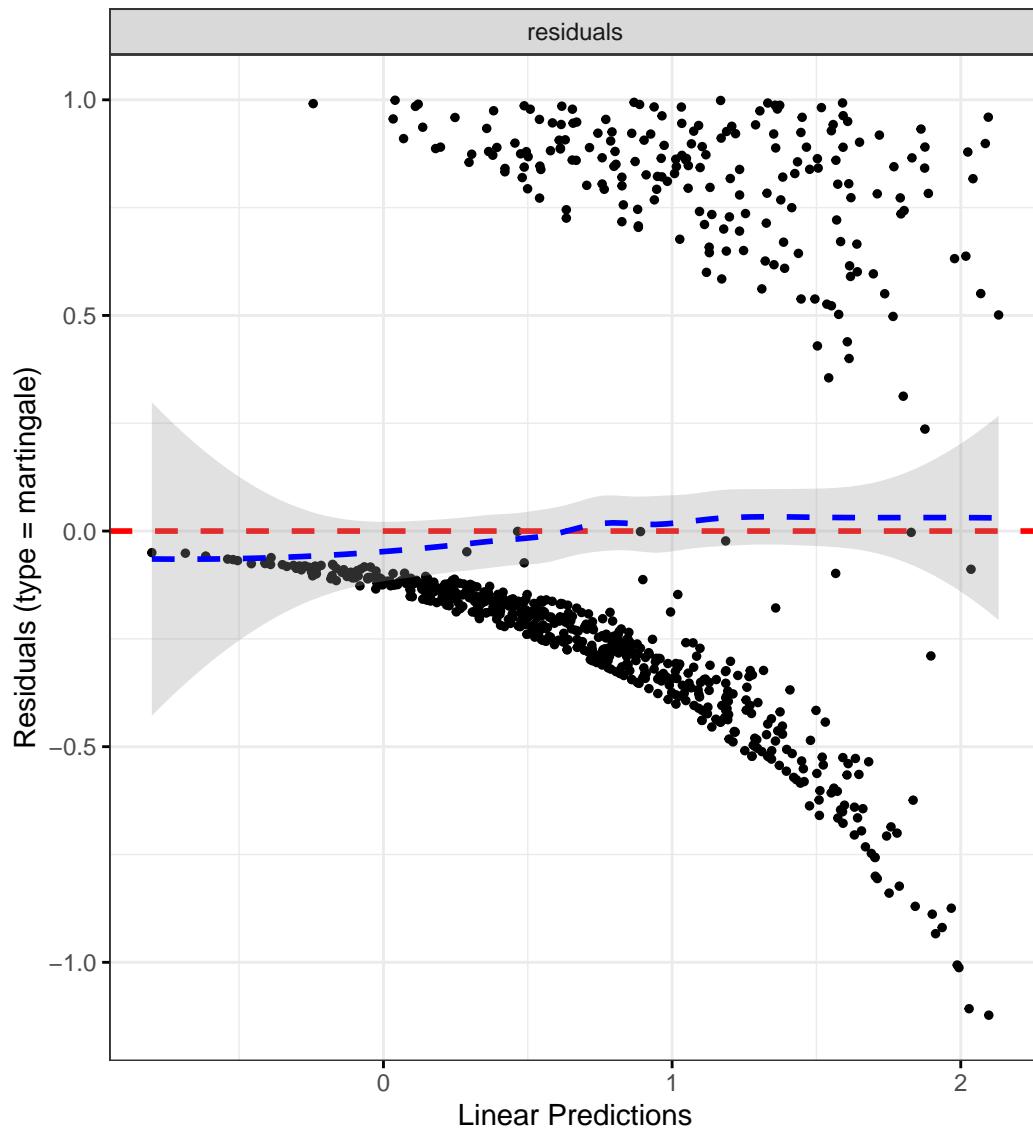
### DF betas

```
`geom_smooth()` using formula = 'y ~ x'
```



Martingale residuals

```
`geom_smooth()` using formula = 'y ~ x'
```



**White cell count**

**Crohn's disease**

**Patient-reported flare**

```
# Categorize White Cell Count by quantiles
flare.cd.df <- categorize_by_quantiles(flare.cd.df, "WCC", reference_data = flare.df)
```

```

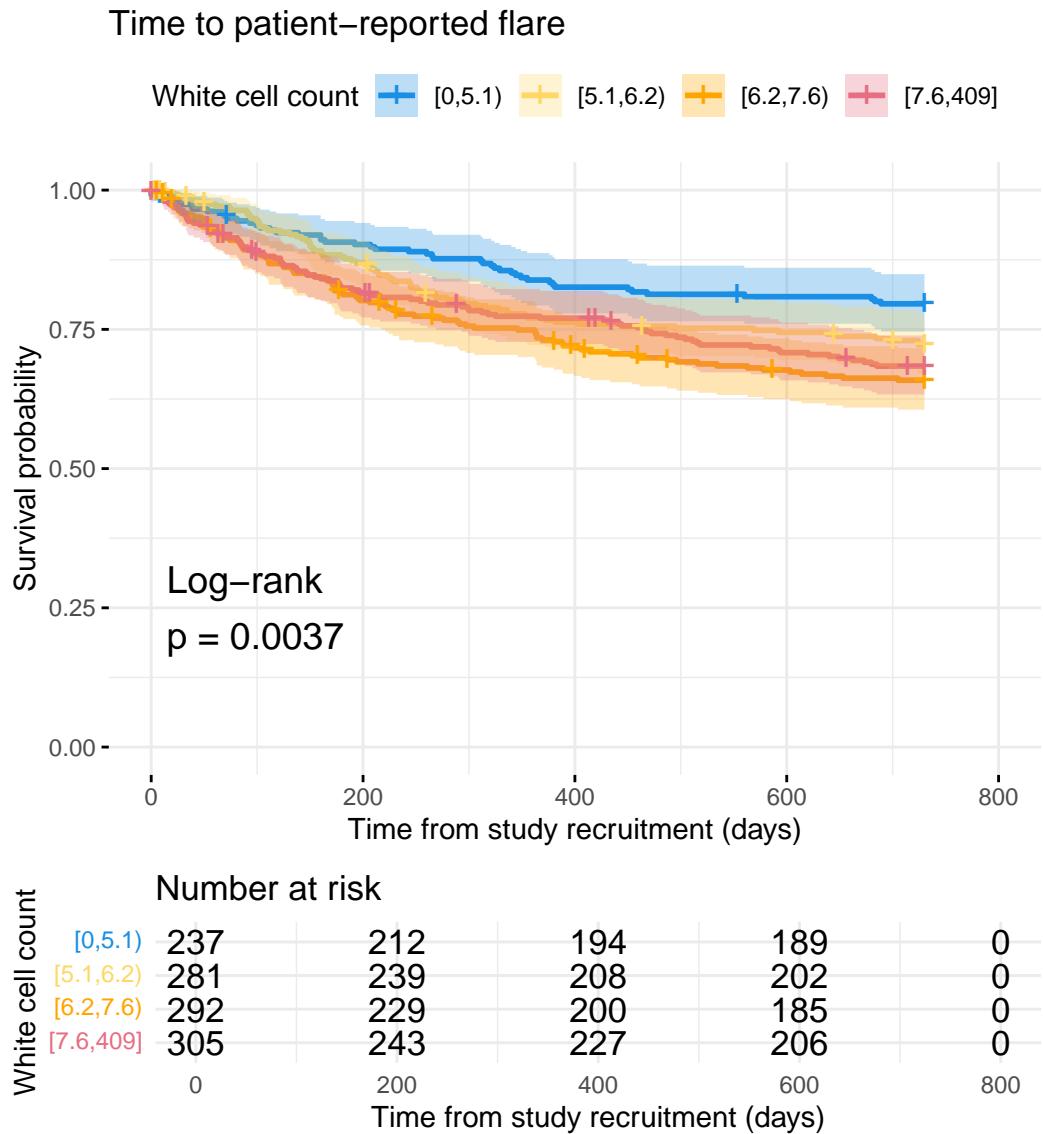
# Run survival analysis using utility function
analysis_result <- run_survival_analysis(
  data = flare.cd.df,
  var_name = "WCC",
  outcome_time = "softflare_time",
  outcome_event = "softflare",
  legend_title = "White cell count",
  plot_base_path = "plots/cd/soft-flare/biochem/wcc",
  break_time_by = 200
)

# Extract hazard ratio for continuous WCC variable
fit.me <- coxph(
  Surv(softflare_time, softflare) ~
    Sex + IMD + cat + WCC + frailty(SiteNo),
  control = coxph.control(outer.max = 20),
  data = flare.cd.df
)

cd.clin.forest <- rbind(
  cd.clin.forest,
  get_HR(fit.me, "WCC")
)

# Display plot and model summary
analysis_result$plot

```



```
invisible(cox_summary(fit.me))
```

Cox model summary:

Variable	HR	Lower 95%	Upper 95%	P-value
SexFemale	2.3006	1.7747	2.9823	0.0000
IMD2	1.1384	0.6732	1.9253	0.6287
IMD3	1.0421	0.6124	1.7731	0.8792

Variable	HR	Lower 95%	Upper 95%	P-value
IMD4	1.1244	0.6710	1.8841	0.6561
IMD5	1.0864	0.6591	1.7906	0.7452
catFC 50-250	1.5984	1.2140	2.1045	0.0008
catFC > 250	2.6410	1.9506	3.5758	0.0000
WCC	1.0101	0.9727	1.0490	0.6023

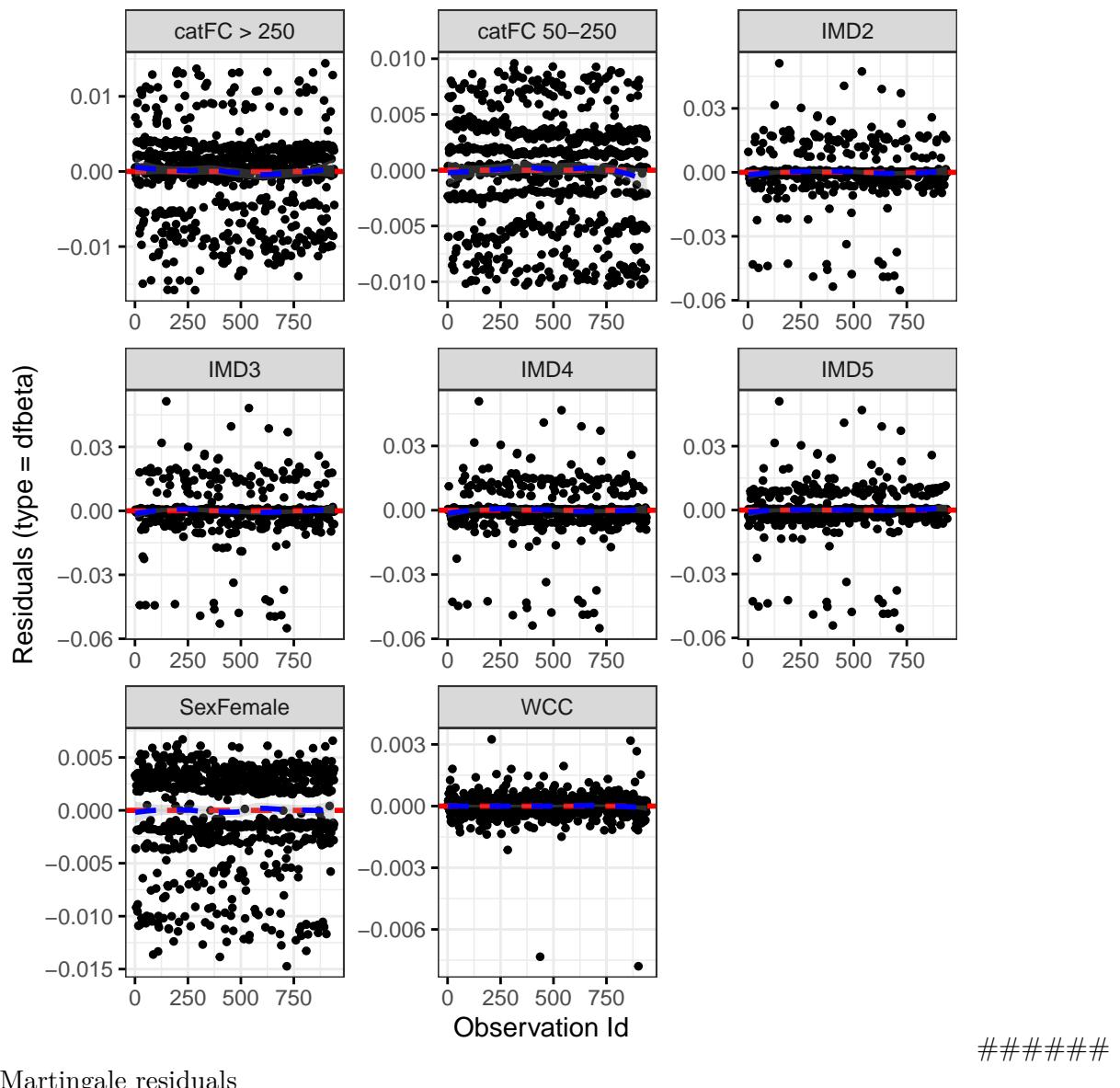
Diagnostics:

#### Proportional hazards assumption test

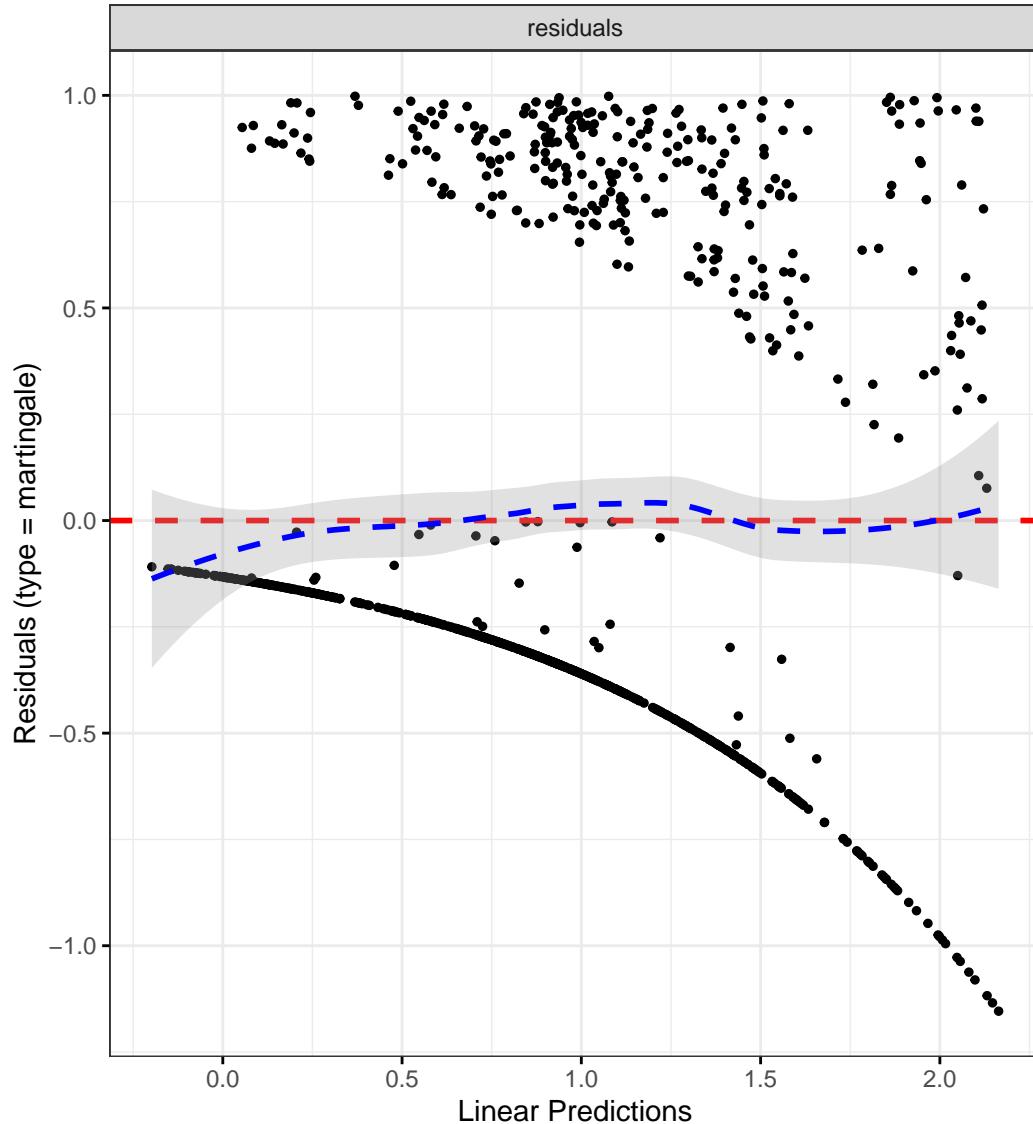
	Chi-squared statistic	DF	P-value
Sex	0.3188	0.9923	0.5691
IMD	4.2450	3.9543	0.3674
cat	4.1388	1.9840	0.1246
WCC	0.5574	0.9899	0.4513
GLOBAL	8.7191	13.4184	0.8182

#### DF betas

```
`geom_smooth()` using formula = 'y ~ x'
```



```
`geom_smooth()` using formula = 'y ~ x'
```



## Hard flare

```
# Generate survival plot and run Cox model for hard flare using utility function
analysis_result <- run_survival_analysis(
  data = flare.cd.df,
  var_name = "WCC",
  outcome_time = "hardflare_time",
  outcome_event = "hardflare",
  legend_title = "White cell count",
```

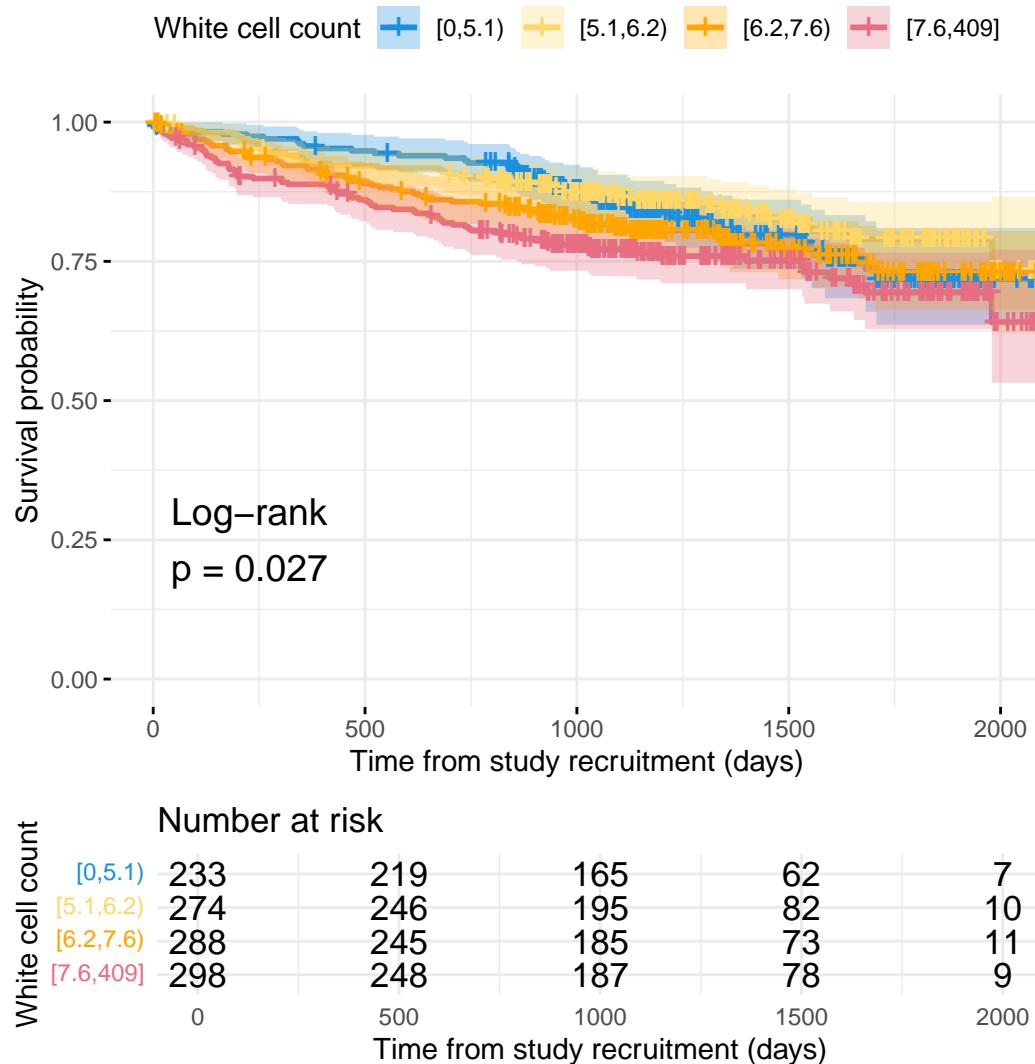
```
plot_base_path = "plots/cd/hard-flare/biochem/wcc",
break_time_by = 500
)

# Extract hazard ratio for continuous WCC variable
fit.me <- coxph(
  Surv(hardflare_time, hardflare) ~
    Sex + IMD + cat + WCC + frailty(SiteNo),
  control = coxph.control(outer.max = 20),
  data = flare.cd.df
)

cd.hard.forest <- rbind(
  cd.hard.forest,
  get_HR(fit.me, "WCC")
)

# Display plot and model summary
analysis_result$plot
```

### Time to hard flare



```
invisible(cox_summary(fit.me))
```

Cox model summary:

Variable	HR	Lower 95%	Upper 95%	P-value
SexFemale	1.4211	1.0561	1.9123	0.0203
IMD2	1.0084	0.5306	1.9167	0.9795
IMD3	1.1067	0.5800	2.1115	0.7585

Variable	HR	Lower 95%	Upper 95%	P-value
IMD4	1.0548	0.5616	1.9811	0.8682
IMD5	1.0054	0.5472	1.8474	0.9861
catFC 50-250	2.0582	1.4509	2.9199	0.0001
catFC > 250	3.5382	2.4336	5.1442	0.0000
WCC	1.0336	0.9885	1.0809	0.1468

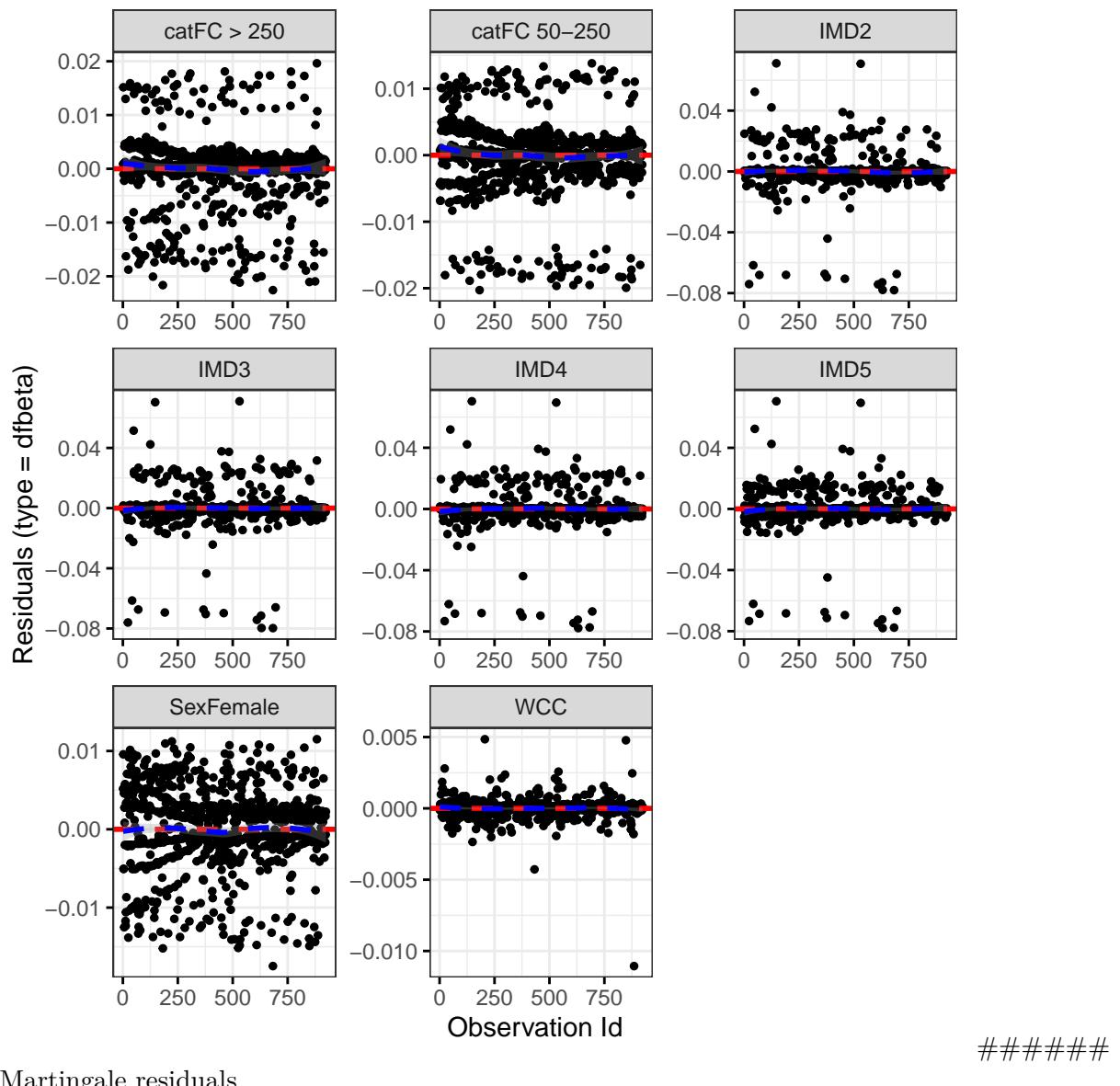
Diagnostics:

#### Proportional hazards assumption test

	Chi-squared statistic	DF	P-value
Sex	2.7970	0.9858	0.0926
IMD	3.6523	3.9454	0.4468
cat	9.7104	1.9851	0.0077
WCC	2.9462	0.9726	0.0829
GLOBAL	19.9546	18.6478	0.3752

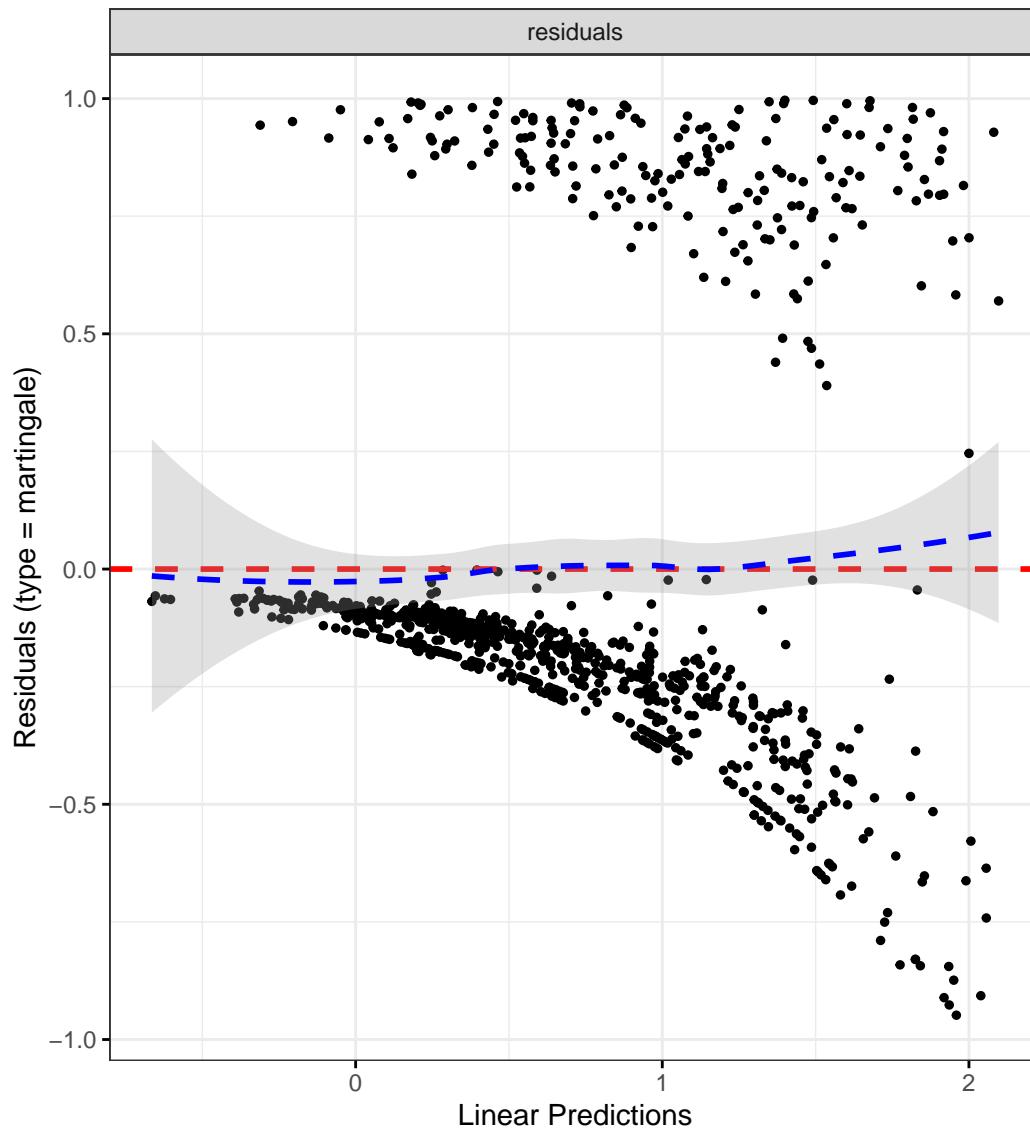
#### DF betas

```
`geom_smooth()` using formula = 'y ~ x'
```



Martingale residuals

```
`geom_smooth()` using formula = 'y ~ x'
```



### Ulcerative colitis

#### Patient-reported flare

```
# Categorize White Cell Count by quantiles
flare.uc.df <- categorize_by_quantiles(flare.uc.df, "WCC", reference_data = flare.df)

# Run survival analysis using utility function
analysis_result <- run_survival_analysis()
```

```

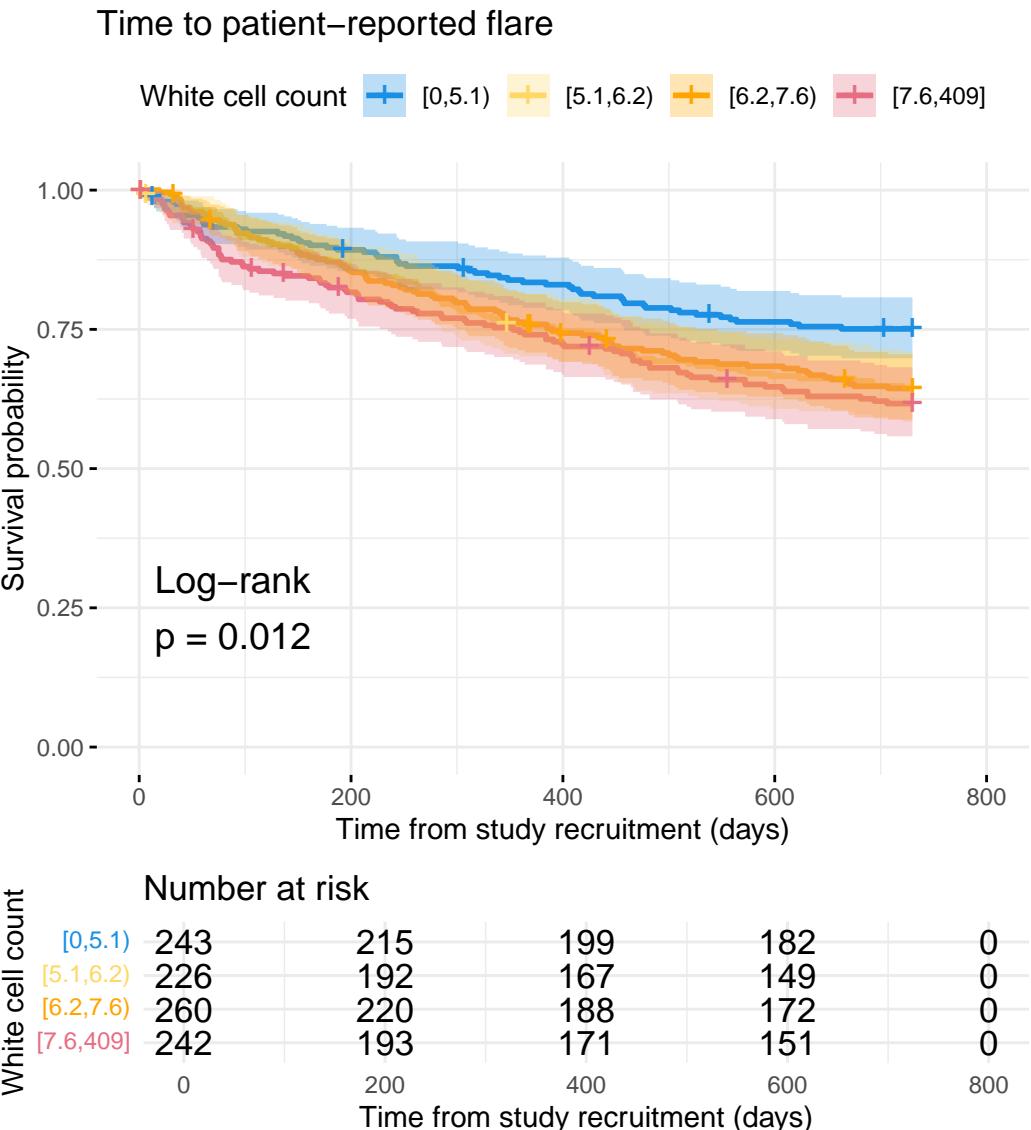
data = flare.uc.df,
var_name = "WCC",
outcome_time = "softflare_time",
outcome_event = "softflare",
legend_title = "White cell count",
plot_base_path = "plots/uc/soft-flare/biochem/wcc",
break_time_by = 200
)

# Extract hazard ratio for continuous WCC variable
fit.me <- coxph(
  Surv(softflare_time, softflare) ~
    Sex + IMD + cat + WCC + frailty(SiteNo),
  control = coxph.control(outer.max = 20),
  data = flare.uc.df
)

uc.clin.forest <- rbind(
  uc.clin.forest,
  get_HR(fit.me, "WCC")
)

# Display plot and model summary
analysis_result$plot

```



```
invisible(cox_summary(fit.me))
```

Cox model summary:

Variable	HR	Lower 95%	Upper 95%	P-value
SexFemale	1.6030	1.2625	2.0353	0.0001
IMD2	1.2114	0.7424	1.9766	0.4427
IMD3	0.9092	0.5532	1.4945	0.7075

Variable	HR	Lower 95%	Upper 95%	P-value
IMD4	1.1339	0.7087	1.8142	0.6002
IMD5	1.1158	0.7084	1.7576	0.6364
catFC 50-250	1.6685	1.2628	2.2047	0.0003
catFC > 250	2.1548	1.5941	2.9127	0.0000
WCC	1.0834	1.0230	1.1475	0.0062

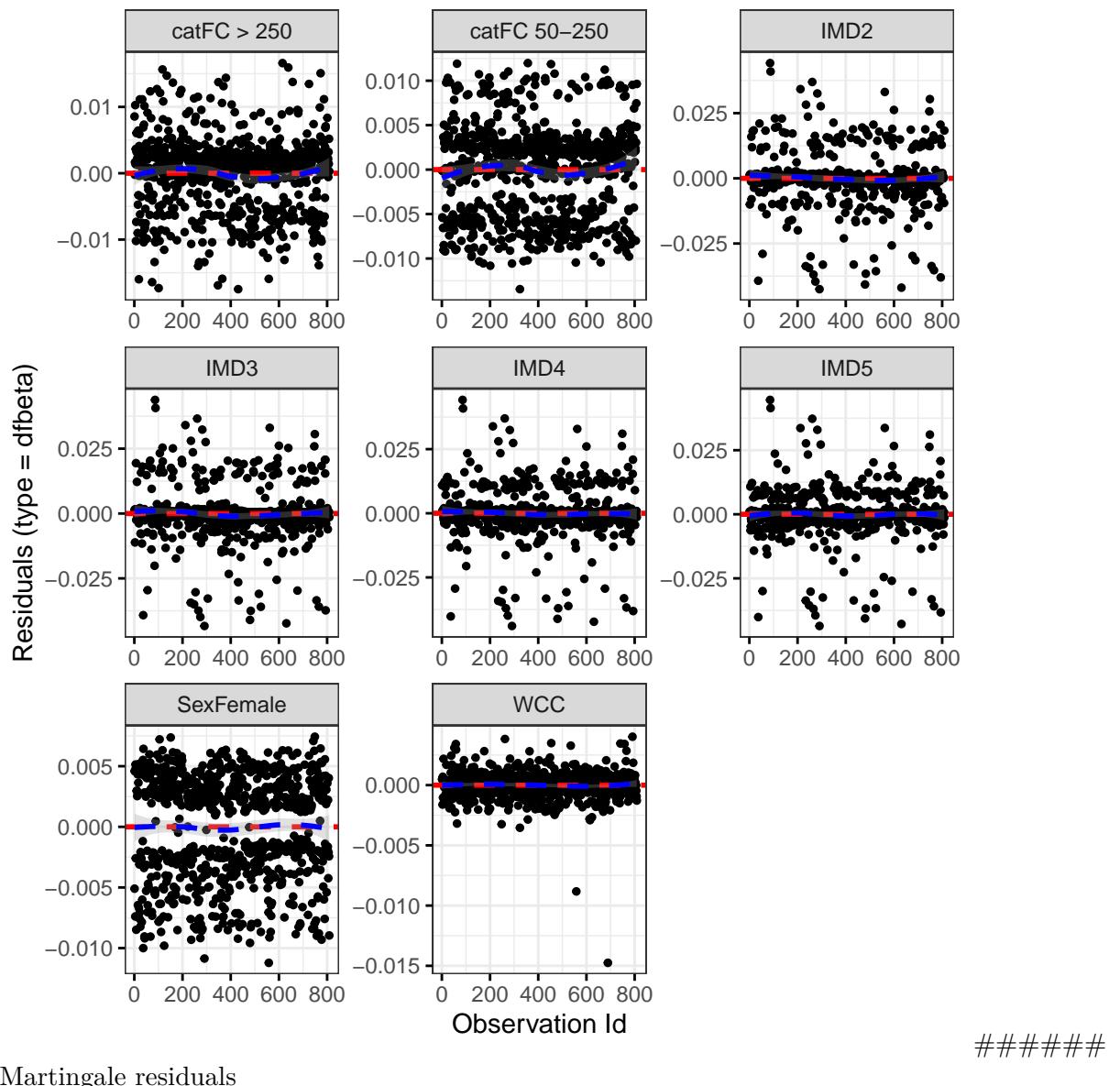
Diagnostics:

### Proportional hazards assumption test

	Chi-squared statistic	DF	P-value
Sex	1.0472	0.9870	0.3020
IMD	3.5238	3.9210	0.4621
cat	4.8428	1.9607	0.0857
WCC	0.3476	0.9723	0.5439
GLOBAL	9.5136	22.3170	0.9916

### DF betas

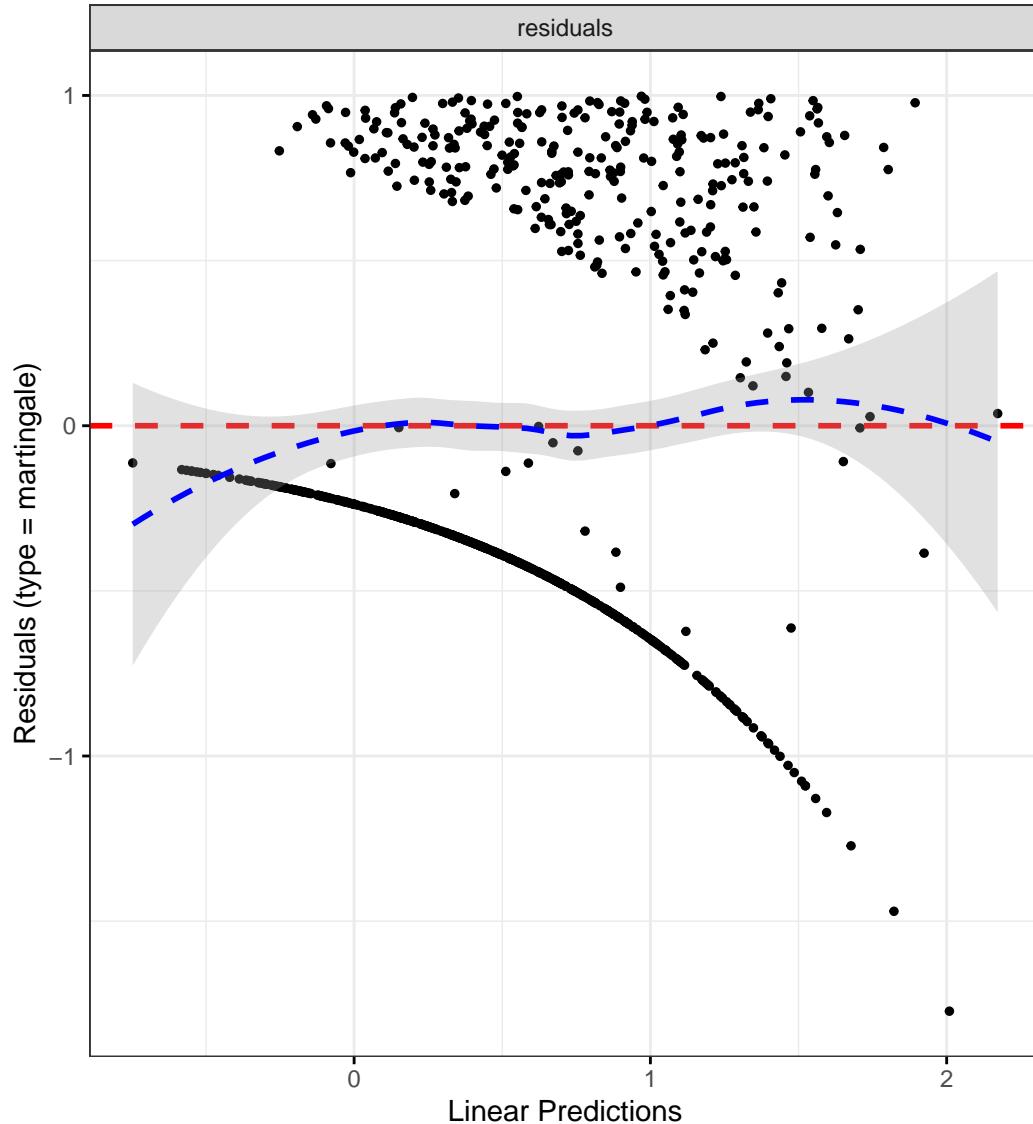
```
`geom_smooth()` using formula = 'y ~ x'
```



#####

Martingale residuals

```
`geom_smooth()` using formula = 'y ~ x'
```



## Hard flare

```
# Generate survival plot and run Cox model for hard flare using utility function
analysis_result <- run_survival_analysis(
  data = flare.uc.df,
  var_name = "WCC",
  outcome_time = "hardflare_time",
  outcome_event = "hardflare",
  legend_title = "White cell count",
```

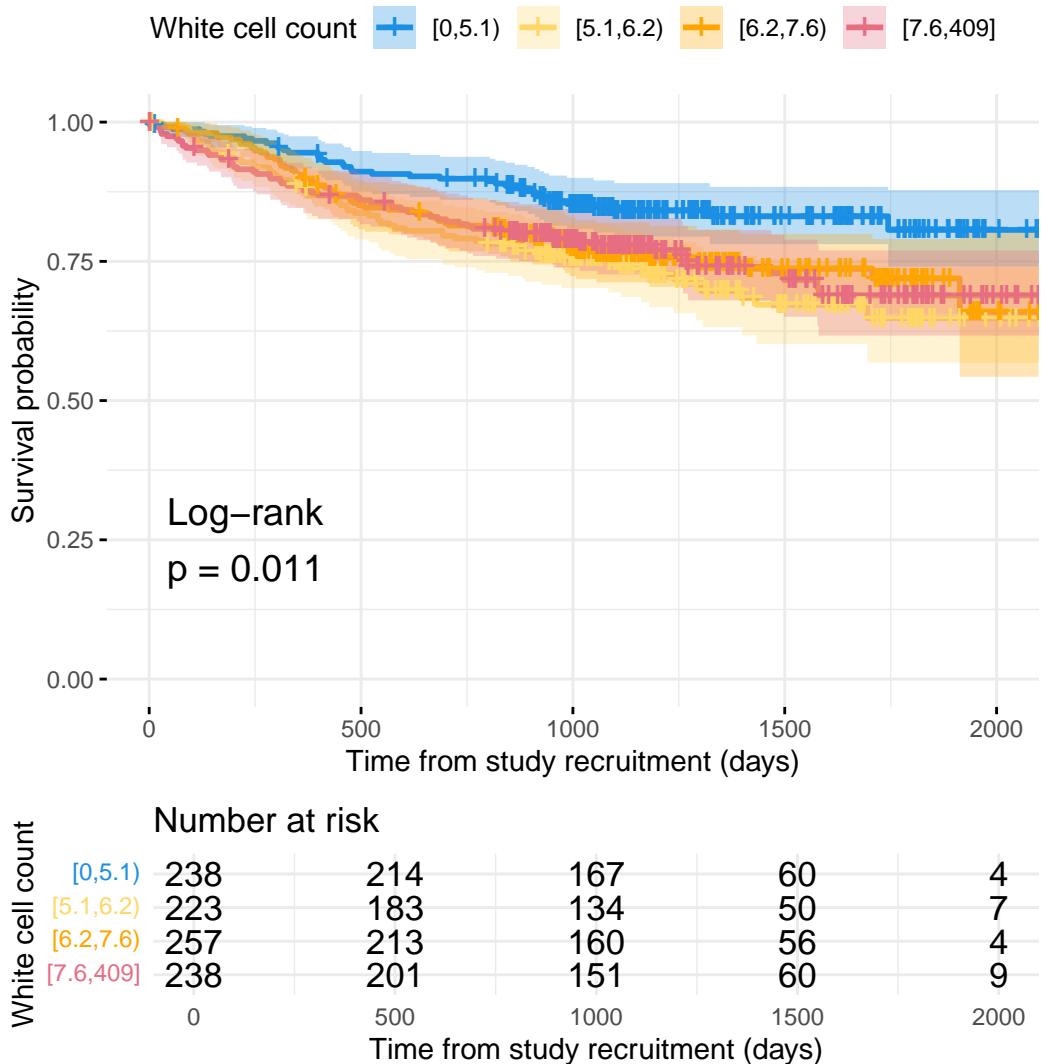
```
plot_base_path = "plots/uc/hard-flare/biochem/wcc",
break_time_by = 500
)

# Extract hazard ratio for continuous WCC variable
fit.me <- coxph(
  Surv(hardflare_time, hardflare) ~
    Sex + IMD + cat + WCC + frailty(SiteNo),
  control = coxph.control(outer.max = 20),
  data = flare.uc.df
)

uc.hard.forest <- rbind(
  uc.hard.forest,
  get_HR(fit.me, "WCC")
)

# Display plot and model summary
analysis_result$plot
```

### Time to hard flare



```
invisible(cox_summary(fit.me))
```

Cox model summary:

Variable	HR	Lower 95%	Upper 95%	P-value
SexFemale	1.3848	1.0406	1.8428	0.0256
IMD2	1.3707	0.7387	2.5434	0.3174
IMD3	1.0676	0.5725	1.9910	0.8370

Variable	HR	Lower 95%	Upper 95%	P-value
IMD4	1.4771	0.8170	2.6706	0.1967
IMD5	1.2764	0.7159	2.2756	0.4082
catFC 50-250	1.9211	1.3662	2.7012	0.0002
catFC > 250	2.8875	2.0164	4.1351	0.0000
WCC	1.0417	0.9716	1.1169	0.2504

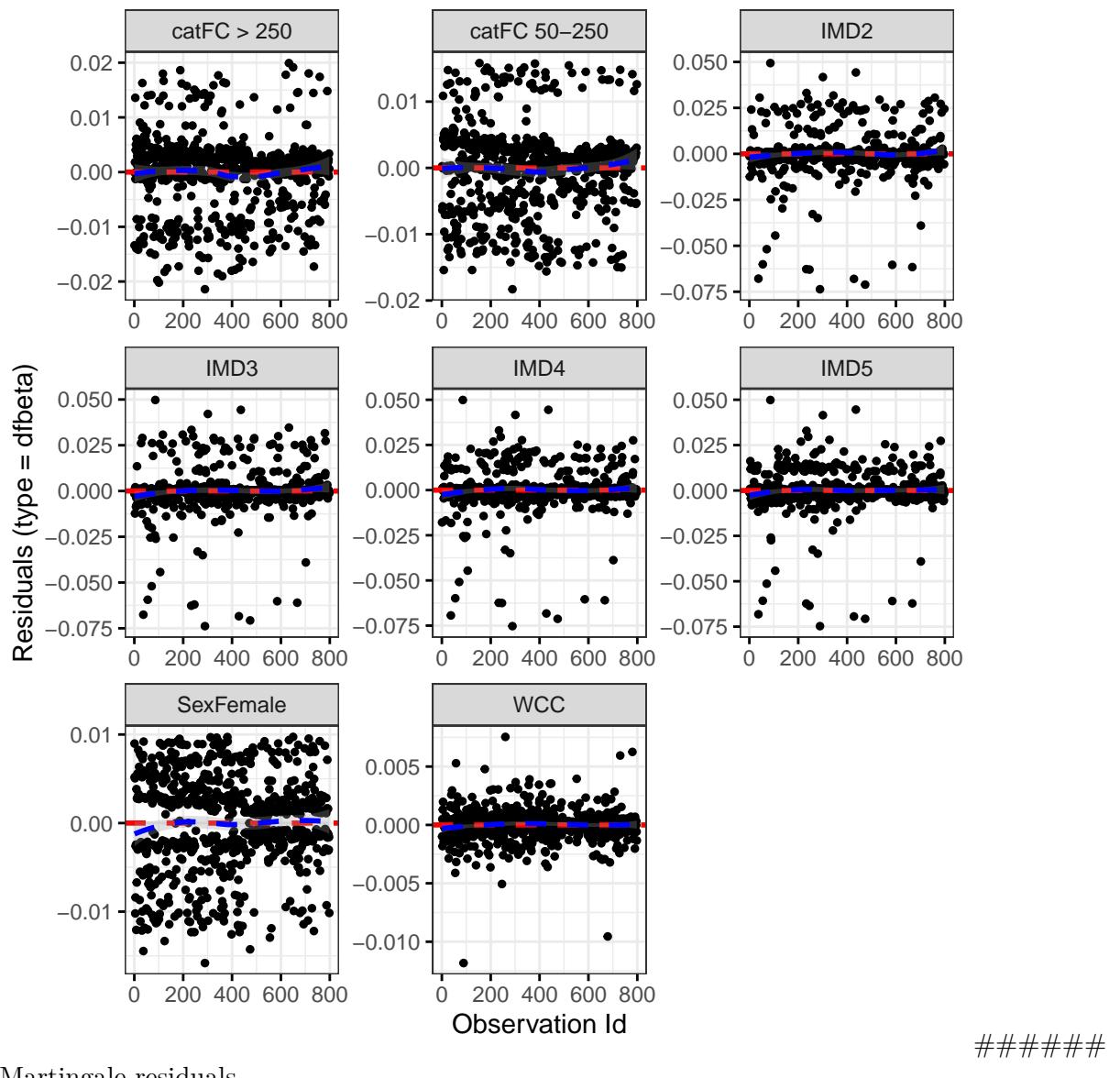
Diagnostics:

### Proportional hazards assumption test

	Chi-squared statistic	DF	P-value
Sex	0.0397	0.9850	0.8373
IMD	1.7908	3.9311	0.7652
cat	5.0693	1.9653	0.0768
WCC	0.4463	0.9791	0.4955
GLOBAL	6.8088	21.7201	0.9990

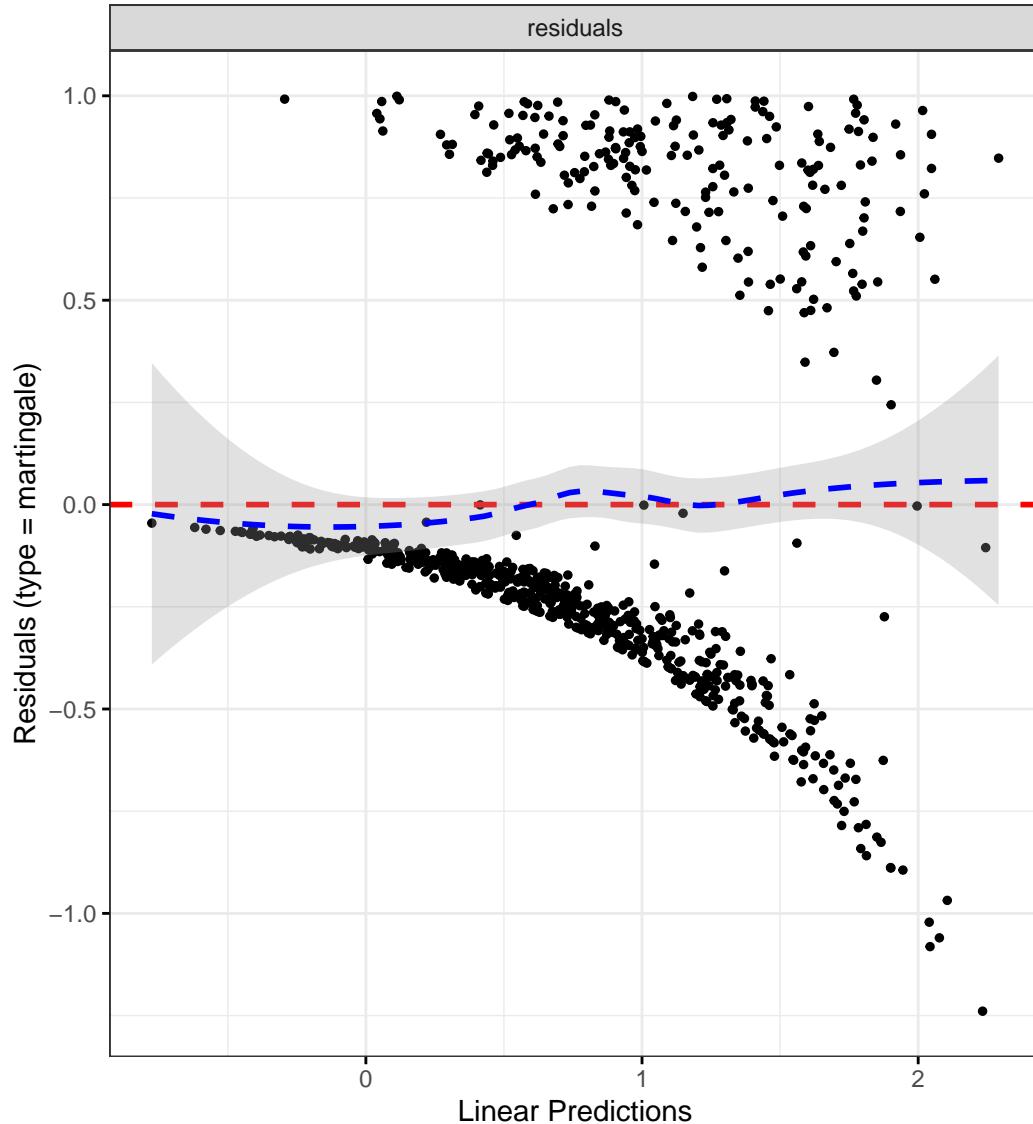
### DF betas

```
`geom_smooth()` using formula = 'y ~ x'
```



Martingale residuals

```
`geom_smooth()` using formula = 'y ~ x'
```



## Platelets

There is significant separation between Kaplan-Meier curves for platelets for both soft and hard flares in CD. However, this significance is lost when controlling for FC and other variables via Cox regression. No significant findings are reported for UC.

## Crohn's disease

### Patient-reported flare

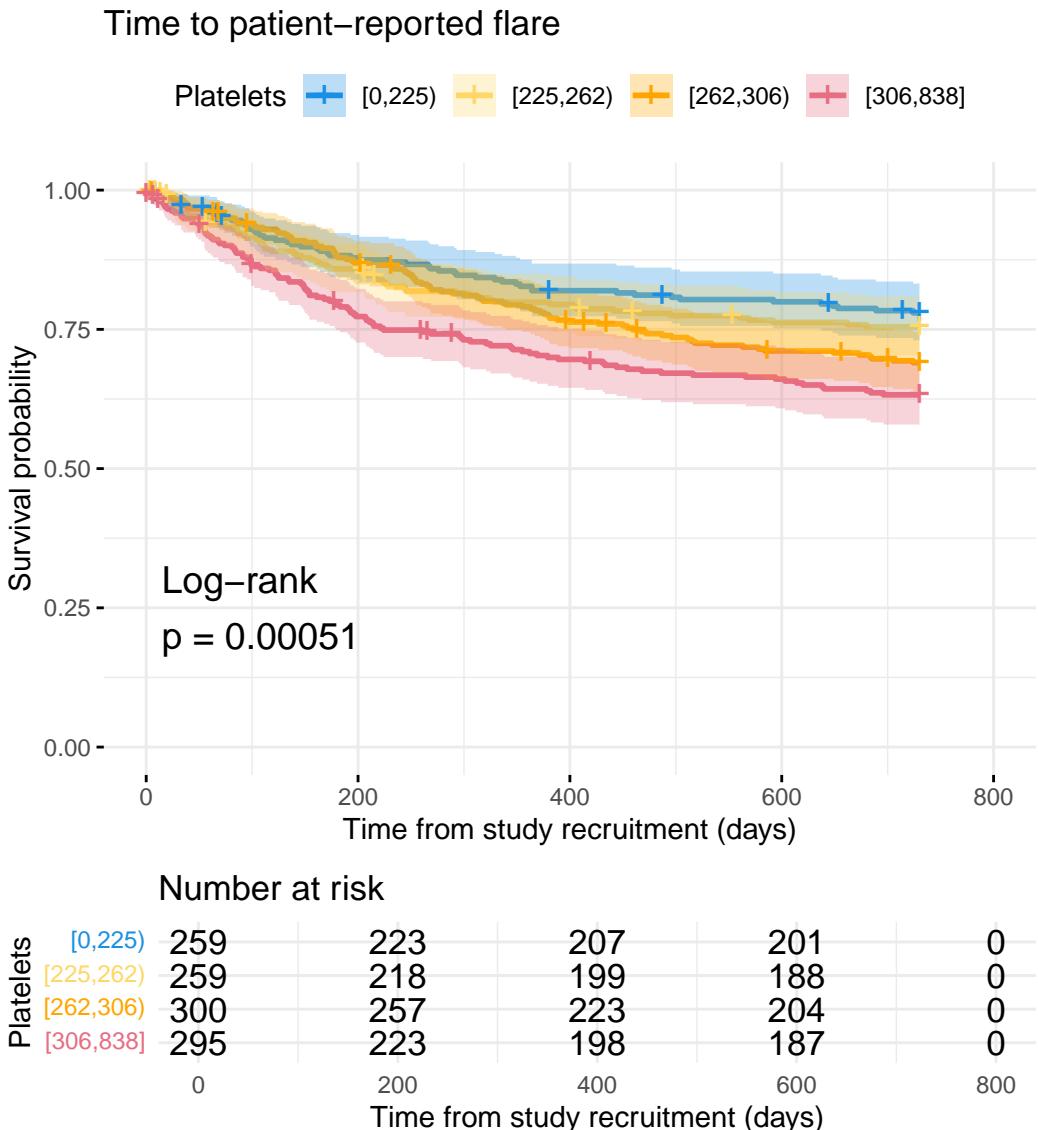
```
# Categorize Platelets by quantiles
flare.cd.df <- categorize_by_quantiles(flare.cd.df, "Platelets", reference_data = flare.df)

# Run survival analysis using utility function
analysis_result <- run_survival_analysis(
  data = flare.cd.df,
  var_name = "Platelets",
  outcome_time = "softflare_time",
  outcome_event = "softflare",
  legend_title = "Platelets",
  plot_base_path = "plots/cd/soft-flare/biochem/Platelets",
  break_time_by = 200
)

# Extract hazard ratio for continuous Platelets variable
fit.me <- coxph(
  Surv(softflare_time, softflare) ~
    Sex + IMD + cat + Platelets + frailty(SiteNo),
  control = coxph.control(outer.max = 20),
  data = flare.cd.df
)

cd.clin.forest <- rbind(
  cd.clin.forest,
  get_HR(fit.me, "Platelets")
)

# Display plot and model summary
analysis_result$plot
```



```
invisible(cox_summary(fit.me))
```

Cox model summary:

Variable	HR	Lower 95%	Upper 95%	P-value
SexFemale	2.2495	1.7228	2.9373	0.0000
IMD2	1.1311	0.6683	1.9143	0.6464
IMD3	1.0335	0.6075	1.7583	0.9033

Variable	HR	Lower 95%	Upper 95%	P-value
IMD4	1.1078	0.6607	1.8576	0.6978
IMD5	1.0735	0.6509	1.7704	0.7811
catFC 50-250	1.5936	1.2107	2.0977	0.0009
catFC > 250	2.5785	1.8953	3.5082	0.0000
Platelets	1.0007	0.9991	1.0023	0.3859

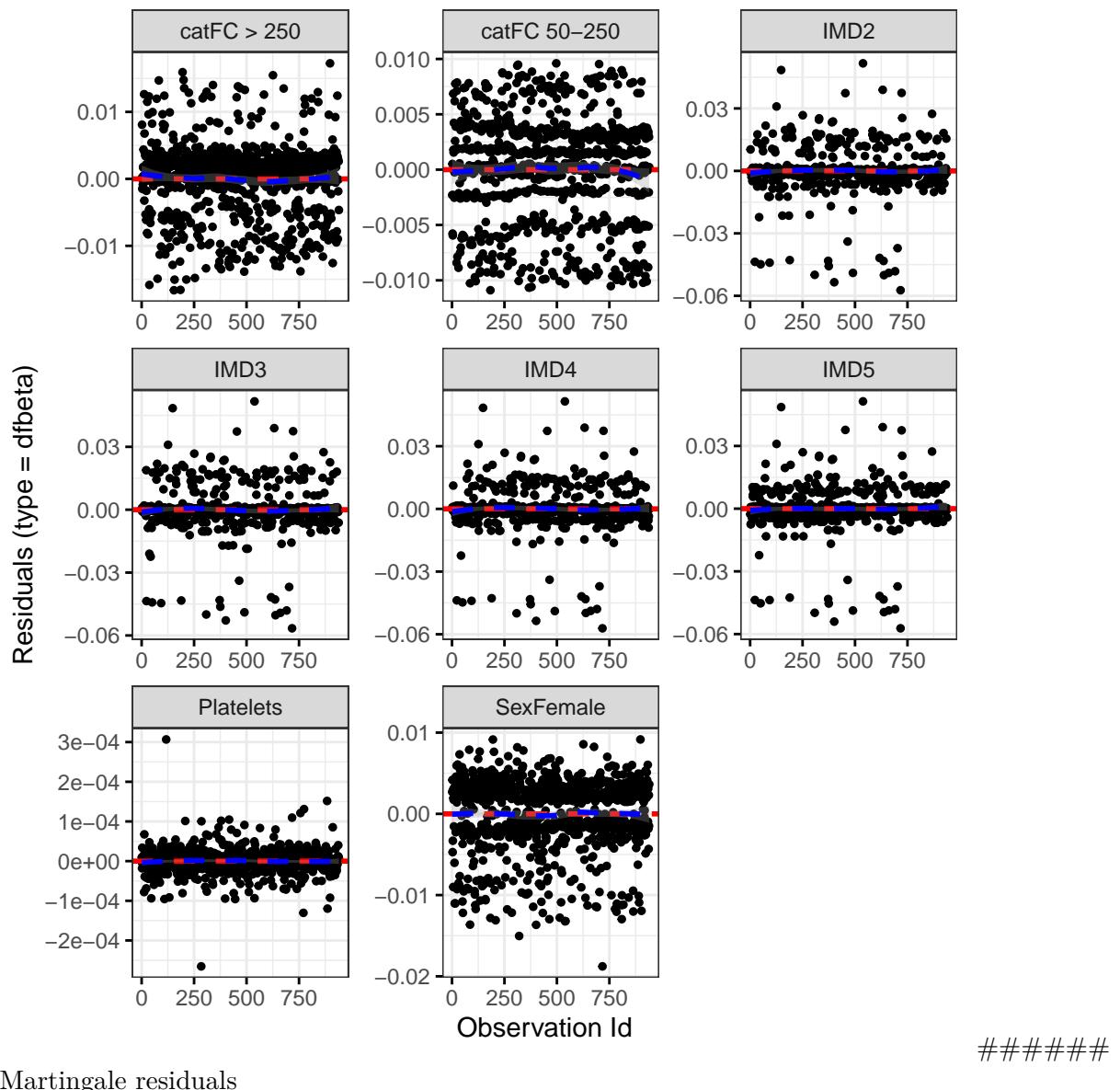
Diagnostics:

### Proportional hazards assumption test

	Chi-squared statistic	DF	P-value
Sex	0.3063	0.9934	0.5772
IMD	4.2714	3.9552	0.3642
cat	4.1818	1.9843	0.1220
Platelets	1.0975	0.9900	0.2917
GLOBAL	9.0260	13.2643	0.7876

### DF betas

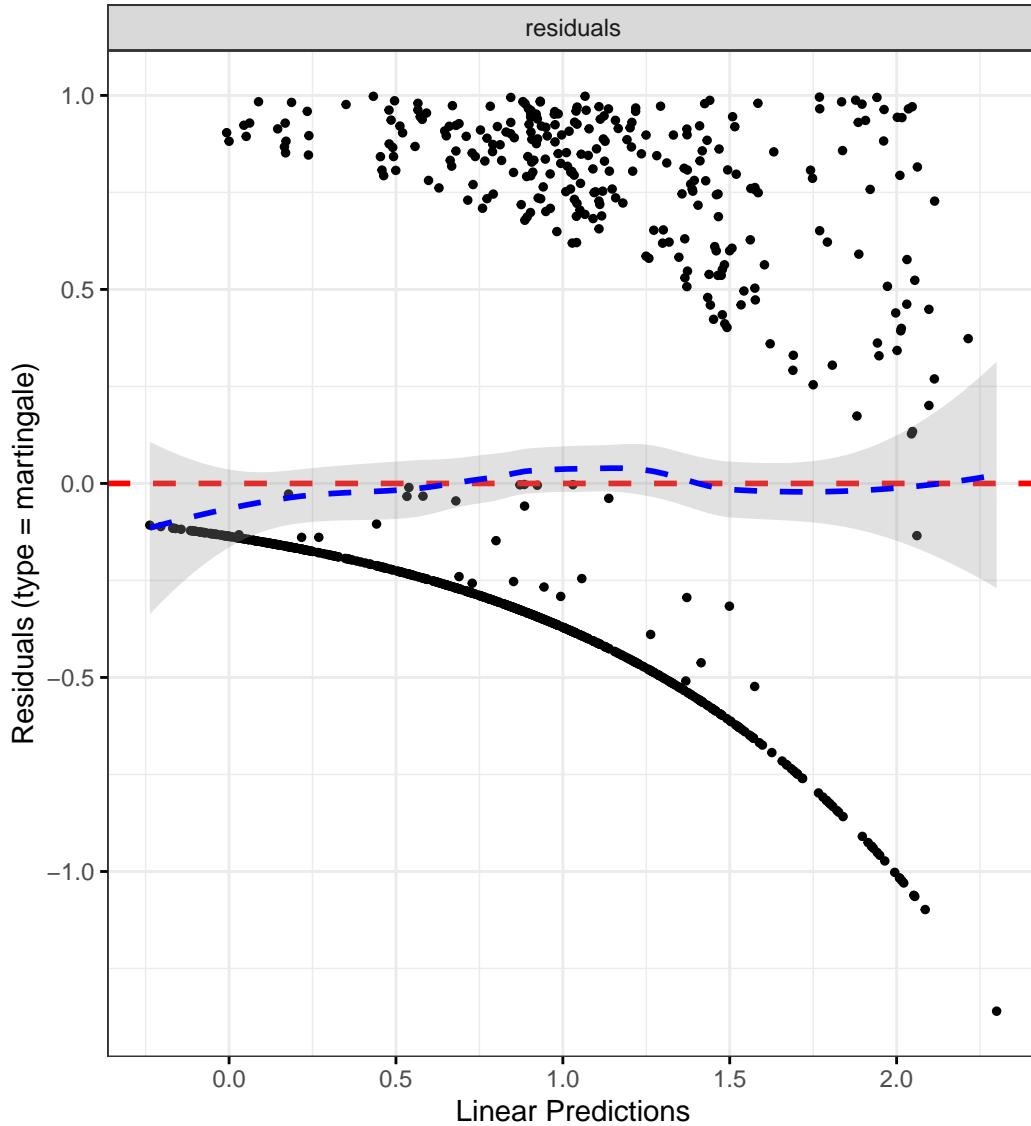
```
`geom_smooth()` using formula = 'y ~ x'
```



#####

Martingale residuals

```
`geom_smooth()` using formula = 'y ~ x'
```



### Hard flare

```
# Generate survival plot and run Cox model for hard flare using utility function
analysis_result <- run_survival_analysis(
  data = flare.cd.df,
  var_name = "Platelets",
  outcome_time = "hardflare_time",
  outcome_event = "hardflare",
  legend_title = "Platelets",
```

```

plot_base_path = "plots/cd/hard-flare/biochem/Platelets",
break_time_by = 500
)

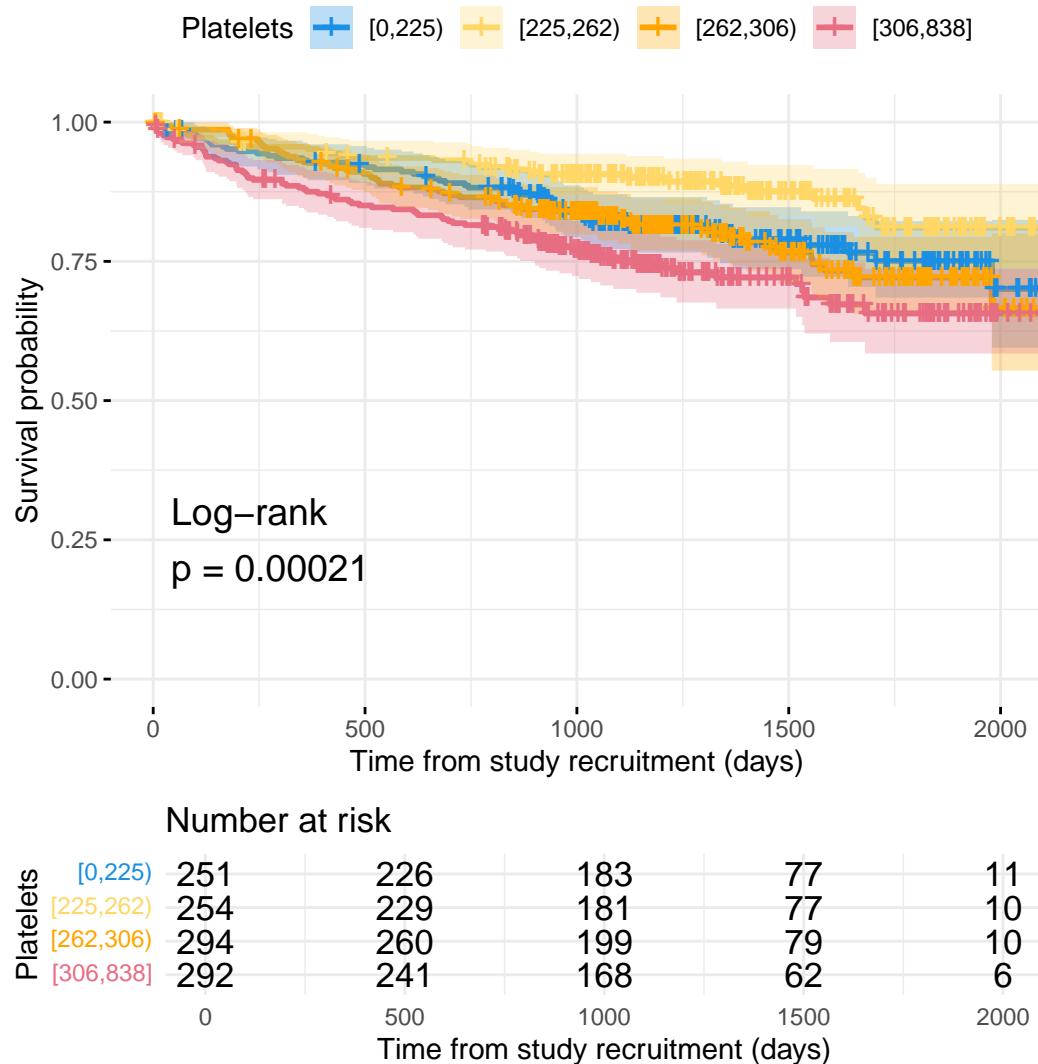
# Extract hazard ratio for continuous Platelets variable
fit.me <- coxph(
  Surv(hardflare_time, hardflare) ~
    Sex + IMD + cat + Platelets + frailty(SiteNo),
  control = coxph.control(outer.max = 20),
  data = flare.cd.df
)

cd.hard.forest <- rbind(
  cd.hard.forest,
  get_HR(fit.me, "Platelets")
)

# Display plot and model summary
analysis_result$plot

```

### Time to hard flare



```
invisible(cox_summary(fit.me))
```

Cox model summary:

Variable	HR	Lower 95%	Upper 95%	P-value
SexFemale	1.3809	1.0170	1.8750	0.0386
IMD2	1.0000	0.5261	1.9011	0.9999
IMD3	1.1066	0.5807	2.1088	0.7582

Variable	HR	Lower 95%	Upper 95%	P-value
IMD4	1.0363	0.5519	1.9458	0.9116
IMD5	0.9874	0.5375	1.8140	0.9674
catFC 50-250	2.0369	1.4350	2.8913	0.0001
catFC > 250	3.4517	2.3653	5.0372	0.0000
Platelets	1.0010	0.9993	1.0028	0.2474

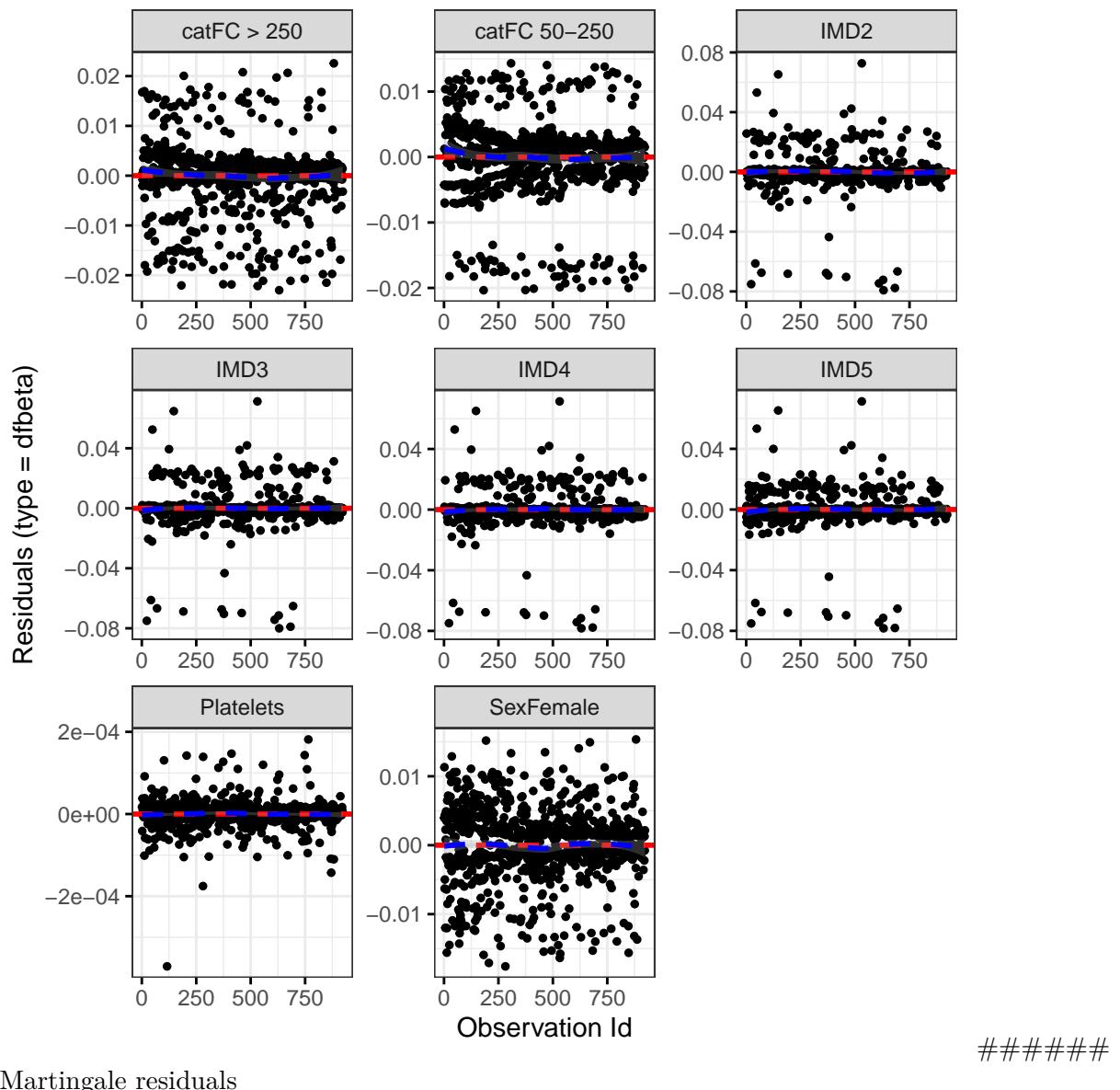
Diagnostics:

### Proportional hazards assumption test

	Chi-squared statistic	DF	P-value
Sex	2.7745	0.9876	0.0942
IMD	3.7163	3.9455	0.4375
cat	9.5984	1.9852	0.0081
Platelets	0.8911	0.9906	0.3420
GLOBAL	17.5943	18.4237	0.5113

### DF betas

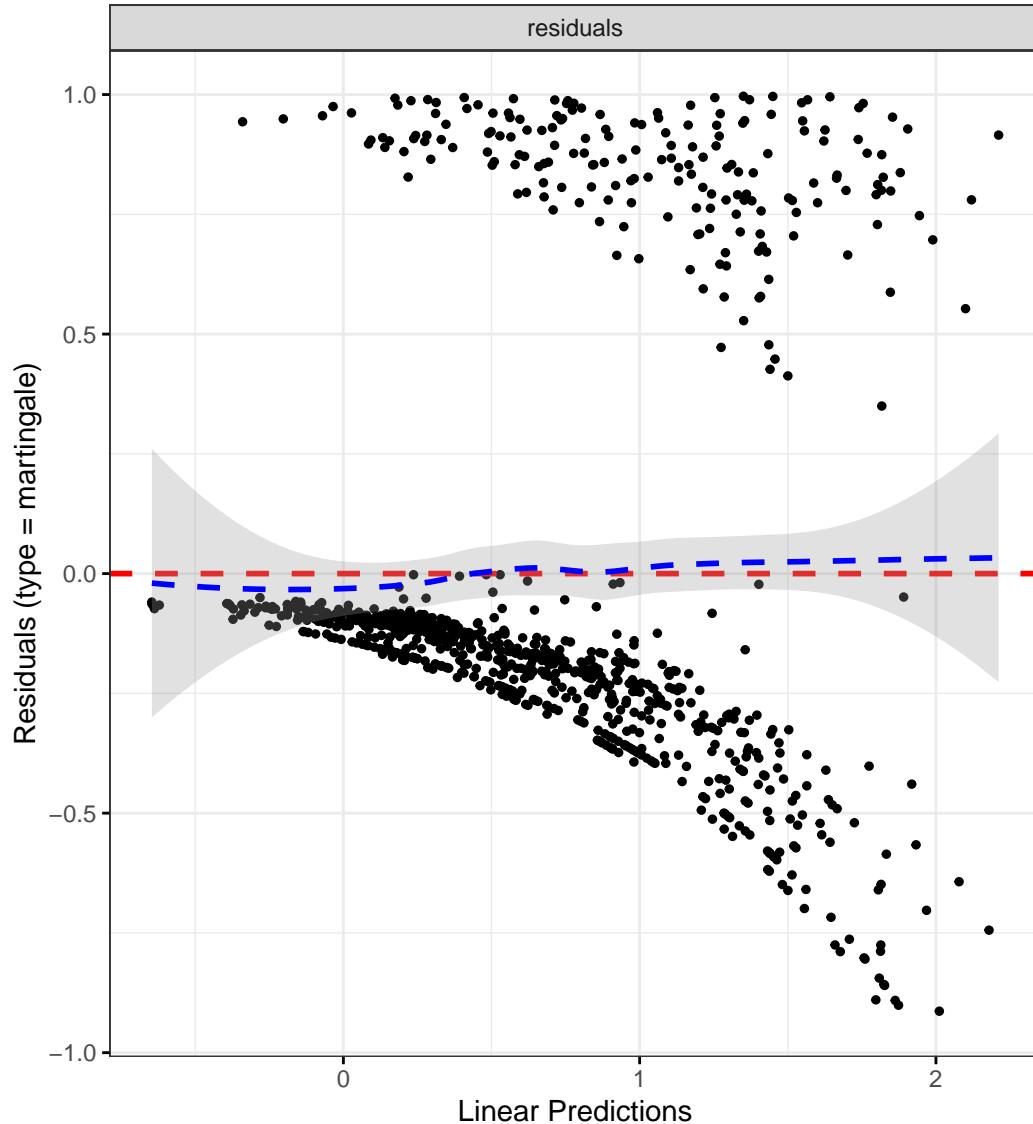
```
`geom_smooth()` using formula = 'y ~ x'
```



#####

Martingale residuals

```
`geom_smooth()` using formula = 'y ~ x'
```



### Ulcerative colitis

#### Patient-reported flare

```
# Categorize Platelets by quantiles
flare.uc.df <- categorize_by_quantiles(flare.uc.df, "Platelets", reference_data = flare.df)

# Run survival analysis using utility function
analysis_result <- run_survival_analysis()
```

```

data = flare.uc.df,
var_name = "Platelets",
outcome_time = "softflare_time",
outcome_event = "softflare",
legend_title = "Platelets",
plot_base_path = "plots/uc/soft-flare/biochem/Platelets",
break_time_by = 200
)

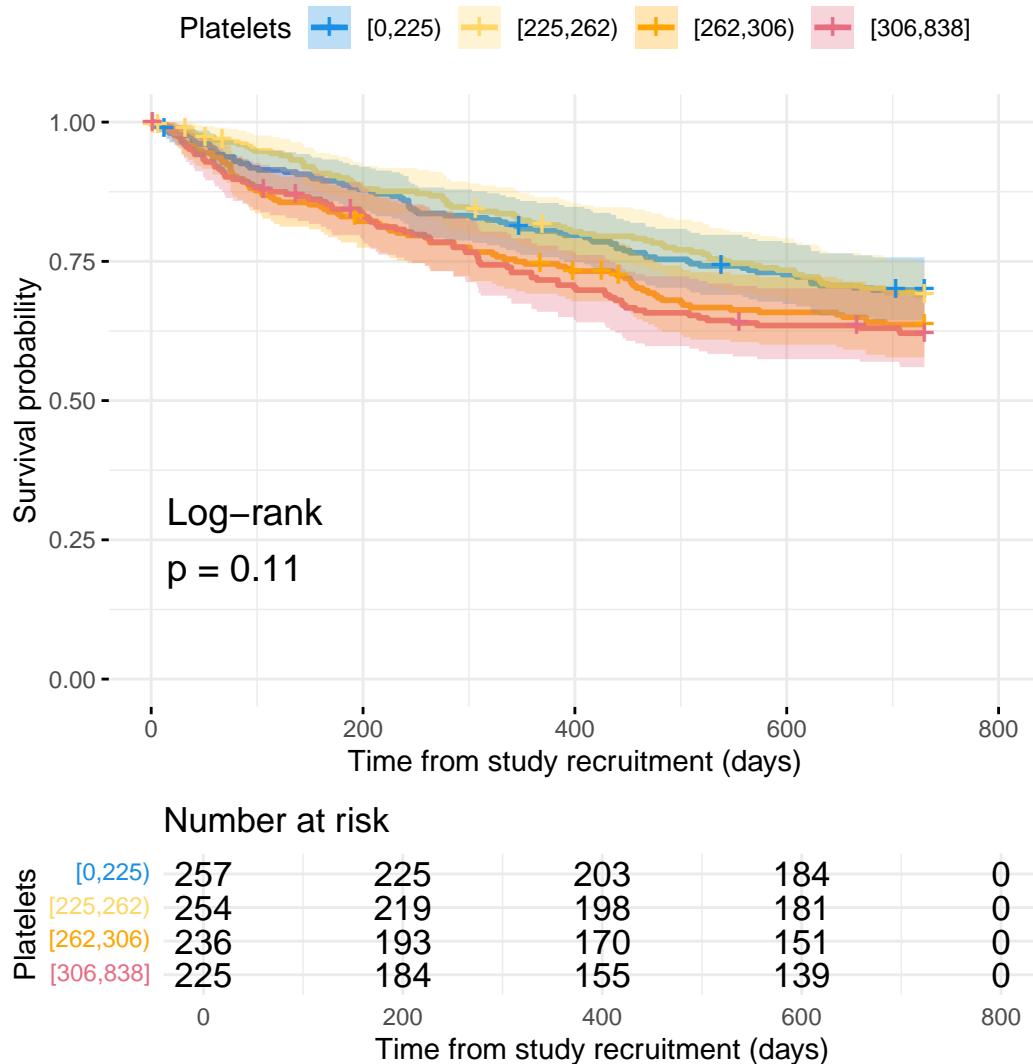
# Extract hazard ratio for continuous Platelets variable
fit.me <- coxph(
  Surv(softflare_time, softflare) ~
    Sex + IMD + cat + Platelets + frailty(SiteNo),
  control = coxph.control(outer.max = 20),
  data = flare.uc.df
)

uc.clin.forest <- rbind(
  uc.clin.forest,
  get_HR(fit.me, "Platelets")
)

# Display plot and model summary
analysis_result$plot

```

### Time to patient-reported flare



```
invisible(cox_summary(fit.me))
```

Cox model summary:

Variable	HR	Lower 95%	Upper 95%	P-value
SexFemale	1.6359	1.2815	2.0883	0.0001
IMD2	1.2115	0.7423	1.9773	0.4427
IMD3	0.9195	0.5589	1.5125	0.7409

Variable	HR	Lower 95%	Upper 95%	P-value
IMD4	1.1625	0.7274	1.8577	0.5291
IMD5	1.1133	0.7074	1.7520	0.6428
catFC 50-250	1.7042	1.2898	2.2517	0.0002
catFC > 250	2.3417	1.7414	3.1489	0.0000
Platelets	0.9996	0.9978	1.0014	0.6470

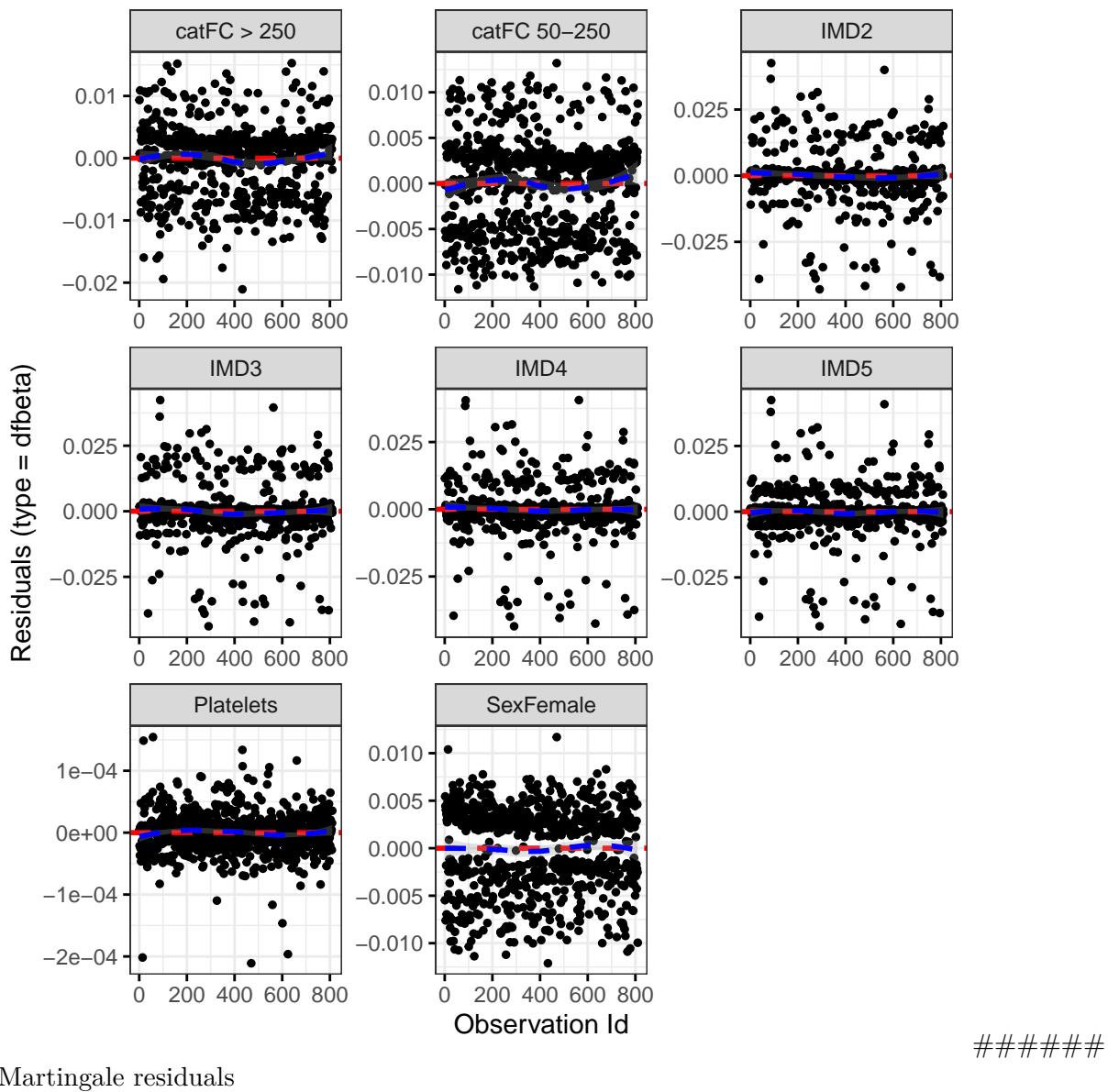
Diagnostics:

### Proportional hazards assumption test

	Chi-squared statistic	DF	P-value
Sex	0.9266	0.9866	0.3313
IMD	3.6484	3.9255	0.4443
cat	5.1798	1.9617	0.0724
Platelets	1.2273	0.9820	0.2627
GLOBAL	10.6943	21.6429	0.9755

### DF betas

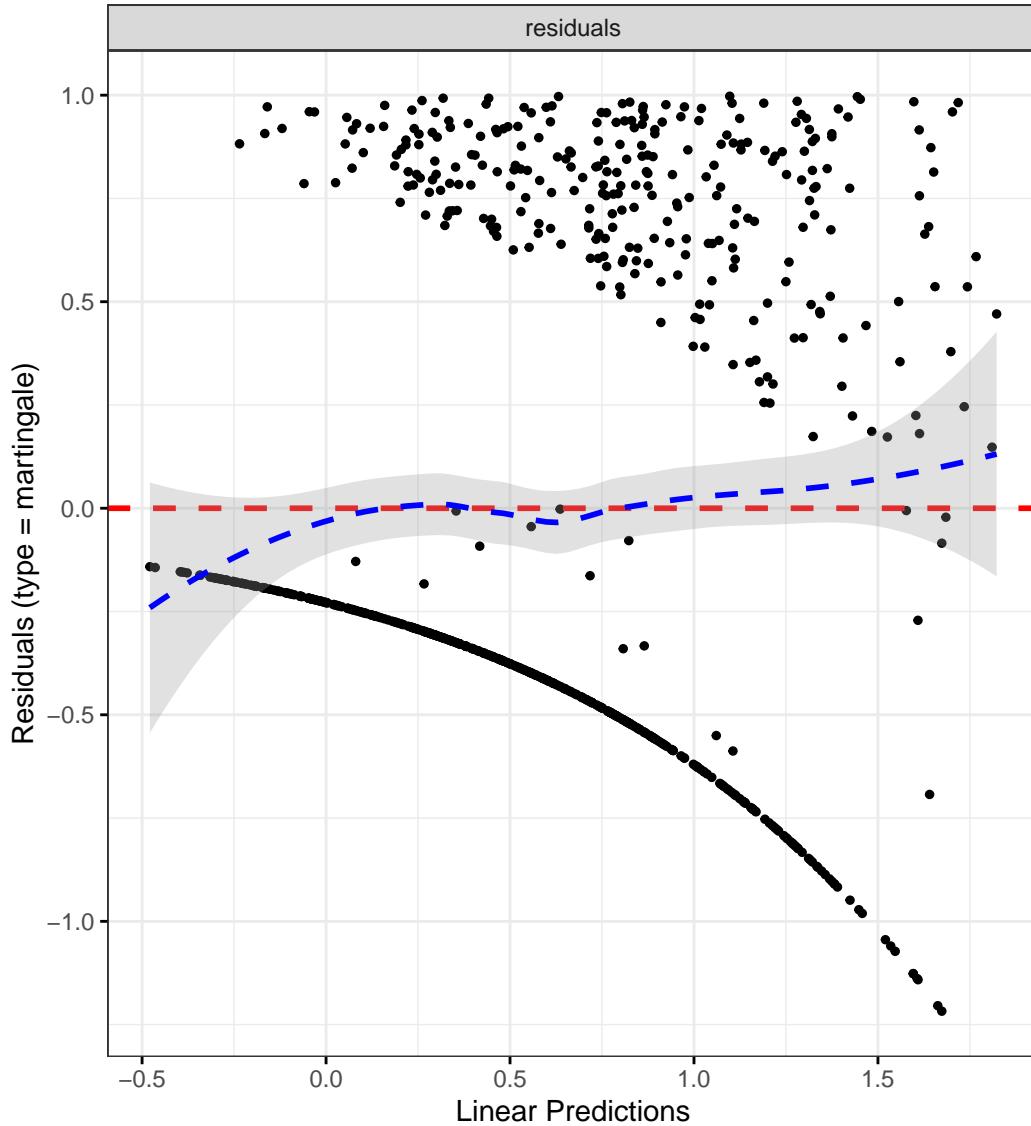
```
`geom_smooth()` using formula = 'y ~ x'
```



#####

Martingale residuals

```
`geom_smooth()` using formula = 'y ~ x'
```



### Hard flare

```
# Generate survival plot and run Cox model for hard flare using utility function
analysis_result <- run_survival_analysis(
  data = flare.uc.df,
  var_name = "Platelets",
  outcome_time = "hardflare_time",
  outcome_event = "hardflare",
  legend_title = "Platelets",
```

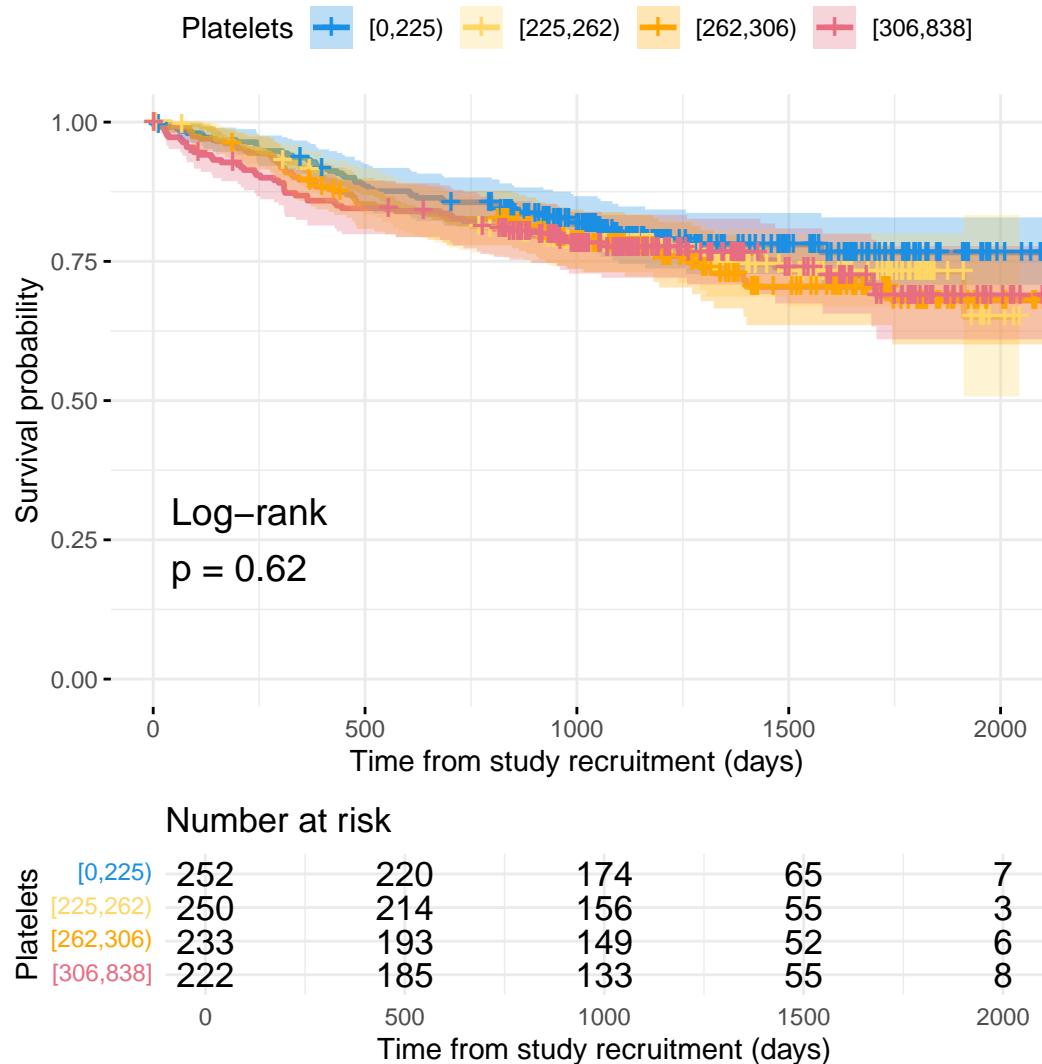
```
plot_base_path = "plots/uc/hard-flare/biochem/Platelets",
break_time_by = 500
)

# Extract hazard ratio for continuous Platelets variable
fit.me <- coxph(
  Surv(hardflare_time, hardflare) ~
    Sex + IMD + cat + Platelets + frailty(SiteNo),
  control = coxph.control(outer.max = 20),
  data = flare.uc.df
)

uc.hard.forest <- rbind(
  uc.hard.forest,
  get_HR(fit.me, "Platelets")
)

# Display plot and model summary
analysis_result$plot
```

### Time to hard flare



```
invisible(cox_summary(fit.me))
```

Cox model summary:

Variable	HR	Lower 95%	Upper 95%	P-value
SexFemale	1.3852	1.0354	1.8533	0.0282
IMD2	1.3635	0.7344	2.5316	0.3260
IMD3	1.0820	0.5795	2.0204	0.8045

Variable	HR	Lower 95%	Upper 95%	P-value
IMD4	1.4818	0.8202	2.6771	0.1926
IMD5	1.2584	0.7059	2.2433	0.4359
catFC 50-250	1.9349	1.3751	2.7228	0.0002
catFC > 250	3.0046	2.1079	4.2827	0.0000
Platelets	1.0001	0.9980	1.0022	0.9407

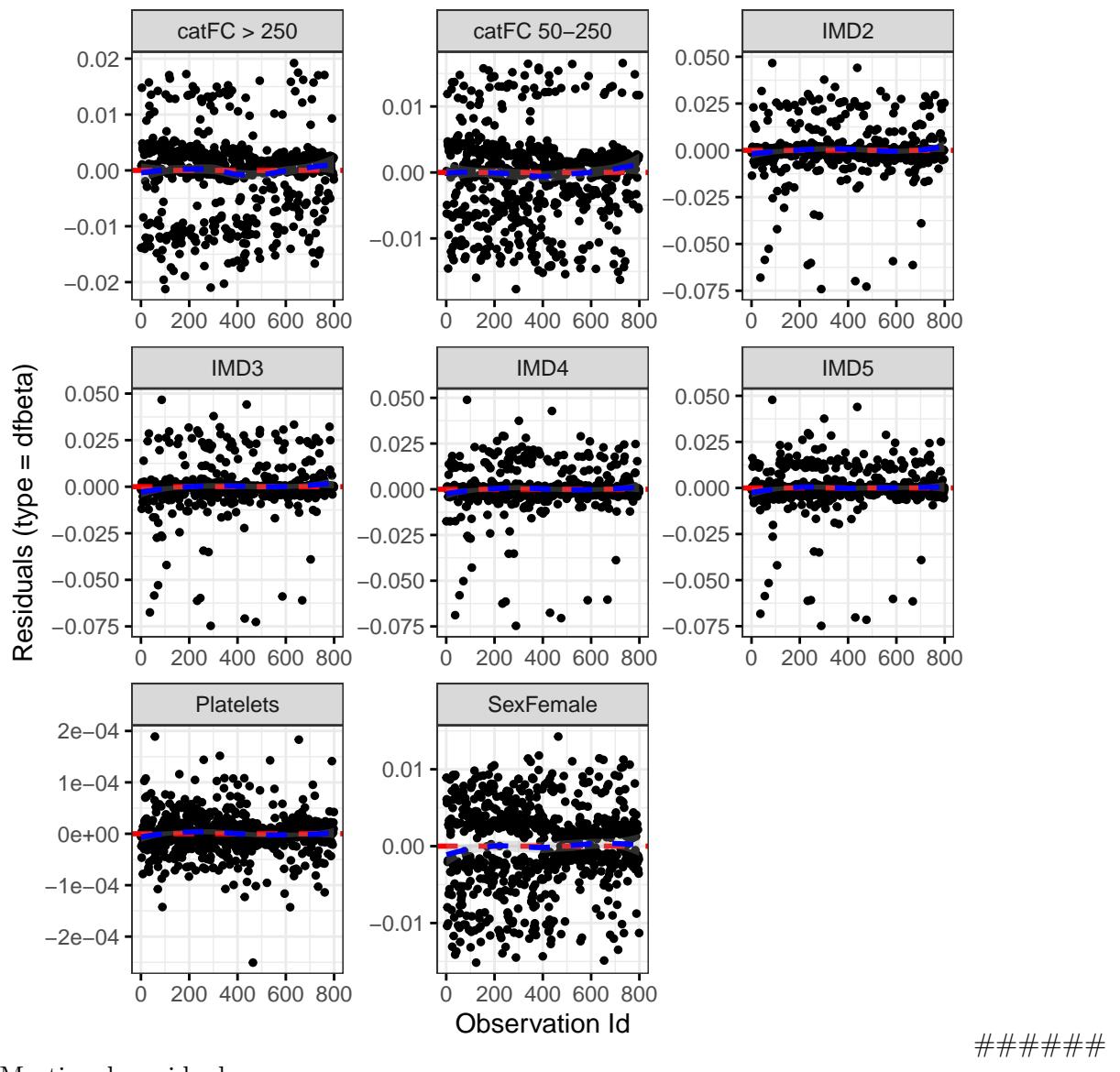
Diagnostics:

### Proportional hazards assumption test

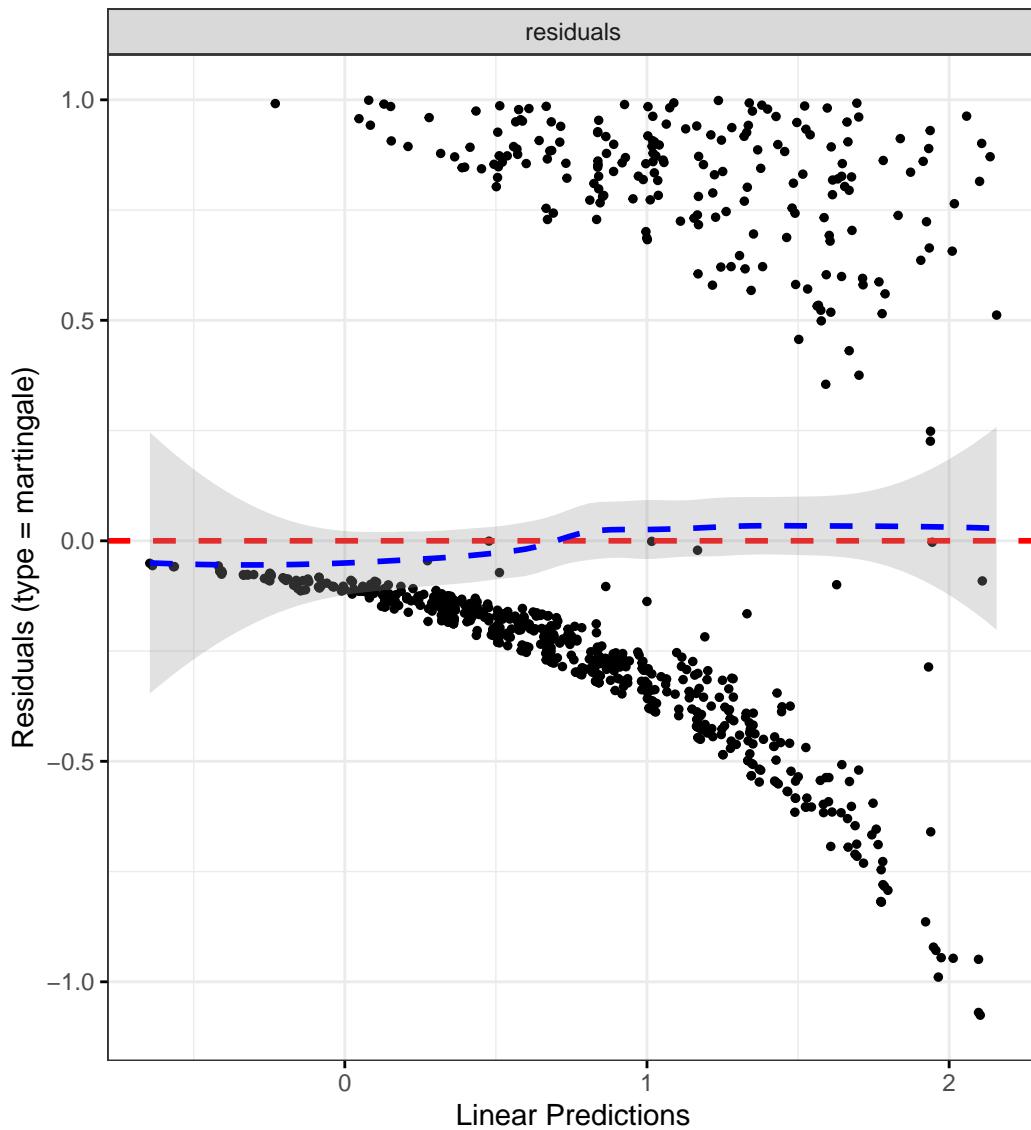
	Chi-squared statistic	DF	P-value
Sex	0.0330	0.9835	0.8508
IMD	1.8166	3.9337	0.7607
cat	4.9938	1.9649	0.0798
Platelets	0.2339	0.9849	0.6224
GLOBAL	6.6236	21.4912	0.9991

### DF betas

```
`geom_smooth()` using formula = 'y ~ x'
```



```
`geom_smooth()` using formula = 'y ~ x'
```



**Albumin**

**Crohn's disease**

**Patient-reported flare**

```
# Categorize Albumin by quantiles
flare.cd.df <- categorize_by_quantiles(flare.cd.df, "Albumin", reference_data = flare.df)
```

```

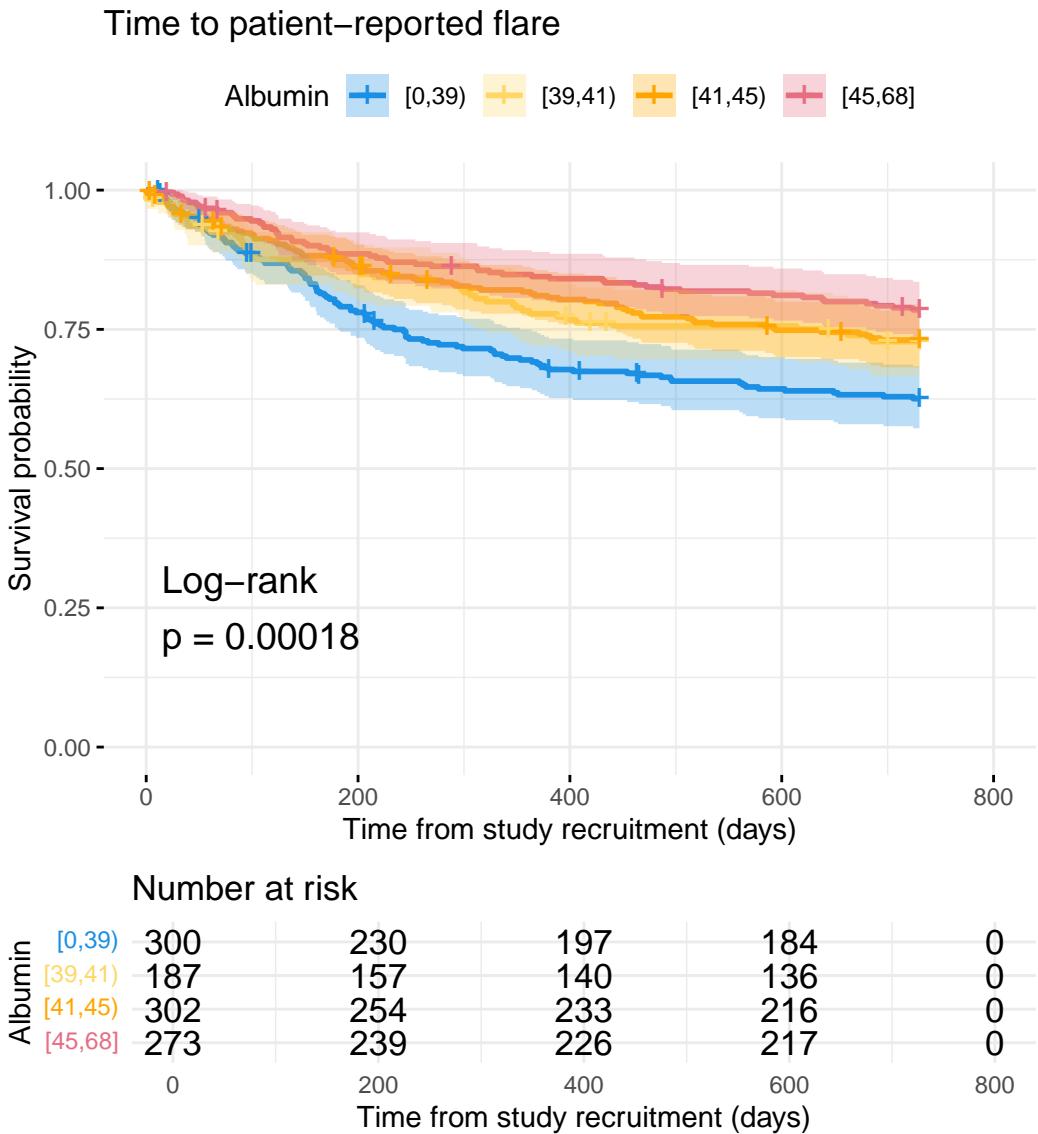
# Run survival analysis using utility function
analysis_result <- run_survival_analysis(
  data = flare.cd.df,
  var_name = "Albumin",
  outcome_time = "softflare_time",
  outcome_event = "softflare",
  legend_title = "Albumin",
  plot_base_path = "plots/cd/soft-flare/biochem/albumin",
  break_time_by = 200
)

# Extract hazard ratio for continuous Albumin variable
fit.me <- coxph(
  Surv(softflare_time, softflare) ~
    Sex + IMD + cat + Albumin + frailty(SiteNo),
  control = coxph.control(outer.max = 20),
  data = flare.cd.df
)

cd.clin.forest <- rbind(
  cd.clin.forest,
  get_HR(fit.me, "Albumin")
)

# Display plot and model summary
analysis_result$plot

```



```
invisible(cox_summary(fit.me))
```

Cox model summary:

Variable	HR	Lower 95%	Upper 95%	P-value
SexFemale	2.2704	1.7358	2.9697	0.0000
IMD2	1.1975	0.7005	2.0474	0.5100
IMD3	1.1051	0.6436	1.8976	0.7170

Variable	HR	Lower 95%	Upper 95%	P-value
IMD4	1.1570	0.6824	1.9617	0.5882
IMD5	1.1152	0.6683	1.8609	0.6765
catFC 50-250	1.6516	1.2460	2.1892	0.0005
catFC > 250	2.4799	1.8009	3.4147	0.0000
Albumin	0.9835	0.9596	1.0081	0.1877

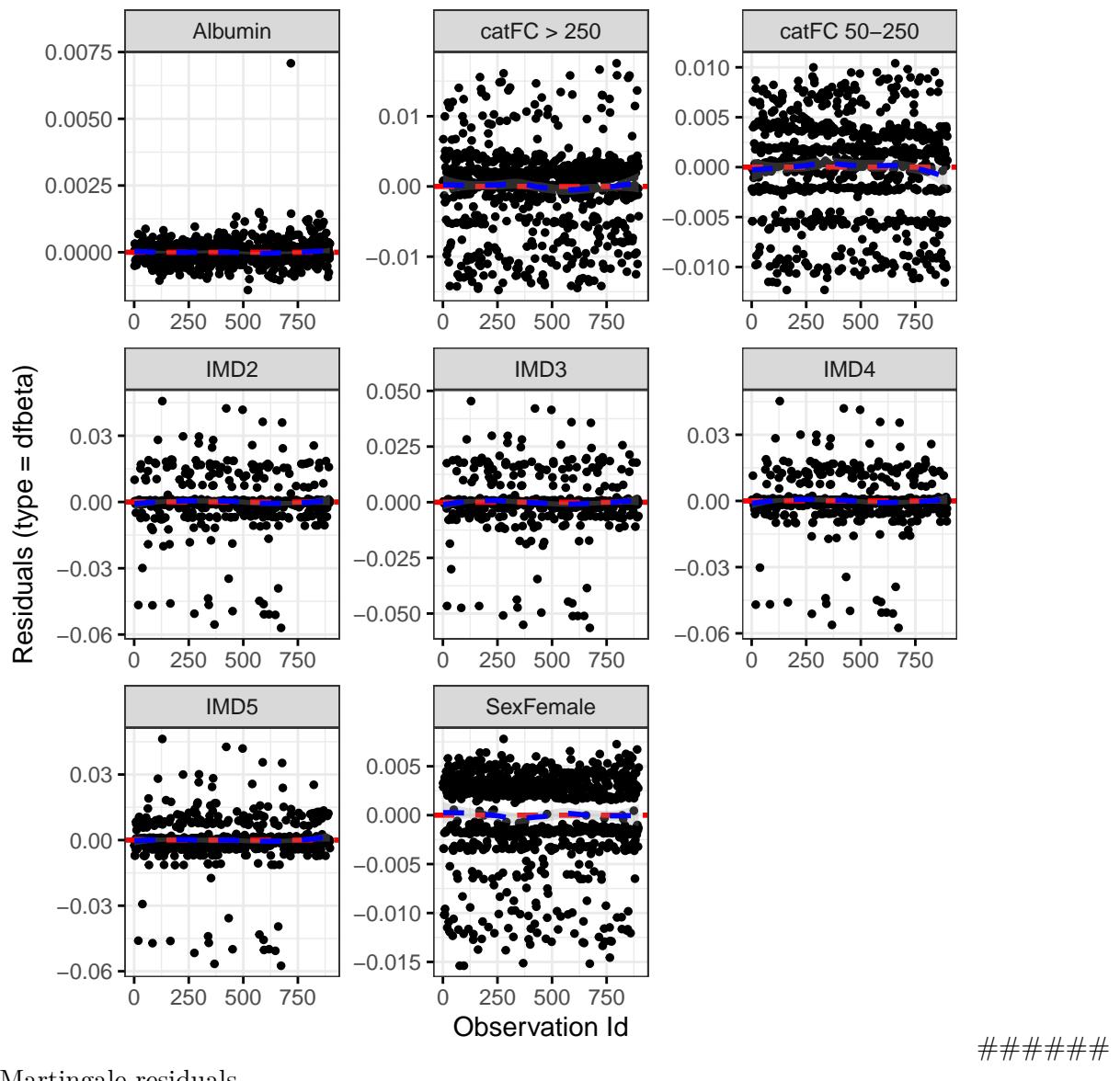
Diagnostics:

### Proportional hazards assumption test

	Chi-squared statistic	DF	P-value
Sex	0.5738	0.9985	0.4481
IMD	3.8357	3.9888	0.4270
cat	6.5059	1.9965	0.0385
Albumin	0.1594	0.9803	0.6816
GLOBAL	10.7187	8.8662	0.2846

### DF betas

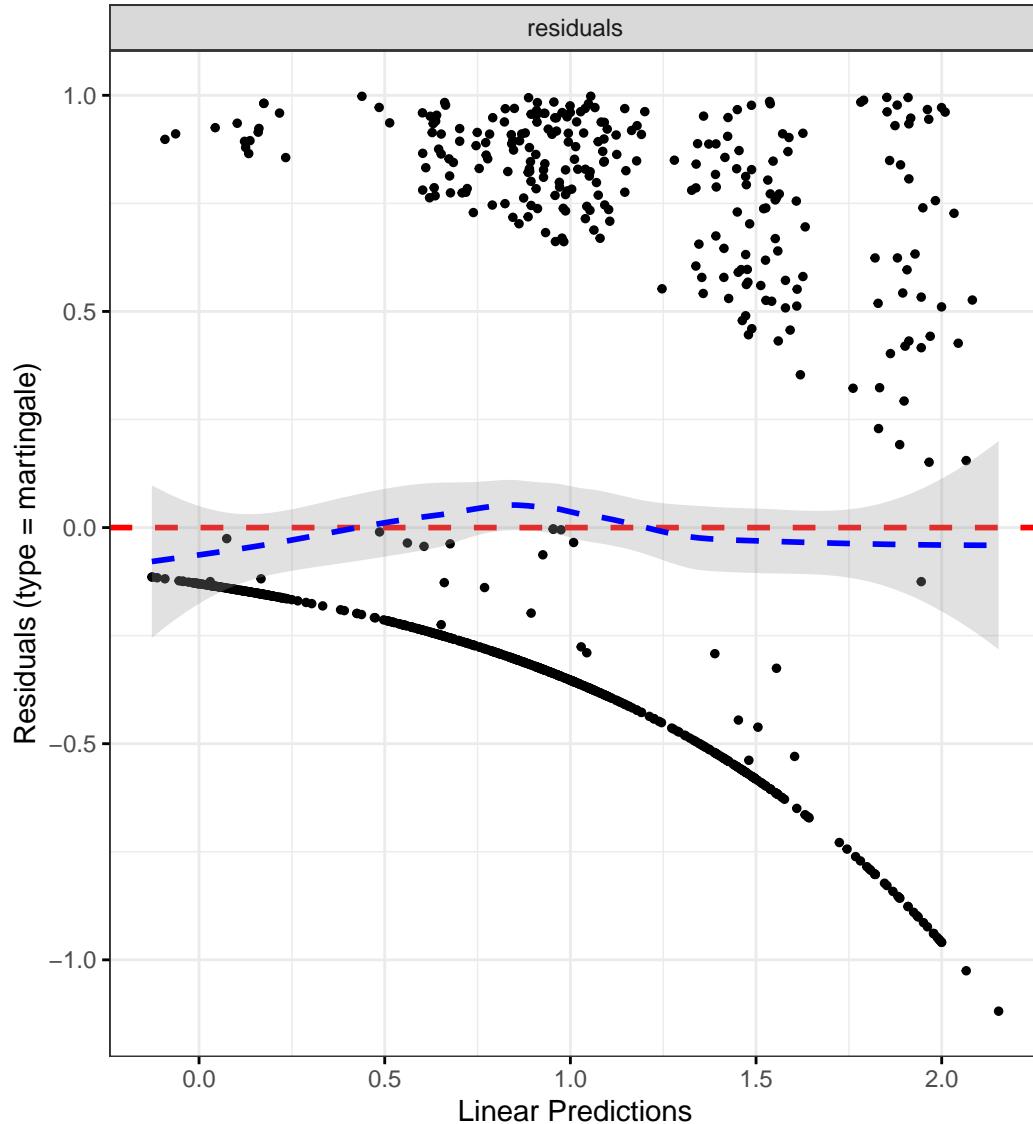
```
`geom_smooth()` using formula = 'y ~ x'
```



Martingale residuals

#####

```
`geom_smooth()` using formula = 'y ~ x'
```



### Hard flare

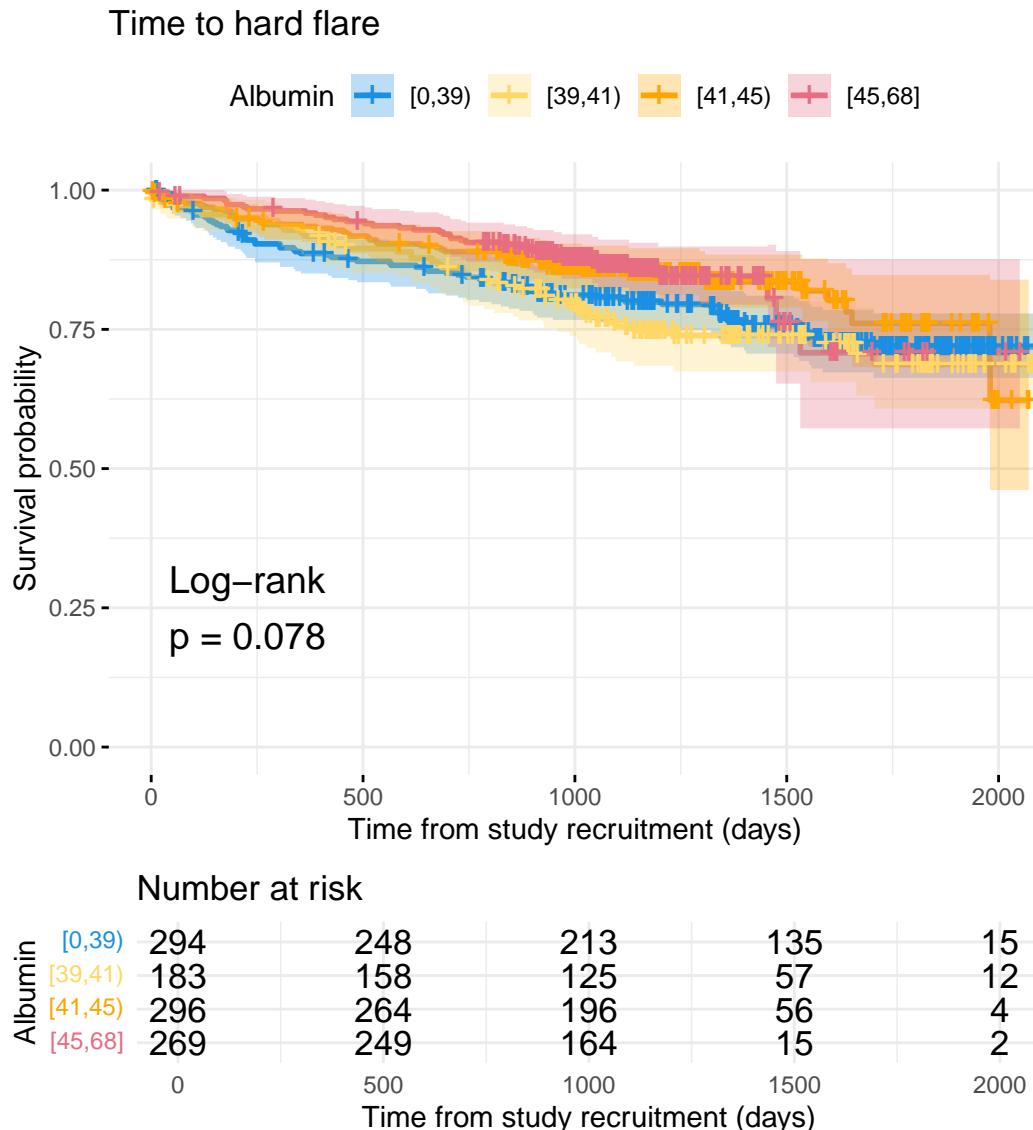
```
# Generate survival plot and run Cox model for hard flare using utility function
analysis_result <- run_survival_analysis(
  data = flare.cd.df,
  var_name = "Albumin",
  outcome_time = "hardflare_time",
  outcome_event = "hardflare",
  legend_title = "Albumin",
```

```
plot_base_path = "plots/cd/hard-flare/biochem/albumin",
break_time_by = 500
)

# Extract hazard ratio for continuous Albumin variable
fit.me <- coxph(
  Surv(hardflare_time, hardflare) ~
    Sex + IMD + cat + Albumin + frailty(SiteNo),
  control = coxph.control(outer.max = 20),
  data = flare.cd.df
)

cd.hard.forest <- rbind(
  cd.hard.forest,
  get_HR(fit.me, "Albumin")
)

# Display plot and model summary
analysis_result$plot
```



```
invisible(cox_summary(fit.me))
```

Cox model summary:

Variable	HR	Lower 95%	Upper 95%	P-value
SexFemale	1.4446	1.0621	1.9646	0.0191
IMD2	0.9964	0.5236	1.8963	0.9913
IMD3	1.0062	0.5244	1.9306	0.9852

Variable	HR	Lower 95%	Upper 95%	P-value
IMD4	0.9927	0.5269	1.8700	0.9818
IMD5	0.9065	0.4904	1.6756	0.7541
catFC 50-250	2.1095	1.4746	3.0179	0.0000
catFC > 250	3.2938	2.2145	4.8992	0.0000
Albumin	1.0071	0.9709	1.0446	0.7043

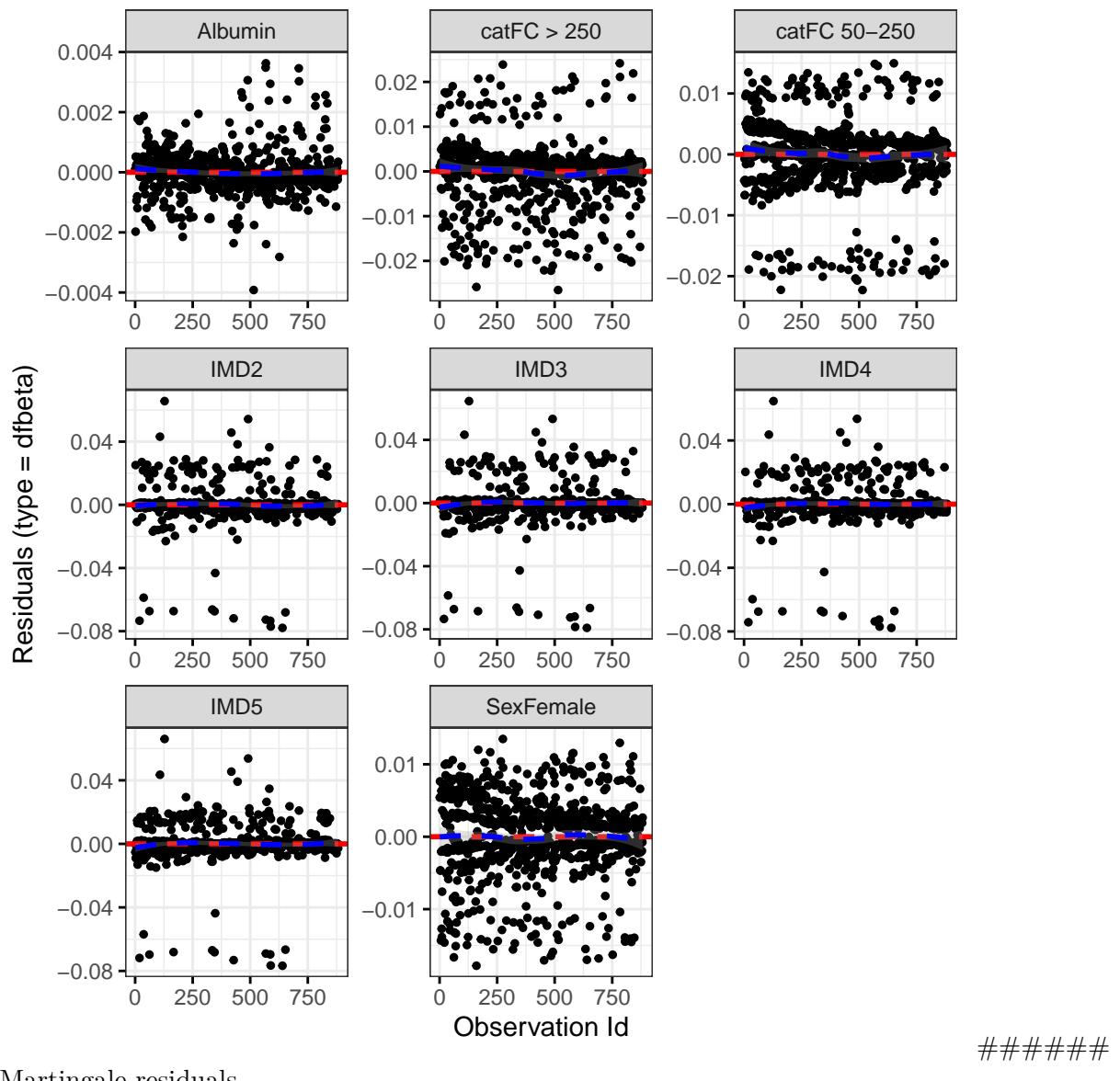
Diagnostics:

### Proportional hazards assumption test

	Chi-squared statistic	DF	P-value
Sex	2.5520	0.9880	0.1084
IMD	6.0553	3.9539	0.1906
cat	8.7501	1.9867	0.0124
Albumin	2.6129	0.9046	0.0928
GLOBAL	19.4962	16.0639	0.2473

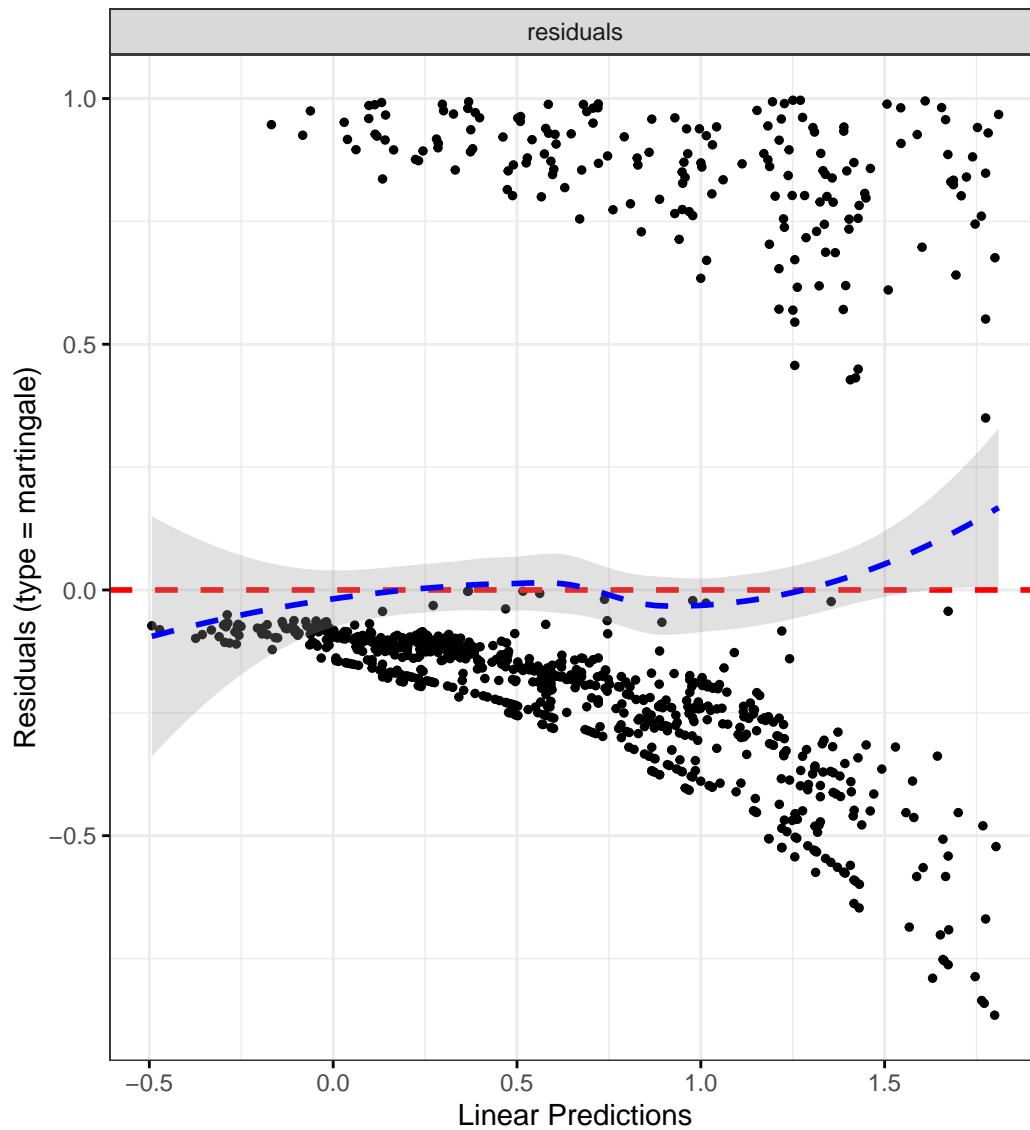
### DF betas

```
`geom_smooth()` using formula = 'y ~ x'
```



Martingale residuals

```
`geom_smooth()` using formula = 'y ~ x'
```



## Ulcerative colitis

### Patient-reported flare

```
# Categorize Albumin by quantiles
flare.uc.df <- categorize_by_quantiles(flare.uc.df, "Albumin", reference_data = flare.df)

# Run survival analysis using utility function
analysis_result <- run_survival_analysis()
```

```

data = flare.uc.df,
var_name = "Albumin",
outcome_time = "softflare_time",
outcome_event = "softflare",
legend_title = "Albumin",
plot_base_path = "plots/uc/soft-flare/biochem/albumin",
break_time_by = 200
)

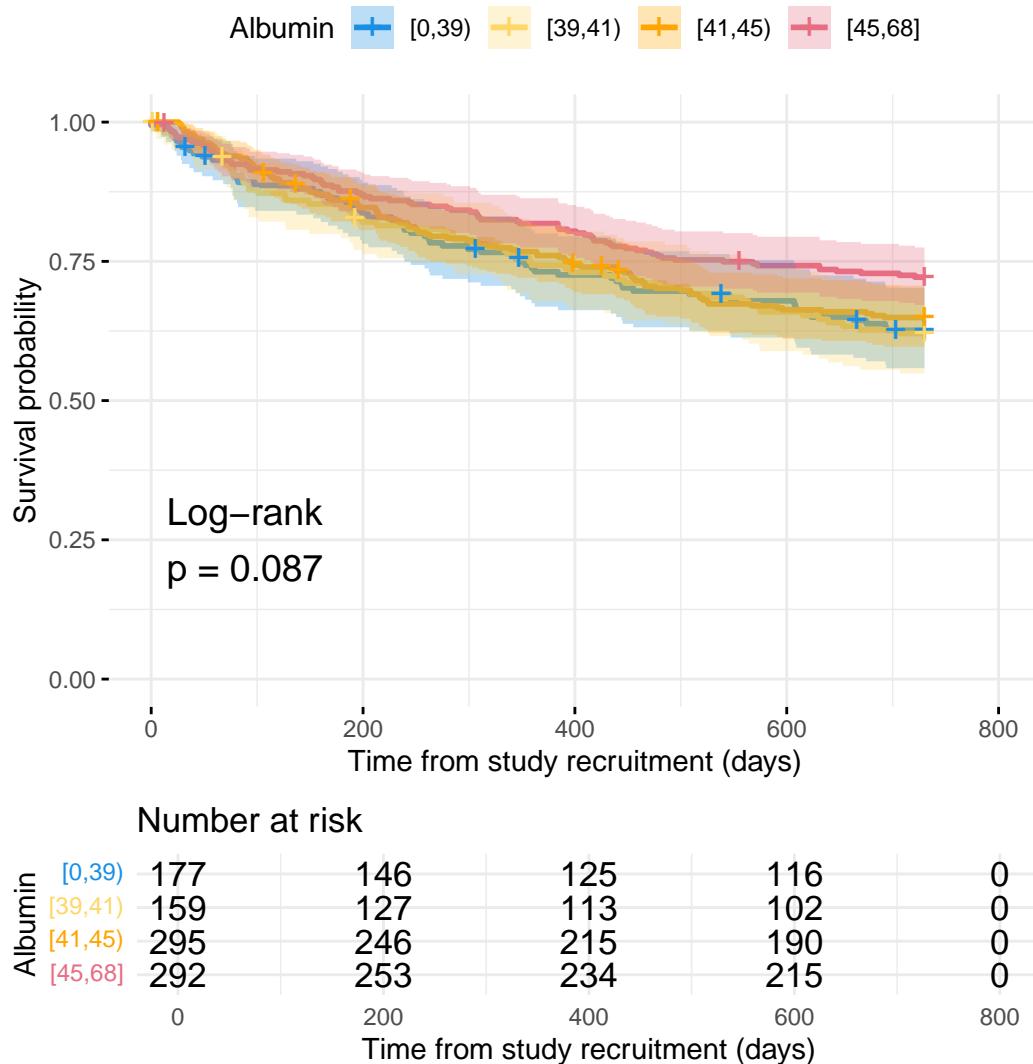
# Extract hazard ratio for continuous Albumin variable
fit.me <- coxph(
  Surv(softflare_time, softflare) ~
    Sex + IMD + cat + Albumin + frailty(SiteNo),
  control = coxph.control(outer.max = 20),
  data = flare.uc.df
)

uc.clin.forest <- rbind(
  uc.clin.forest,
  get_HR(fit.me, "Albumin")
)

# Display plot and model summary
analysis_result$plot

```

### Time to patient-reported flare



```
invisible(cox_summary(fit.me))
```

Cox model summary:

Variable	HR	Lower 95%	Upper 95%	P-value
SexFemale	1.5516	1.2175	1.9773	0.0004
IMD2	1.1036	0.6697	1.8185	0.6990
IMD3	0.9271	0.5653	1.5205	0.7644

Variable	HR	Lower 95%	Upper 95%	P-value
IMD4	1.1778	0.7403	1.8739	0.4897
IMD5	1.0943	0.6955	1.7218	0.6966
catFC 50-250	1.7423	1.3140	2.3103	0.0001
catFC > 250	2.1971	1.6288	2.9637	0.0000
Albumin	0.9775	0.9514	1.0043	0.0996

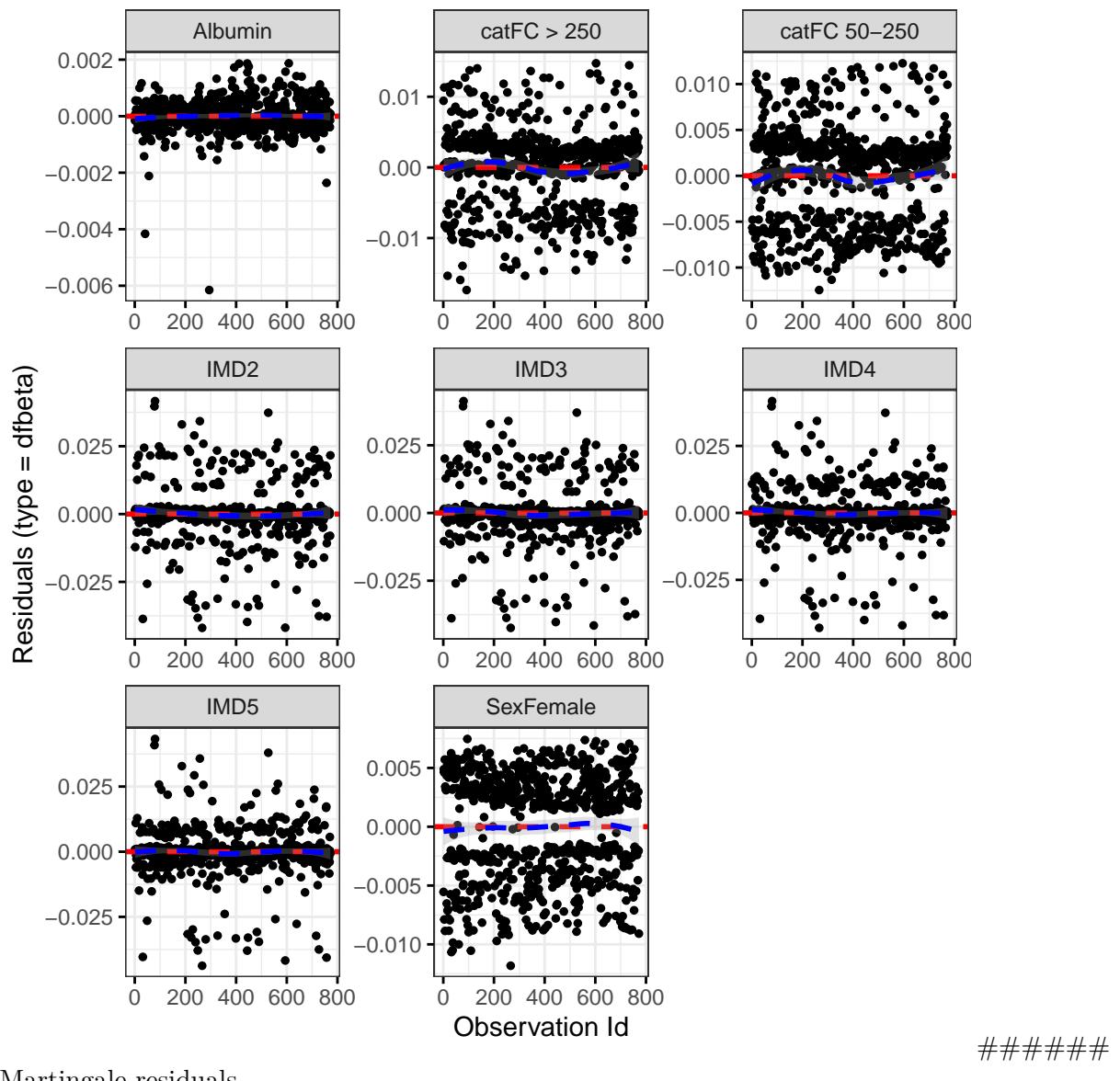
Diagnostics:

### Proportional hazards assumption test

	Chi-squared statistic	DF	P-value
Sex	0.9351	0.9883	0.3296
IMD	5.0458	3.9409	0.2754
cat	4.4764	1.9667	0.1036
Albumin	0.0124	0.8768	0.8787
GLOBAL	10.2249	18.1402	0.9279

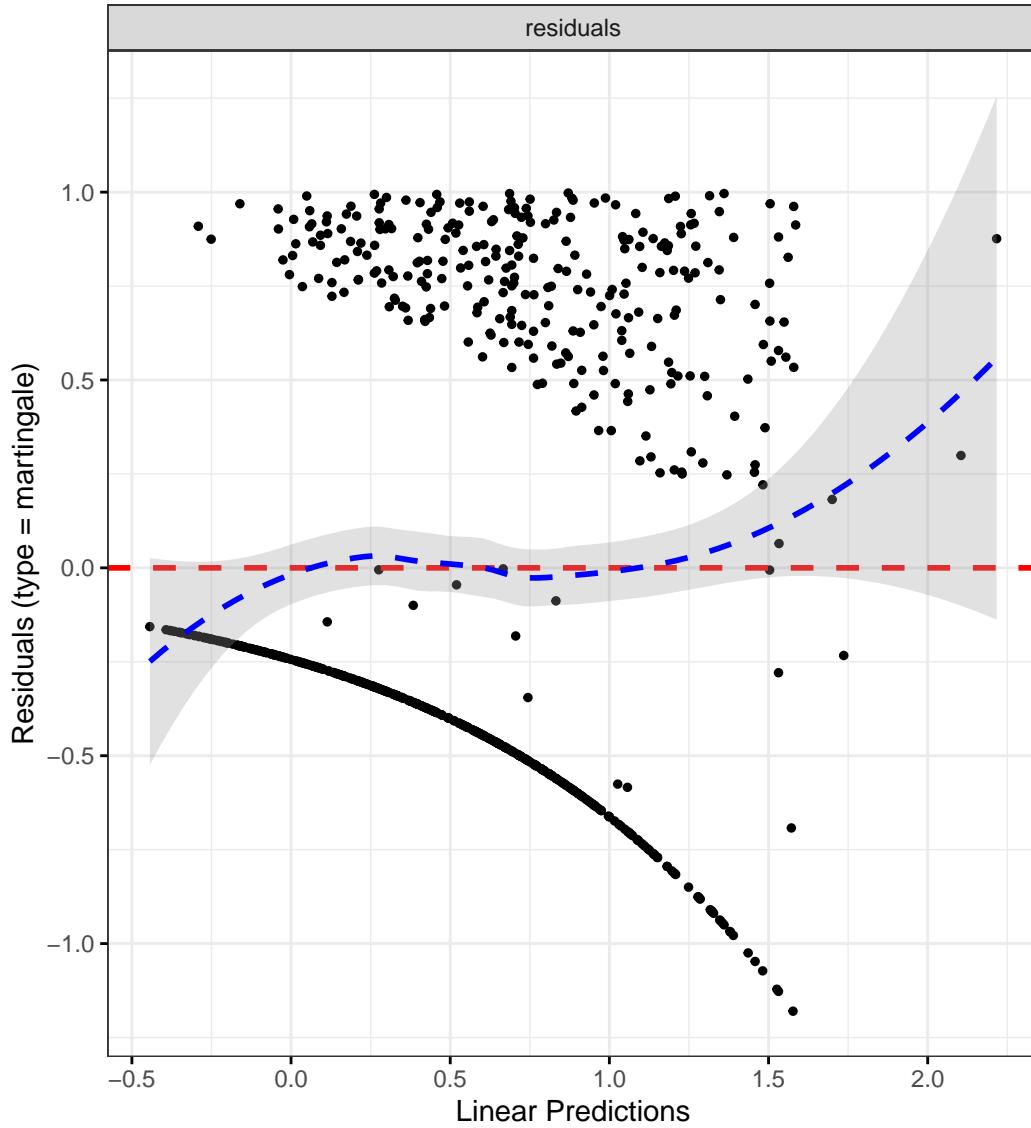
### DF betas

```
`geom_smooth()` using formula = 'y ~ x'
```



Martingale residuals

```
`geom_smooth()` using formula = 'y ~ x'
```



### Hard flare

```
# Generate survival plot and run Cox model for hard flare using utility function
analysis_result <- run_survival_analysis(
  data = flare.uc.df,
  var_name = "Albumin",
  outcome_time = "hardflare_time",
  outcome_event = "hardflare",
  legend_title = "Albumin",
```

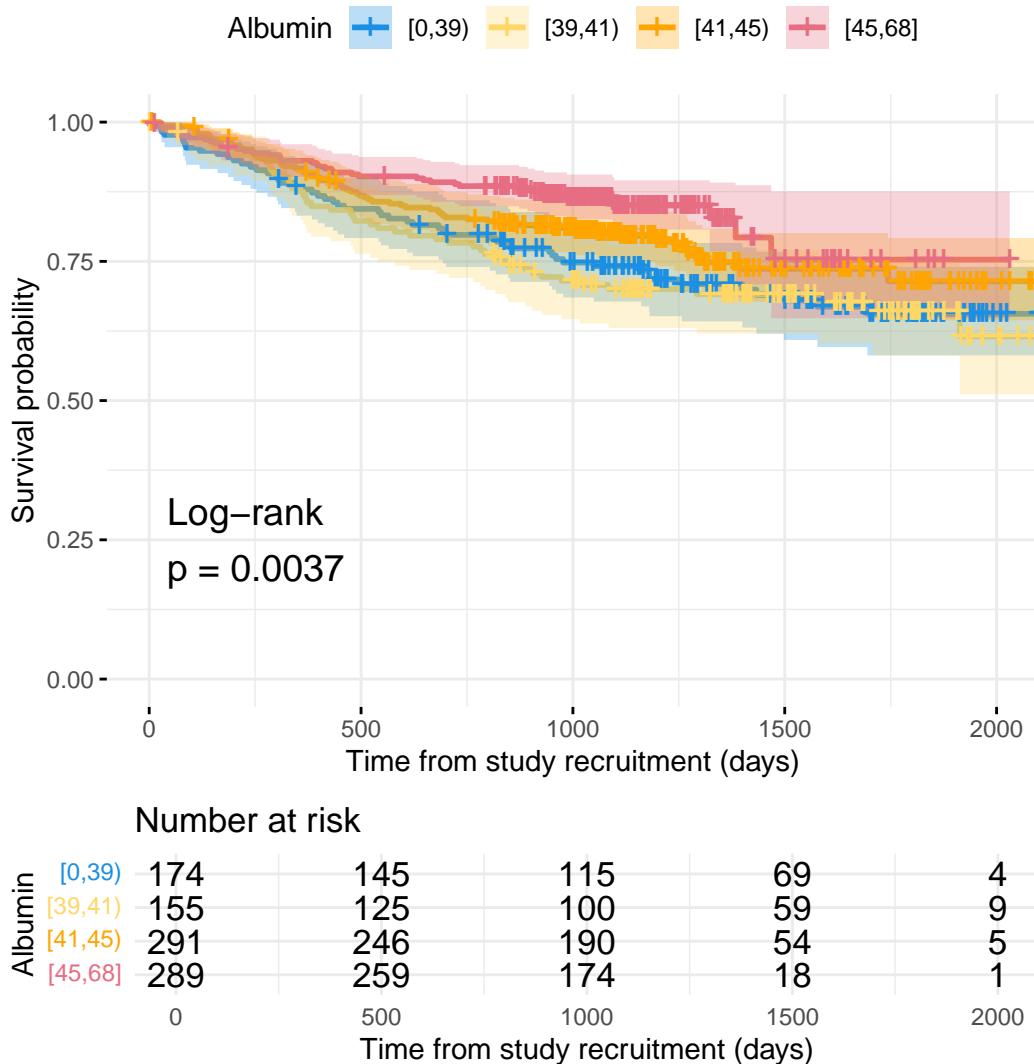
```
plot_base_path = "plots/uc/hard-flare/biochem/albumin",
break_time_by = 500
)

# Extract hazard ratio for continuous Albumin variable
fit.me <- coxph(
  Surv(hardflare_time, hardflare) ~
    Sex + IMD + cat + Albumin + frailty(SiteNo),
  control = coxph.control(outer.max = 20),
  data = flare.uc.df
)

uc.hard.forest <- rbind(
  uc.hard.forest,
  get_HR(fit.me, "Albumin")
)

# Display plot and model summary
analysis_result$plot
```

### Time to hard flare



```
invisible(cox_summary(fit.me))
```

Cox model summary:

Variable	HR	Lower 95%	Upper 95%	P-value
SexFemale	1.3243	0.9879	1.7754	0.0603
IMD2	1.2614	0.6707	2.3723	0.4711
IMD3	1.0544	0.5654	1.9662	0.8678

Variable	HR	Lower 95%	Upper 95%	P-value
IMD4	1.4427	0.7995	2.6033	0.2236
IMD5	1.2662	0.7108	2.2558	0.4230
catFC 50-250	1.9480	1.3707	2.7685	0.0002
catFC > 250	2.8956	2.0192	4.1525	0.0000
Albumin	0.9866	0.9573	1.0168	0.3798

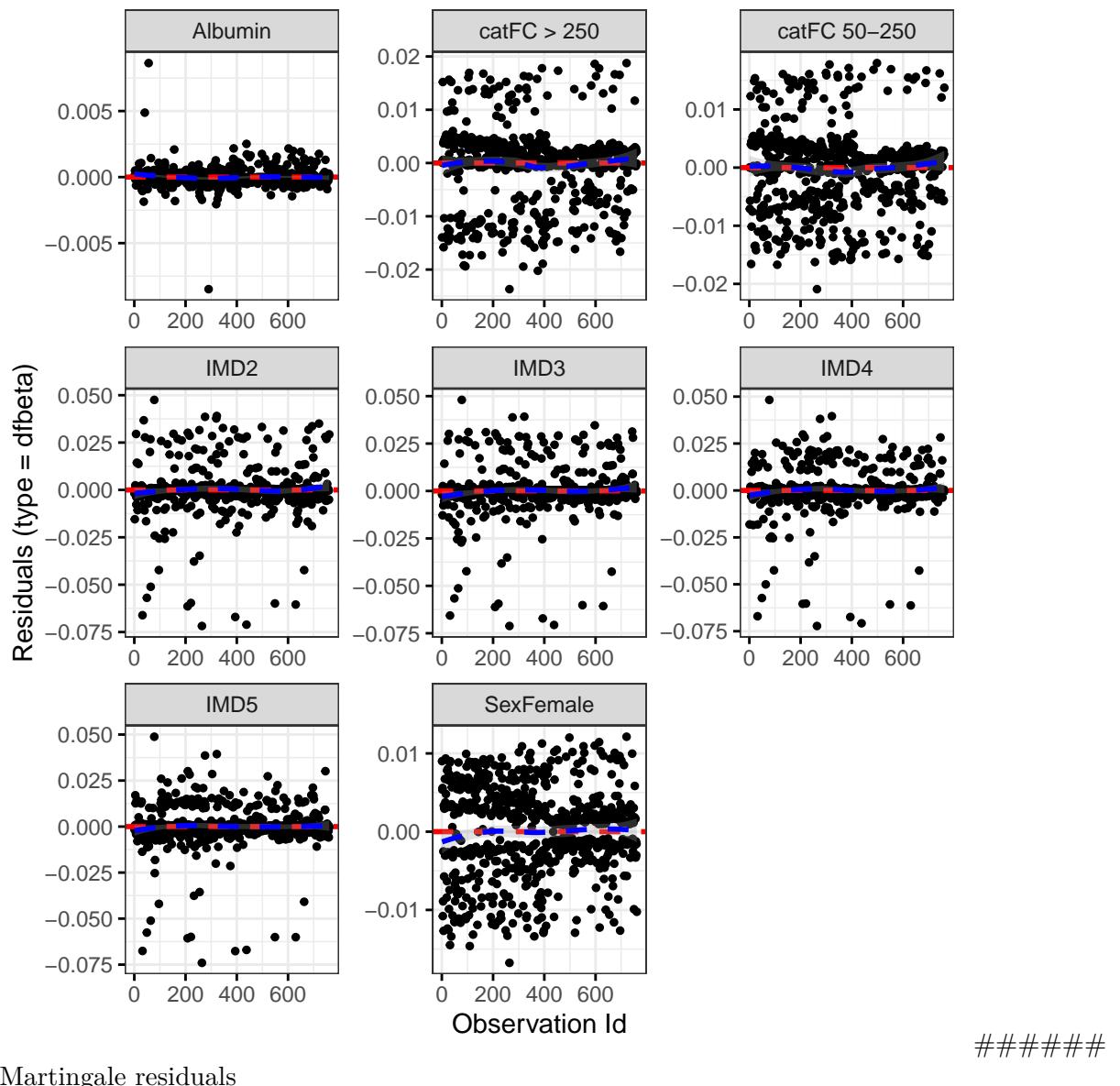
Diagnostics:

### Proportional hazards assumption test

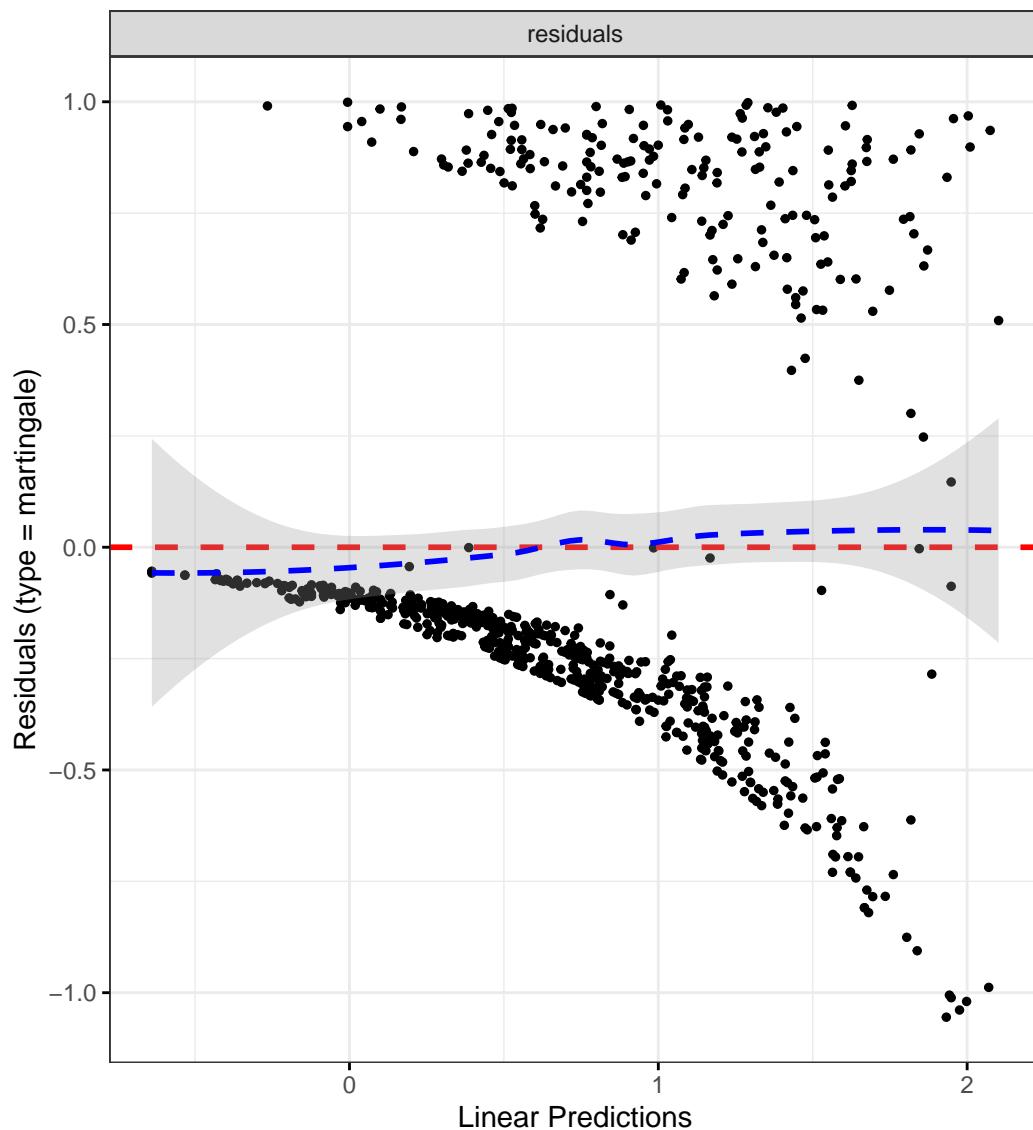
	Chi-squared statistic	DF	P-value
Sex	0.1590	0.9858	0.6843
IMD	1.5380	3.9411	0.8130
cat	4.2686	1.9686	0.1152
Albumin	0.3438	0.9039	0.5165
GLOBAL	6.1498	19.9849	0.9987

### DF betas

```
`geom_smooth()` using formula = 'y ~ x'
```



```
`geom_smooth()` using formula = 'y ~ x'
```



## Reproduction and reproducibility

Session info

R version 4.4.0 (2024-04-24)

Platform: aarch64-unknown-linux-gnu

locale: LC\_CTYPE=en\_US.UTF-8, LC\_NUMERIC=C, LC\_TIME=en\_US.UTF-8,  
LC\_COLLATE=en\_US.UTF-8, LC\_MONETARY=en\_US.UTF-8, LC\_MESSAGES=en\_US.UTF-

8,  $LC\_PAPER=en\_US.UTF-8$ ,  $LC\_NAME=C$ ,  $LC\_ADDRESS=C$ ,  $LC\_TELEPHONE=C$ ,  $LC\_MEASUREMENT=en\_US.UTF-8$  and  $LC\_IDENTIFICATION=C$

**attached base packages:** *stats, graphics, grDevices, utils, datasets, methods* and *base*

**other attached packages:** *gtsummary(v.1.7.2)*, *DescTools(v.0.99.54)*, *finalfit(v.1.0.7)*, *coxme(v.2.2-20)*, *bdsmatrix(v.1.3-7)*, *pander(v.0.6.5)*, *survminer(v.0.4.9)*, *ggpubr(v.0.6.0)*, *survival(v.3.5-8)*, *datefixR(v.1.6.1)*, *lubridate(v.1.9.3)*, *forcats(v.1.0.0)*, *stringr(v.1.5.1)*, *dplyr(v.1.1.4)*, *purrr(v.1.0.2)*, *readr(v.2.1.5)*, *tidyR(v.1.3.1)*, *tibble(v.3.2.1)*, *ggplot2(v.3.5.1)*, *tidyverse(v.2.0.0)* and *readxl(v.1.4.3)*

**loaded via a namespace (and not attached):** *gridExtra(v.2.3)*, *gld(v.2.6.6)*, *rlang(v.1.1.3)*, *magrittr(v.2.0.3)*, *e1071(v.1.7-14)*, *compiler(v.4.4.0)*, *mgcv(v.1.9-1)*, *vcvers(v.0.6.5)*, *pkgconfig(v.2.0.3)*, *shape(v.1.4.6.1)*, *fastmap(v.1.2.0)*, *backports(v.1.5.0)*, *labeling(v.0.4.3)*, *KMsurv(v.0.1-5)*, *utf8(v.1.2.4)*, *rmarkdown(v.2.27)*, *markdown(v.1.12)*, *tzdb(v.0.4.0)*, *nloptr(v.2.0.3)*, *xfun(v.0.44)*, *glmnet(v.4.1-8)*, *jomo(v.2.7-6)*, *jsonlite(v.1.8.8)*, *pan(v.1.9)*, *broom(v.1.0.6)*, *R6(v.2.5.1)*, *stringi(v.1.8.4)*, *car(v.3.1-2)*, *boot(v.1.3-30)*, *rpart(v.4.1.23)*, *cellranger(v.1.1.0)*, *Rcpp(v.1.0.12)*, *iterators(v.1.0.14)*, *knitr(v.1.47)*, *zoo(v.1.8-12)*, *Matrix(v.1.7-0)*, *splines(v.4.4.0)*, *nnet(v.7.3-19)*, *timechange(v.0.3.0)*, *tidyselect(v.1.2.1)*, *rstudioapi(v.0.16.0)*, *abind(v.1.4-5)*, *yaml(v.2.3.8)*, *ggtext(v.0.1.2)*, *codetools(v.0.2-20)*, *lattice(v.0.22-6)*, *withr(v.3.0.0)*, *evaluate(v.0.23)*, *proxy(v.0.4-27)*, *xml2(v.1.3.6)*, *survMisc(v.0.5.6)*, *pillar(v.1.9.0)*, *carData(v.3.0-5)*, *mice(v.3.16.0)*, *foreach(v.1.5.2)*, *generics(v.0.1.3)*, *hms(v.1.1.3)*, *commonmark(v.1.9.1)*, *munsell(v.0.5.1)*, *scales(v.1.3.0)*, *rootSolve(v.1.8.2.4)*, *minqa(v.1.2.7)*, *xtable(v.1.8-4)*, *class(v.7.3-22)*, *glue(v.1.7.0)*, *lmom(v.3.0)*, *tools(v.4.4.0)*, *data.table(v.1.15.4)*, *lme4(v.1.1-35.3)*, *ggsignif(v.0.6.4)*, *Exact(v.3.2)*, *mvtnorm(v.1.2-5)*, *grid(v.4.4.0)*, *colorspace(v.2.1-0)*, *nlme(v.3.1-164)*, *cli(v.3.6.2)*, *km.ci(v.0.5-6)*, *fansi(v.1.0.6)*, *expm(v.0.999-9)*, *broom.helpers(v.1.15.0)*, *gt(v.0.10.1)*, *gtable(v.0.3.5)*, *rstatix(v.0.7.2)*, *digest(v.0.6.35)*, *farver(v.2.1.2)*, *htmltools(v.0.5.8.1)*, *lifecycle(v.1.0.4)*, *httr(v.1.4.7)*, *mitml(v.0.4-5)*, *gridtext(v.0.1.5)* and *MASS(v.7.3-60.2)*

Licensed by CC BY unless otherwise stated.

```
saveRDS(flare.df, paste0(paths$outdir, "flares-biochem.RDS"))
saveRDS(flare.cd.df, paste0(paths$outdir, "flares-biochem-cd.RDS"))
saveRDS(flare.uc.df, paste0(paths$outdir, "flares-biochem-uc.RDS"))

saveRDS(cd.clin.forest, paste0(paths$outdir, "cd-clin-biochem.RDS"))
saveRDS(cd.hard.forest, paste0(paths$outdir, "cd-hard-biochem.RDS"))
saveRDS(uc.clin.forest, paste0(paths$outdir, "uc-clin-biochem.RDS"))
saveRDS(uc.hard.forest, paste0(paths$outdir, "uc-hard-biochem.RDS"))
```