

# Biochemistry

Nathan Constantine-Cooke

2025-11-05

## Table of contents

Introduction . . . . .	4
C-reactive protein . . . . .	5
Crohn's disease . . . . .	5
Patient-reported flare . . . . .	5
Proportional hazards assumption test . . . . .	6
DF betas . . . . .	6
Martingale residuals . . . . .	6
Objective flare . . . . .	6
Proportional hazards assumption test . . . . .	8
DF betas . . . . .	8
Martingale residuals . . . . .	8
Ulcerative colitis . . . . .	8
Patient-reported flare . . . . .	8
Proportional hazards assumption test . . . . .	9
DF betas . . . . .	10
Martingale residuals . . . . .	10
Objective flare . . . . .	10
Proportional hazards assumption test . . . . .	11
DF betas . . . . .	11
Martingale residuals . . . . .	11
Haemoglobin . . . . .	11
Crohn's disease . . . . .	11
Patient-reported flare . . . . .	11
Proportional hazards assumption test . . . . .	13
DF betas . . . . .	13
Martingale residuals . . . . .	13
Objective flare . . . . .	13
Proportional hazards assumption test . . . . .	14

DF betas . . . . .	15
Martingale residuals . . . . .	15
Ulcerative colitis . . . . .	15
Patient-reported flare . . . . .	15
Proportional hazards assumption test . . . . .	16
DF betas . . . . .	16
Martingale residuals . . . . .	16
Objective flare . . . . .	16
Proportional hazards assumption test . . . . .	18
DF betas . . . . .	18
Martingale residuals . . . . .	18
White cell count . . . . .	18
Crohn's disease . . . . .	18
Patient-reported flare . . . . .	18
Proportional hazards assumption test . . . . .	19
DF betas . . . . .	20
Martingale residuals . . . . .	20
Objective flare . . . . .	20
Proportional hazards assumption test . . . . .	21
DF betas . . . . .	21
Martingale residuals . . . . .	21
Ulcerative colitis . . . . .	21
Patient-reported flare . . . . .	21
Proportional hazards assumption test . . . . .	23
DF betas . . . . .	23
Martingale residuals . . . . .	23
Objective flare . . . . .	23
Proportional hazards assumption test . . . . .	24
DF betas . . . . .	25
Martingale residuals . . . . .	25
Platelets . . . . .	25
Crohn's disease . . . . .	25
Patient-reported flare . . . . .	25
Proportional hazards assumption test . . . . .	26
DF betas . . . . .	26
Martingale residuals . . . . .	27
Objective flare . . . . .	27
Proportional hazards assumption test . . . . .	28
DF betas . . . . .	28
Martingale residuals . . . . .	28
Ulcerative colitis . . . . .	28
Patient-reported flare . . . . .	28
Proportional hazards assumption test . . . . .	30

DF betas . . . . .	30
Martingale residuals . . . . .	30
Objective flare . . . . .	30
Proportional hazards assumption test . . . . .	31
DF betas . . . . .	31
Martingale residuals . . . . .	32
Albumin . . . . .	32
Crohn's disease . . . . .	32
Patient-reported flare . . . . .	32
Proportional hazards assumption test . . . . .	33
DF betas . . . . .	33
Martingale residuals . . . . .	33
Objective flare . . . . .	33
Proportional hazards assumption test . . . . .	35
DF betas . . . . .	35
Martingale residuals . . . . .	35
Ulcerative colitis . . . . .	35
Patient-reported flare . . . . .	35
Proportional hazards assumption test . . . . .	36
DF betas . . . . .	37
Martingale residuals . . . . .	37
Objective flare . . . . .	37
Proportional hazards assumption test . . . . .	38
DF betas . . . . .	38
Martingale residuals . . . . .	38
Proportional hazards assumption test . . . . .	39
DF betas . . . . .	40
Proportional hazards assumption test . . . . .	43
DF betas . . . . .	44
Proportional hazards assumption test . . . . .	46
DF betas . . . . .	47
Proportional hazards assumption test . . . . .	50
DF betas . . . . .	51
Proportional hazards assumption test . . . . .	53
DF betas . . . . .	54
Proportional hazards assumption test . . . . .	57
DF betas . . . . .	58
Proportional hazards assumption test . . . . .	61
DF betas . . . . .	62
Proportional hazards assumption test . . . . .	64
DF betas . . . . .	65
Proportional hazards assumption test . . . . .	68
DF betas . . . . .	69

Proportional hazards assumption test	71
DF betas	72
Proportional hazards assumption test	75
DF betas	76
Proportional hazards assumption test	79
DF betas	80
Proportional hazards assumption test	82
DF betas	83
Proportional hazards assumption test	86
DF betas	87
Proportional hazards assumption test	89
DF betas	90
Proportional hazards assumption test	93
DF betas	94
Proportional hazards assumption test	97
DF betas	98
Proportional hazards assumption test	100
DF betas	101
Proportional hazards assumption test	104
DF betas	105
Proportional hazards assumption test	107
DF betas	108
Reproduction and reproducibility	110

## 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
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
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'

### Objective flare

```

# Generate survival plot and run Cox model for objective 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
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'
```

### Martingale residuals

```
`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, 10.0000001, 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
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'
```

## Martingale residuals

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

## Objective flare

```
# Generate survival plot and run Cox model for objective 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
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
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
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
IMD4	1.1227	0.6698	1.8817	0.6605
IMD5	1.0728	0.6506	1.7690	0.7830

Variable	HR	Lower 95%	Upper 95%	P-value
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'
```

### Objective flare

```
# Generate survival plot and run Cox model for objective 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
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
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'
```

## Martingale residuals

```
`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
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
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'
```

### Objective flare

```

# Generate survival plot and run Cox model for objective 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
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
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
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'
```

## Martingale residuals

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

## Objective flare

```
# Generate survival plot and run Cox model for objective 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
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
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
IMD4	1.1339	0.7087	1.8142	0.6002
IMD5	1.1158	0.7084	1.7576	0.6364

Variable	HR	Lower 95%	Upper 95%	P-value
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'
```

### Objective flare

```
# Generate survival plot and run Cox model for objective 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
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
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 objective 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
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'
```

## Objective flare

```
# Generate survival plot and run Cox model for objective 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
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
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
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'
```

## Objective flare

```
# Generate survival plot and run Cox model for objective 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
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
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'
```

## Martingale residuals

```
`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
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'
```

### Objective flare

```

# Generate survival plot and run Cox model for objective 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
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
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
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'
```

## Objective flare

```
# Generate survival plot and run Cox model for objective 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
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
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'
```

### Martingale residuals

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

Control for additional covariates

Crohn's disease

Patient-reported flare

C-reactive protein

```

fit.me <- coxph(
  Surv(softflare_time, softflare) ~
    Sex +
    IMD +
    cat +
    `IBD Duration` +
    BMI +
    Treatment +
    Age +
    CReactiveProtein +
    frailty(SiteNo),
  control = coxph.control(outer.max = 20),
  data = flare.cd.df
)

invisible(cox_summary(fit.me))

```

Cox model summary:

Variable	HR	Lower 95%	Upper 95%	P-value
SexFemale	2.6107	1.9418	3.5101	0.0000
IMD2	1.1133	0.6282	1.9729	0.7132
IMD3	1.1022	0.6125	1.9836	0.7454
IMD4	1.2829	0.7323	2.2475	0.3838
IMD5	1.1860	0.6870	2.0476	0.5403
catFC 50-250	1.4791	1.0947	1.9984	0.0108
catFC > 250	2.2698	1.6193	3.1817	0.0000
IBD Duration	0.9995	0.9870	1.0122	0.9378
BMI	1.0287	1.0040	1.0541	0.0224
TreatmentMono biologic	0.9692	0.6503	1.4446	0.8780
TreatmentCombo therapy	0.7105	0.4388	1.1504	0.1645
Treatment5-ASA	1.2912	0.6772	2.4616	0.4376
TreatmentNone reported	1.0152	0.6996	1.4732	0.9366
Age	1.0022	0.9922	1.0122	0.6722
CReactiveProtein	1.0008	0.9875	1.0142	0.9120

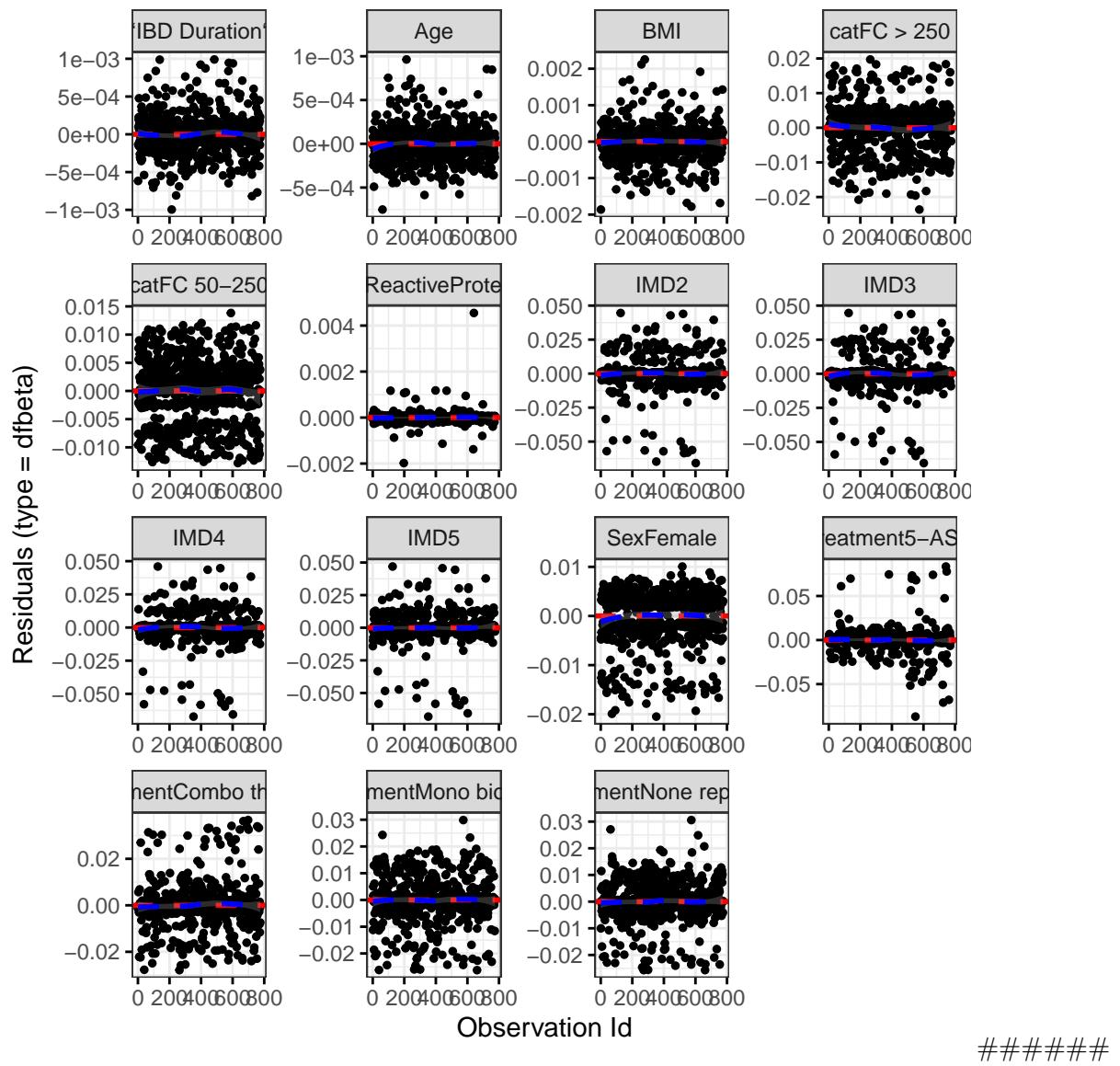
Diagnostics:

### Proportional hazards assumption test

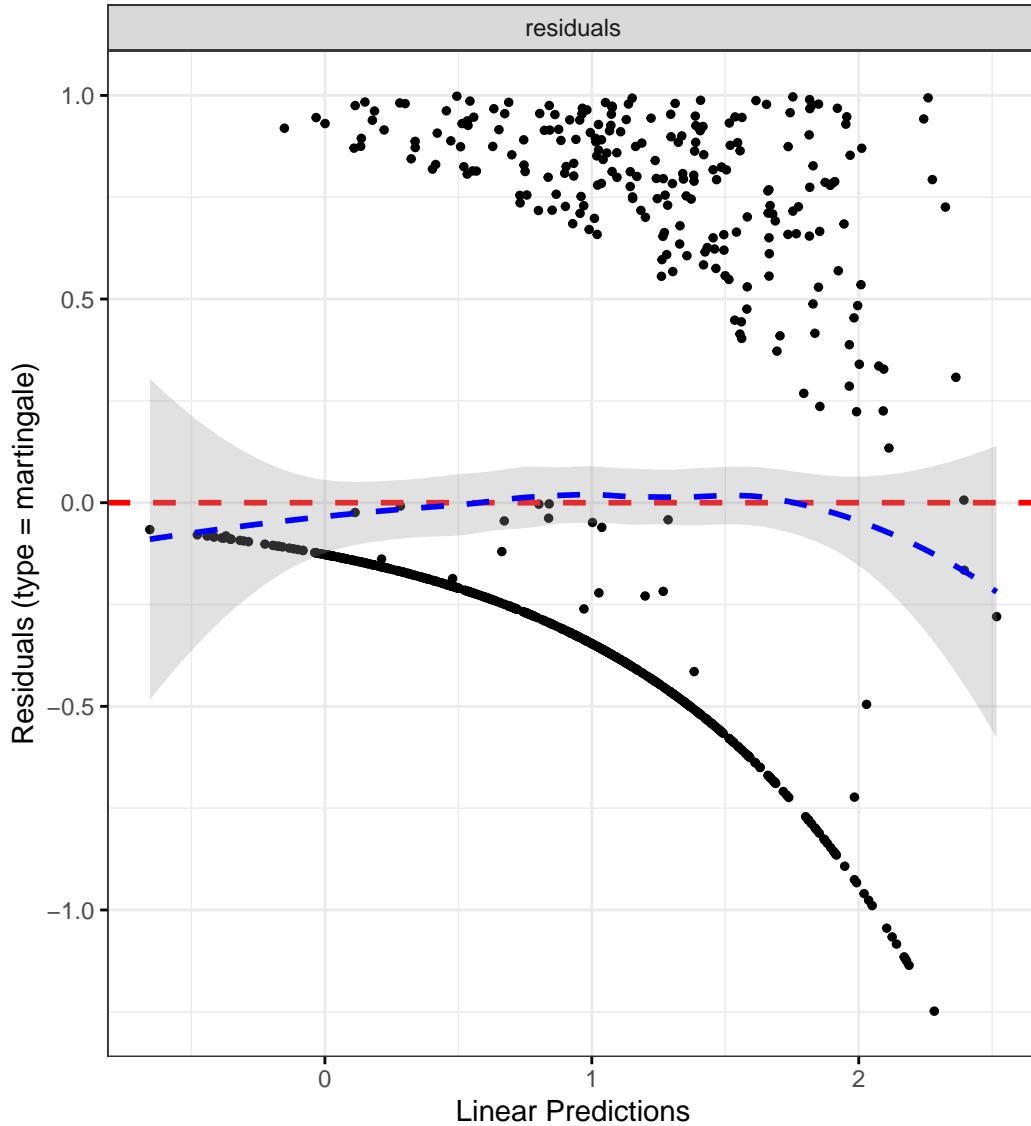
	Chi-squared statistic	DF	P-value
Sex	0.0866	0.9950	0.7667
IMD	2.7166	3.9677	0.6013
cat	1.8128	1.9874	0.4011
IBD Duration	3.5405	0.9973	0.0597
BMI	2.5239	0.9945	0.1113
Treatment	5.0810	3.9262	0.2701
Age	0.6056	0.9969	0.4353
CRactiveProtein	2.3676	0.9959	0.1232
GLOBAL	17.6149	17.3384	0.4364

## DF betas

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



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



Haemoglobin

```
fit.me <- coxph(
  Surv(softflare_time, softflare) ~
    Sex +
    IMD +
    cat +
    `IBD Duration` +
    BMI +
    Treatment +
```

```

Age +
Haemoglobin +
frailty(SiteNo),
control = coxph.control(outer.max = 20),
data = flare.cd.df
)

invisible(cox_summary(fit.me))

```

Cox model summary:

Variable	HR	Lower 95%	Upper 95%	P-value
SexFemale	2.1597	1.5666	2.9775	0.0000
IMD2	1.0259	0.5960	1.7659	0.9265
IMD3	1.0234	0.5856	1.7886	0.9351
IMD4	1.1447	0.6724	1.9488	0.6187
IMD5	1.0682	0.6380	1.7884	0.8019
catFC 50-250	1.5057	1.1301	2.0063	0.0052
catFC > 250	2.2631	1.6365	3.1298	0.0000
IBD Duration	0.9955	0.9835	1.0077	0.4718
BMI	1.0222	0.9978	1.0472	0.0755
TreatmentMono biologic	1.0985	0.7527	1.6032	0.6261
TreatmentCombo therapy	0.7624	0.4819	1.2060	0.2463
Treatment5-ASA	1.2865	0.6827	2.4243	0.4358
TreatmentNone reported	1.0489	0.7352	1.4964	0.7925
Age	1.0017	0.9922	1.0113	0.7258
Haemoglobin	0.9915	0.9814	1.0018	0.1044

Diagnostics:

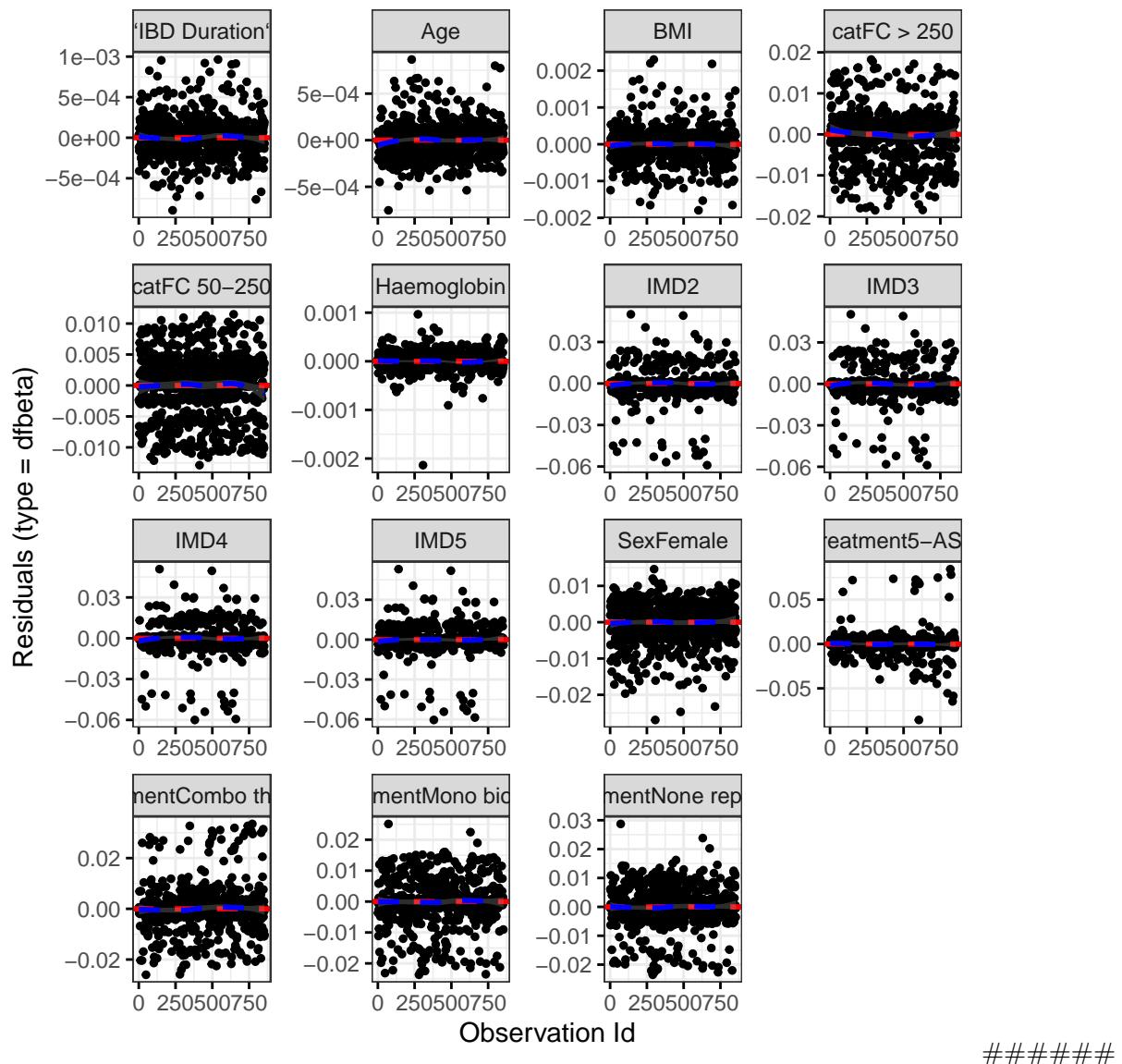
### Proportional hazards assumption test

	Chi-squared statistic	DF	P-value
Sex	0.7656	0.9922	0.3788
IMD	4.7397	3.9670	0.3107
cat	2.5886	1.9874	0.2718
IBD Duration	3.4103	0.9959	0.0644
BMI	3.9782	0.9949	0.0457
Treatment	3.9996	3.9264	0.3952

	Chi-squared statistic	DF	P-value
Age	0.7858	0.9948	0.3735
Haemoglobin	1.2319	0.9947	0.2655
GLOBAL	19.9851	17.9736	0.3321

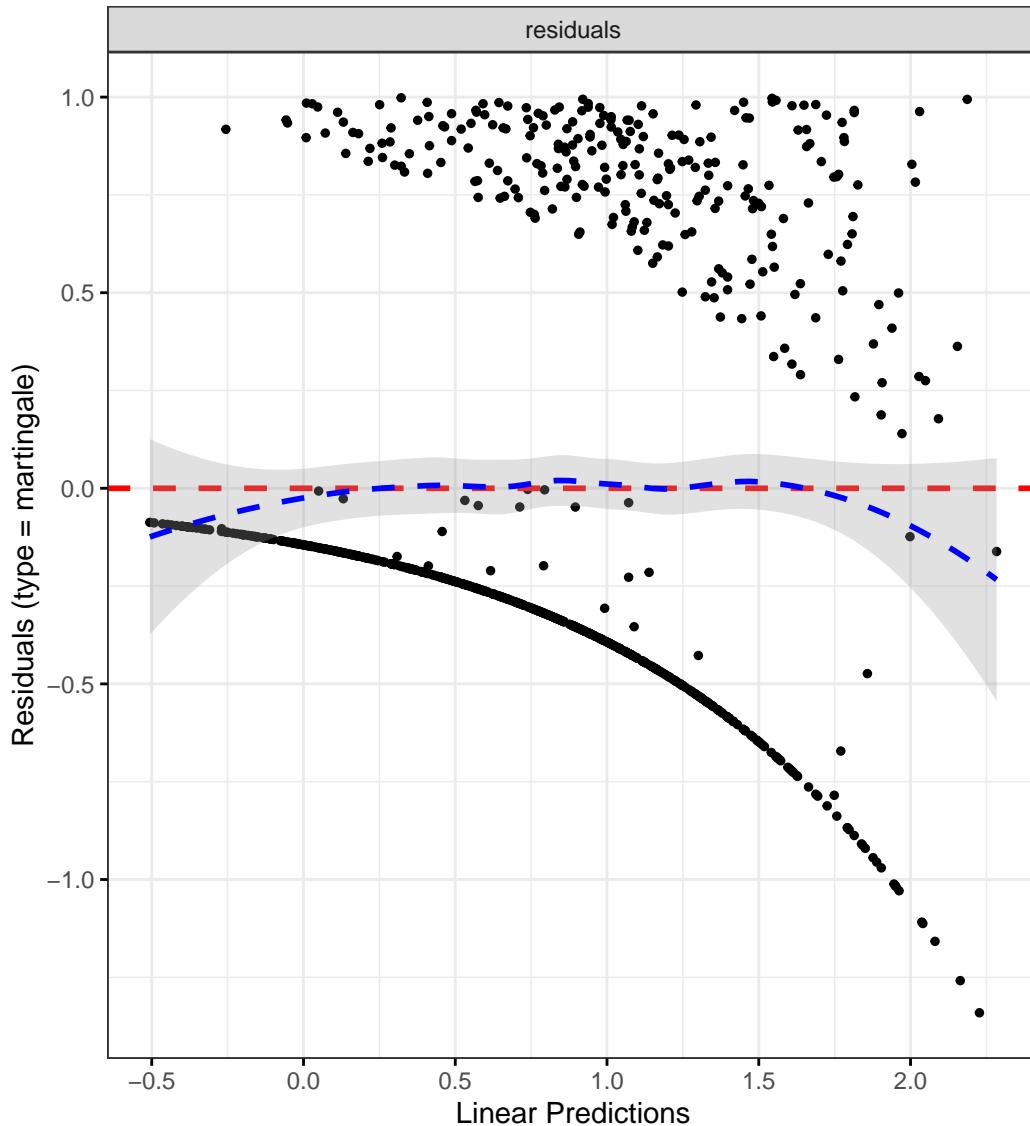
## DF betas

`geom\_smooth()` using formula = 'y ~ x'



Martingale residuals

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



White cell count

```
fit.me <- coxph(  
  Surv(softflare_time, softflare) ~  
    Sex +
```

```

IMD +
cat +
`IBD Duration` +
BMI +
Treatment +
Age +
WCC +
frailty(SiteNo),
control = coxph.control(outer.max = 20),
data = flare.cd.df
)

invisible(cox_summary(fit.me))

```

Cox model summary:

Variable	HR	Lower 95%	Upper 95%	P-value
SexFemale	2.4582	1.8550	3.2574	0.0000
IMD2	1.0346	0.6012	1.7807	0.9022
IMD3	1.0013	0.5728	1.7505	0.9964
IMD4	1.1351	0.6667	1.9326	0.6407
IMD5	1.0768	0.6431	1.8032	0.7784
catFC 50-250	1.5111	1.1335	2.0145	0.0049
catFC > 250	2.3572	1.7088	3.2516	0.0000
IBD Duration	0.9953	0.9833	1.0075	0.4499
BMI	1.0199	0.9956	1.0448	0.1097
TreatmentMono biologic	1.0885	0.7444	1.5916	0.6620
TreatmentCombo therapy	0.7724	0.4881	1.2223	0.2701
Treatment5-ASA	1.2253	0.6502	2.3092	0.5297
TreatmentNone reported	1.0393	0.7272	1.4855	0.8323
Age	1.0023	0.9927	1.0119	0.6404
WCC	0.9974	0.9537	1.0431	0.9095

Diagnostics:

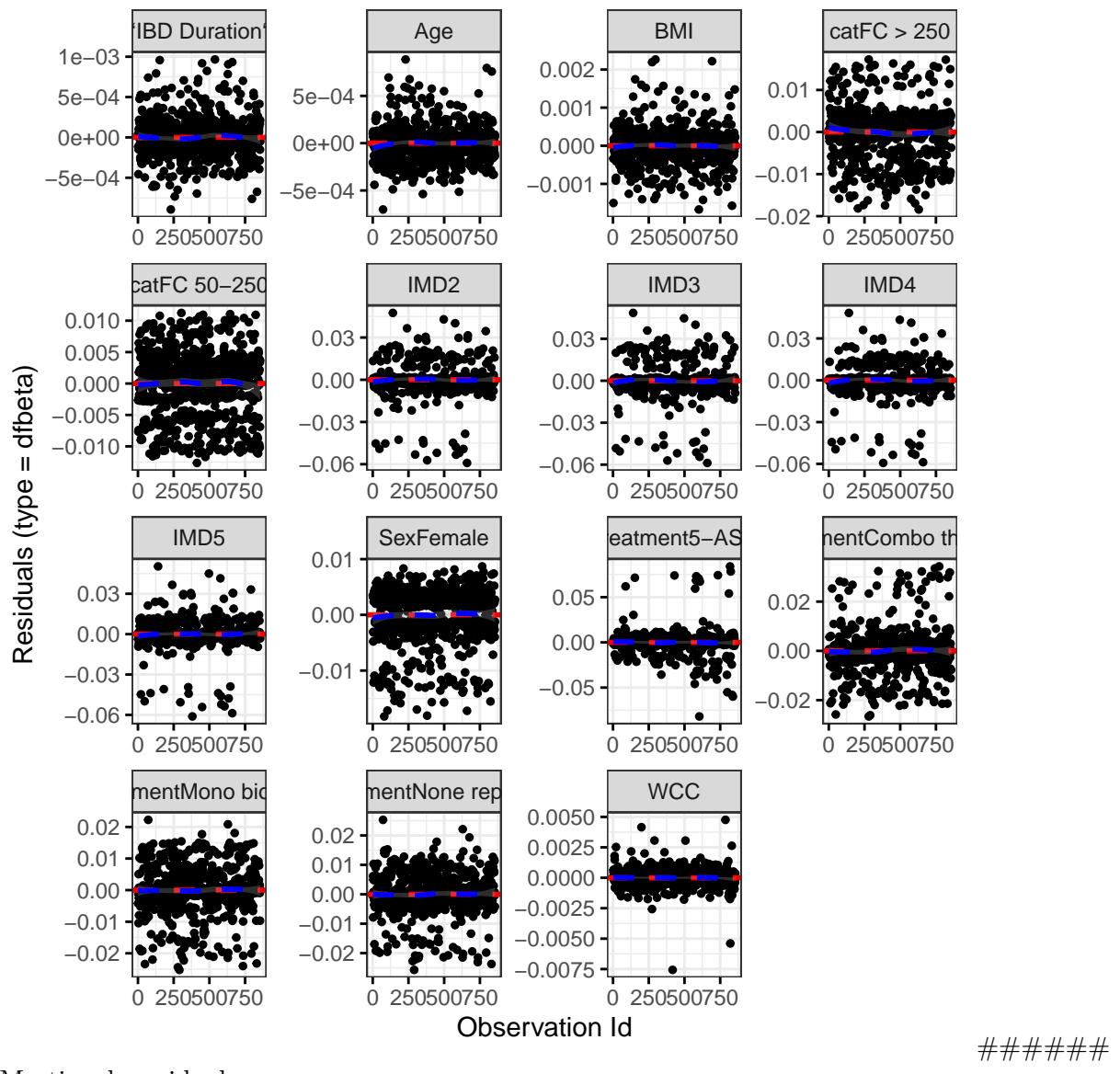
### Proportional hazards assumption test

	Chi-squared statistic	DF	P-value
Sex	0.7886	0.9942	0.3725

	Chi-squared statistic	DF	P-value
IMD	4.6795	3.9666	0.3174
cat	2.5678	1.9881	0.2748
IBD Duration	3.5210	0.9964	0.0603
BMI	4.0258	0.9950	0.0445
Treatment	4.0799	3.9290	0.3850
Age	0.8093	0.9951	0.3666
WCC	0.8075	0.9956	0.3673
GLOBAL	20.8247	17.8669	0.2809

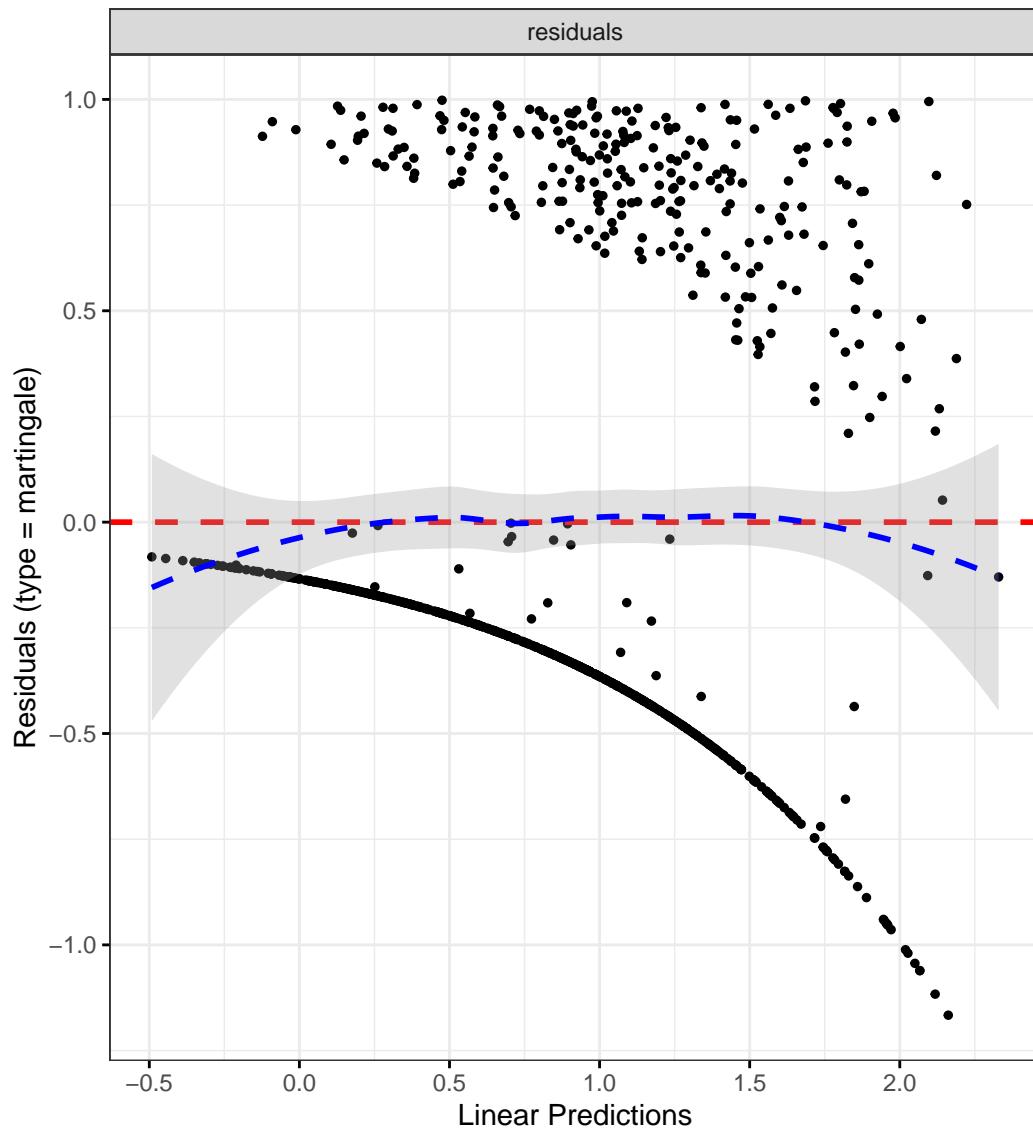
## DF betas

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



Martingale residuals

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



Platelets

```
fit.me <- coxph(
  Surv(softflare_time, softflare) ~
    Sex +
    IMD +
    cat +
    `IBD Duration` +
    BMI +
    Treatment +
```

```

Age +
Platelets +
frailty(SiteNo),
control = coxph.control(outer.max = 20),
data = flare.cd.df
)

invisible(cox_summary(fit.me))

```

Cox model summary:

Variable	HR	Lower 95%	Upper 95%	P-value
SexFemale	2.4111	1.8051	3.2205	0.0000
IMD2	1.0358	0.6016	1.7832	0.8991
IMD3	0.9909	0.5668	1.7321	0.9743
IMD4	1.1281	0.6625	1.9211	0.6572
IMD5	1.0707	0.6392	1.7935	0.7953
catFC 50-250	1.5006	1.1261	1.9996	0.0056
catFC > 250	2.3071	1.6638	3.1991	0.0000
IBD Duration	0.9955	0.9834	1.0077	0.4677
BMI	1.0193	0.9950	1.0442	0.1205
TreatmentMono biologic	1.0943	0.7504	1.5958	0.6396
TreatmentCombo therapy	0.7712	0.4878	1.2191	0.2661
Treatment5-ASA	1.2113	0.6438	2.2793	0.5522
TreatmentNone reported	1.0409	0.7298	1.4847	0.8247
Age	1.0027	0.9931	1.0125	0.5805
Platelets	1.0005	0.9988	1.0022	0.5775

Diagnostics:

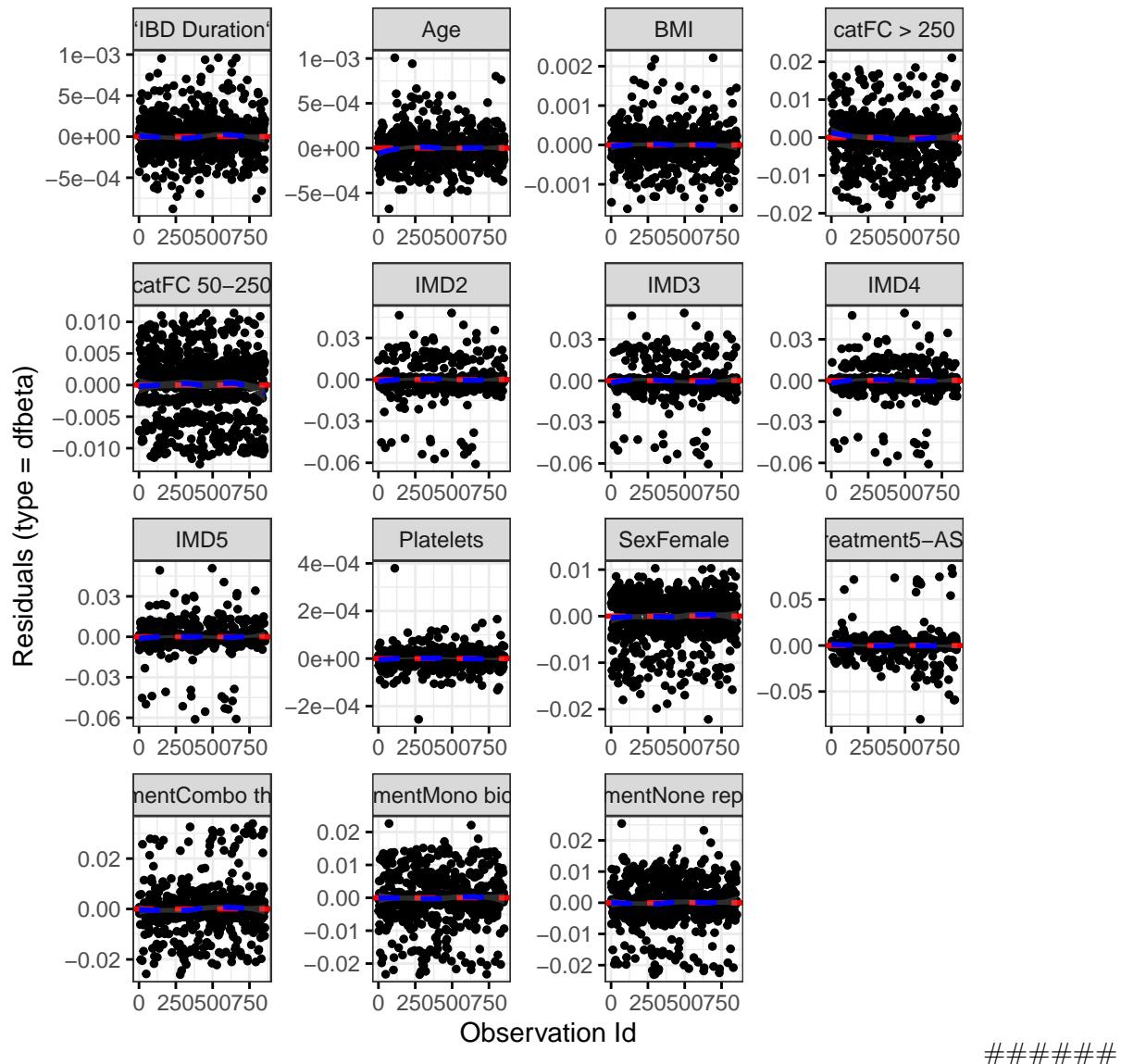
### Proportional hazards assumption test

	Chi-squared statistic	DF	P-value
Sex	0.7853	0.9954	0.3739
IMD	4.7137	3.9681	0.3138
cat	2.5936	1.9885	0.2713
IBD Duration	3.5158	0.9965	0.0605
BMI	4.0297	0.9950	0.0444
Treatment	4.0671	3.9330	0.3872

	Chi-squared statistic	DF	P-value
Age	0.7788	0.9958	0.3760
Platelets	1.8028	0.9945	0.1782
GLOBAL	21.6810	17.6807	0.2302

## DF betas

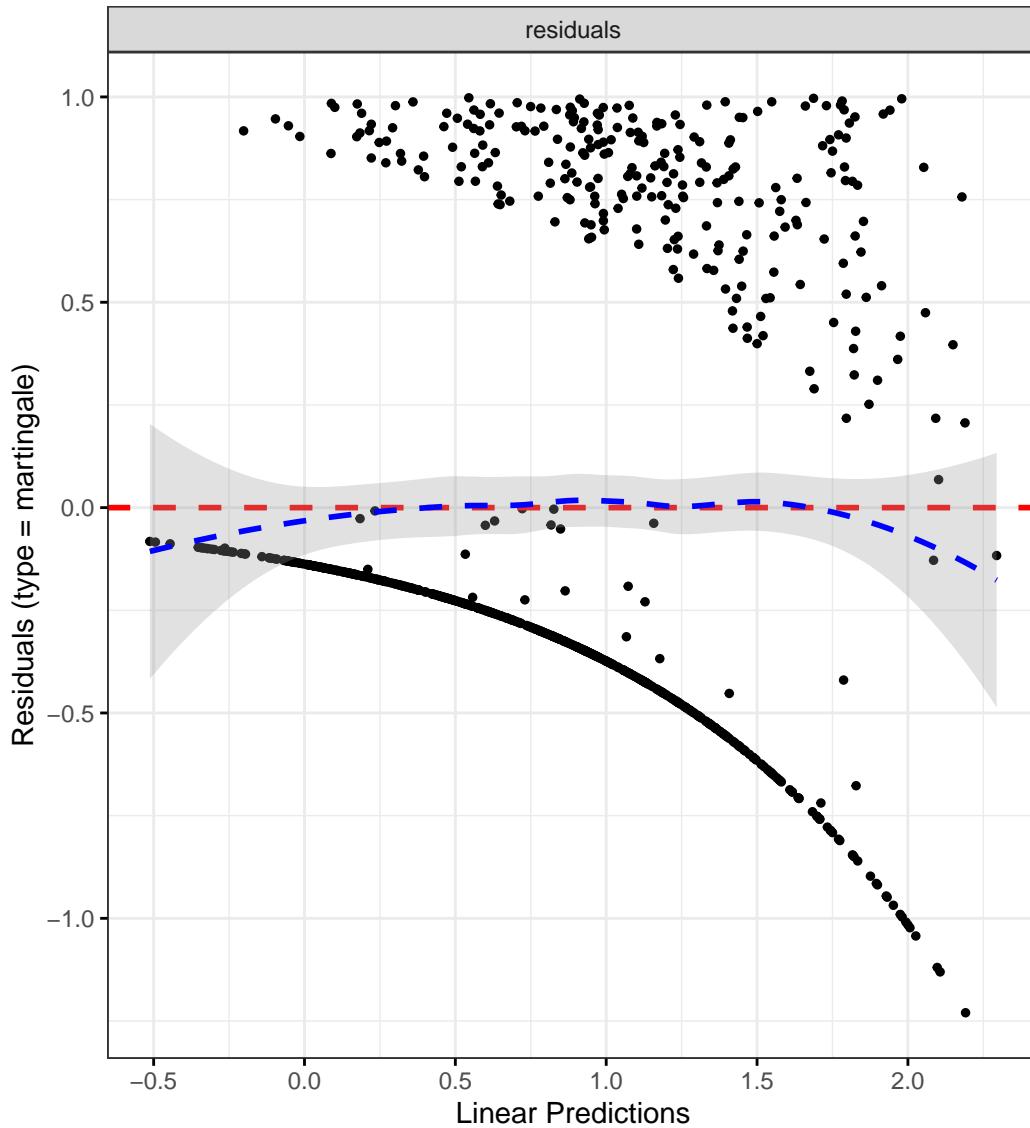
`geom\_smooth()` using formula = 'y ~ x'



#####

Martingale residuals

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



Albumin

```
fit.me <- coxph(  
  Surv(softflare_time, softflare) ~  
    Sex +
```

```

IMD +
cat +
`IBD Duration` +
BMI +
Treatment +
Age +
Albumin +
frailty(SiteNo),
control = coxph.control(outer.max = 20),
data = flare.cd.df
)

invisible(cox_summary(fit.me))

```

Cox model summary:

Variable	HR	Lower 95%	Upper 95%	P-value
SexFemale	2.4223	1.8119	3.2382	0.0000
IMD2	1.0815	0.6202	1.8859	0.7824
IMD3	1.0549	0.5959	1.8675	0.8544
IMD4	1.1648	0.6755	2.0088	0.5831
IMD5	1.0816	0.6366	1.8375	0.7719
catFC 50-250	1.5522	1.1568	2.0826	0.0034
catFC > 250	2.1298	1.5104	3.0033	0.0000
IBD Duration	0.9934	0.9813	1.0058	0.2958
BMI	1.0158	0.9909	1.0413	0.2154
TreatmentMono biologic	1.0781	0.7318	1.5883	0.7035
TreatmentCombo therapy	0.7625	0.4757	1.2223	0.2601
Treatment5-ASA	1.2974	0.6819	2.4687	0.4276
TreatmentNone reported	1.0413	0.7238	1.4981	0.8272
Age	1.0049	0.9950	1.0148	0.3357
Albumin	0.9742	0.9428	1.0066	0.1168

Diagnostics:

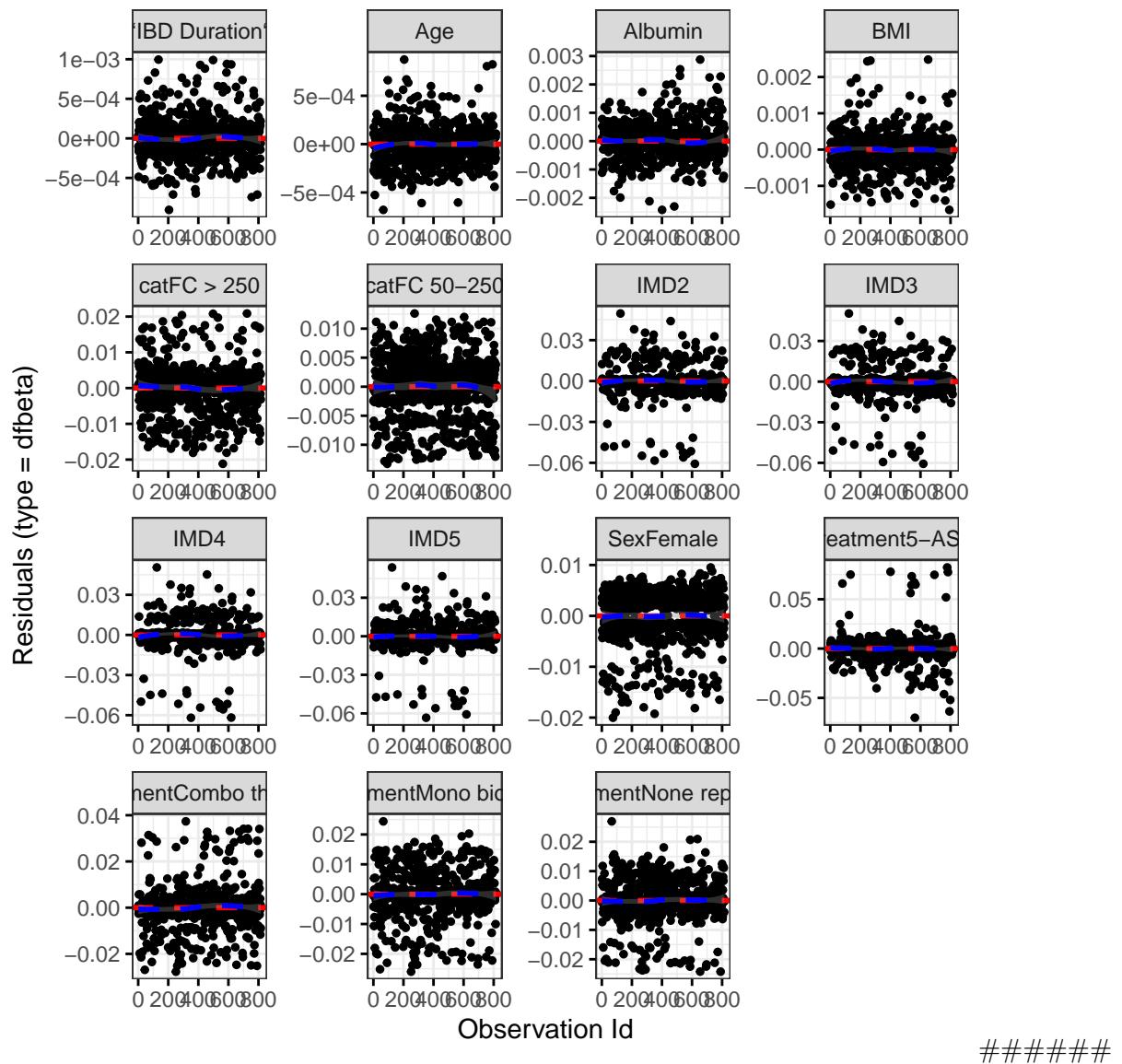
#### Proportional hazards assumption test

	Chi-squared statistic	DF	P-value
Sex	1.0516	1.0000	0.3051

	Chi-squared statistic	DF	P-value
IMD	4.3697	4.0000	0.3583
cat	4.4235	2.0000	0.1095
IBD Duration	3.1653	1.0000	0.0752
BMI	2.6329	1.0000	0.1047
Treatment	4.1227	4.0000	0.3897
Age	0.3211	1.0000	0.5709
Albumin	0.7032	1.0000	0.4017
GLOBAL	19.8301	15.0001	0.1786

## DF betas

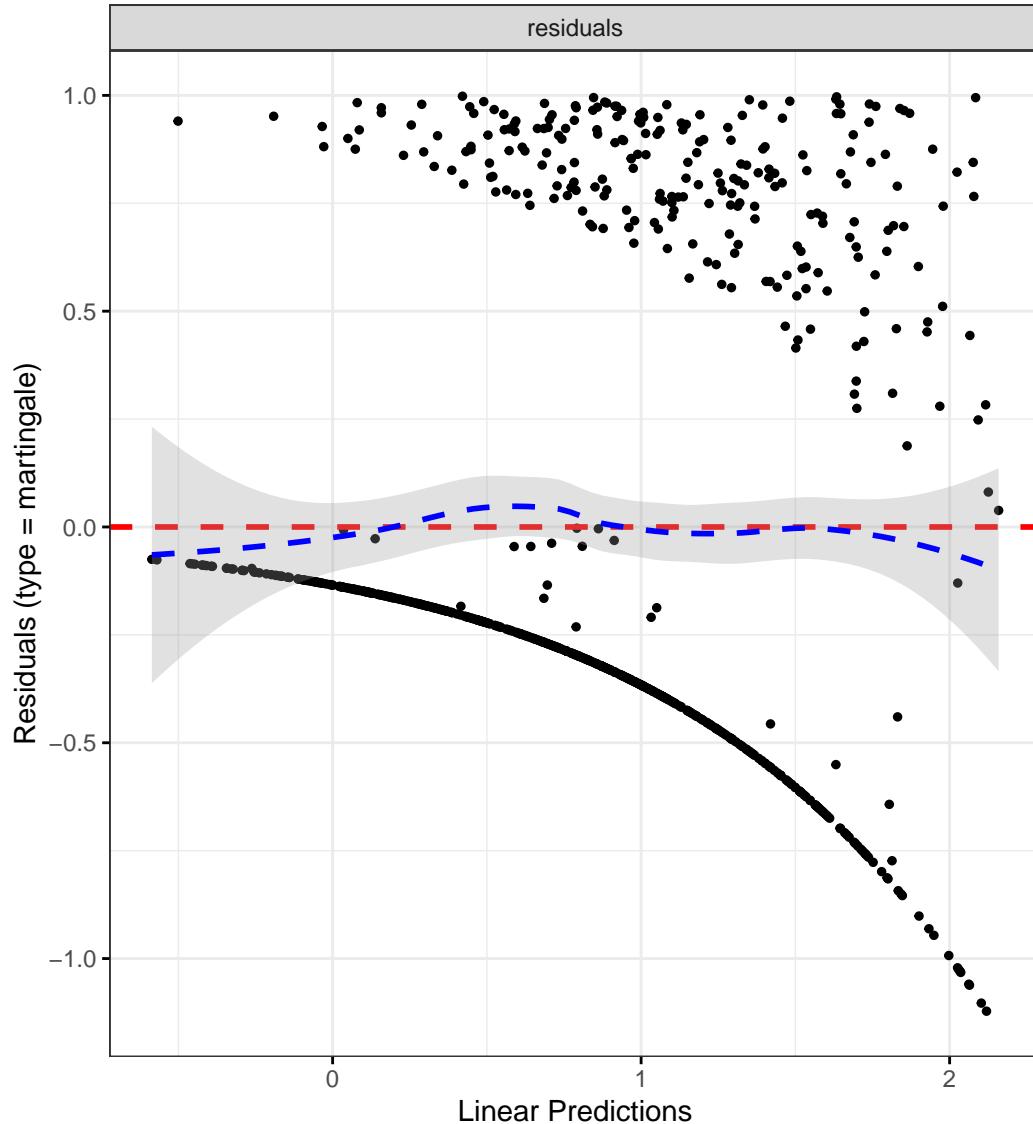
```
`geom_smooth()` using formula = 'y ~ x'
```



#####

Martingale residuals

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



Objective flare

C-reactive protein

```
fit.me <- coxph(
  Surv(hardflare_time, hardflare) ~
    Sex +
    IMD +
    cat +
    `IBD Duration` +
```

```

    BMI +
Treatment +
Age +
CReactiveProtein +
frailty(SiteNo),
control = coxph.control(outer.max = 20),
data = flare.cd.df
)

invisible(cox_summary(fit.me))

```

Cox model summary:

Variable	HR	Lower 95%	Upper 95%	P-value
SexFemale	1.8009	1.2824	2.5291	0.0007
IMD2	1.2982	0.6151	2.7400	0.4935
IMD3	1.1456	0.5291	2.4807	0.7302
IMD4	1.3636	0.6493	2.8636	0.4127
IMD5	1.5807	0.7763	3.2184	0.2069
catFC 50-250	1.9148	1.3007	2.8189	0.0010
catFC > 250	3.0442	2.0116	4.6068	0.0000
IBD Duration	0.9875	0.9711	1.0042	0.1425
BMI	1.0150	0.9836	1.0474	0.3528
TreatmentMono biologic	0.9277	0.5912	1.4556	0.7439
TreatmentCombo therapy	0.5426	0.3015	0.9764	0.0414
Treatment5-ASA	1.2132	0.5261	2.7974	0.6503
TreatmentNone reported	0.6823	0.4395	1.0592	0.0884
Age	0.9894	0.9775	1.0014	0.0831
CReactiveProtein	1.0079	0.9989	1.0170	0.0868

Diagnostics:

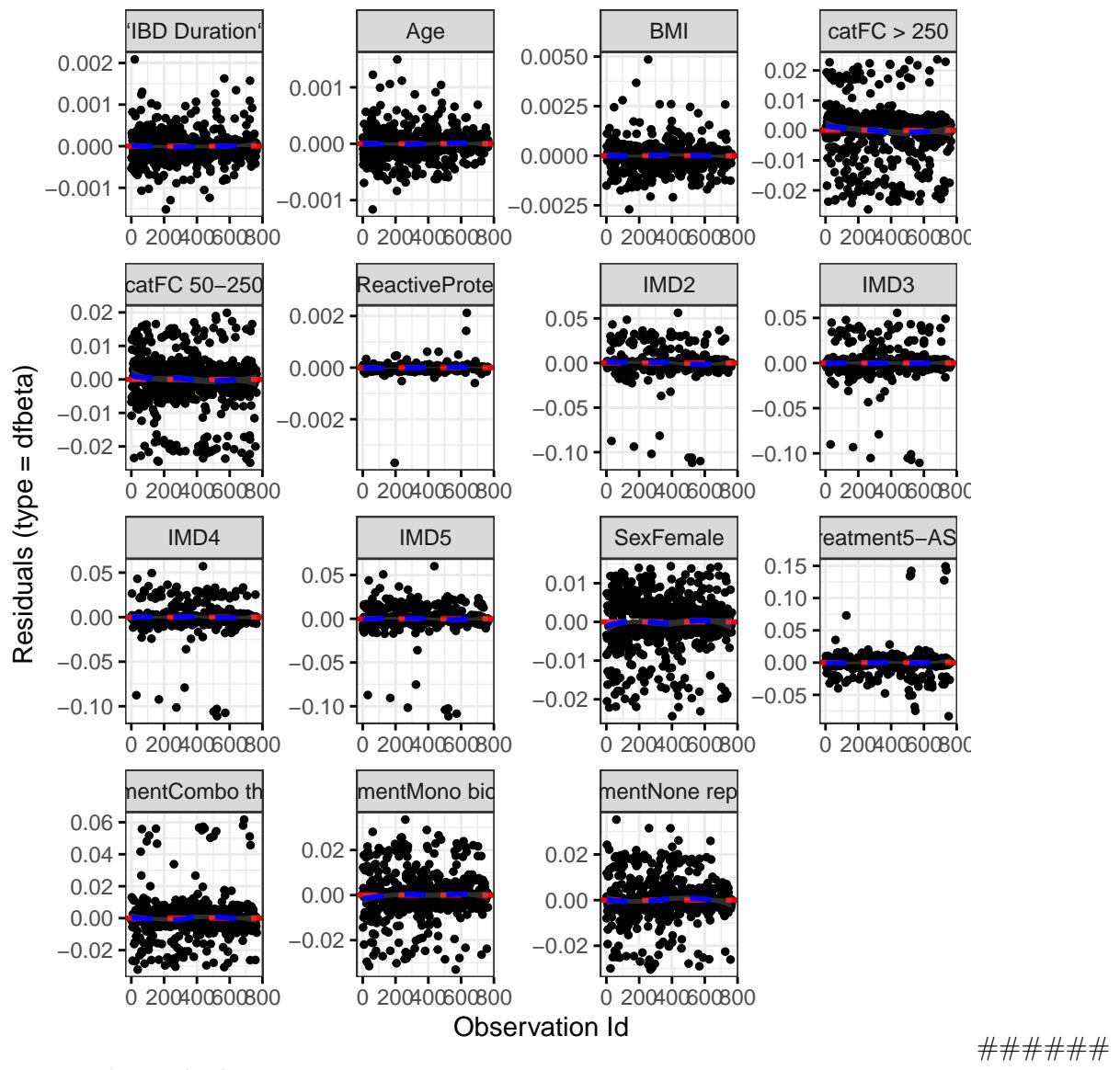
#### Proportional hazards assumption test

	Chi-squared statistic	DF	P-value
Sex	3.3788	0.9933	0.0654
IMD	1.4745	3.9766	0.8285
cat	6.6521	1.9943	0.0357
IBD Duration	0.5193	0.9984	0.4705

	Chi-squared statistic	DF	P-value
BMI	0.5636	0.9973	0.4518
Treatment	3.7682	3.9447	0.4299
Age	4.2899	0.9986	0.0383
CRactiveProtein	2.5542	0.9967	0.1095
GLOBAL	24.5659	16.5537	0.0921

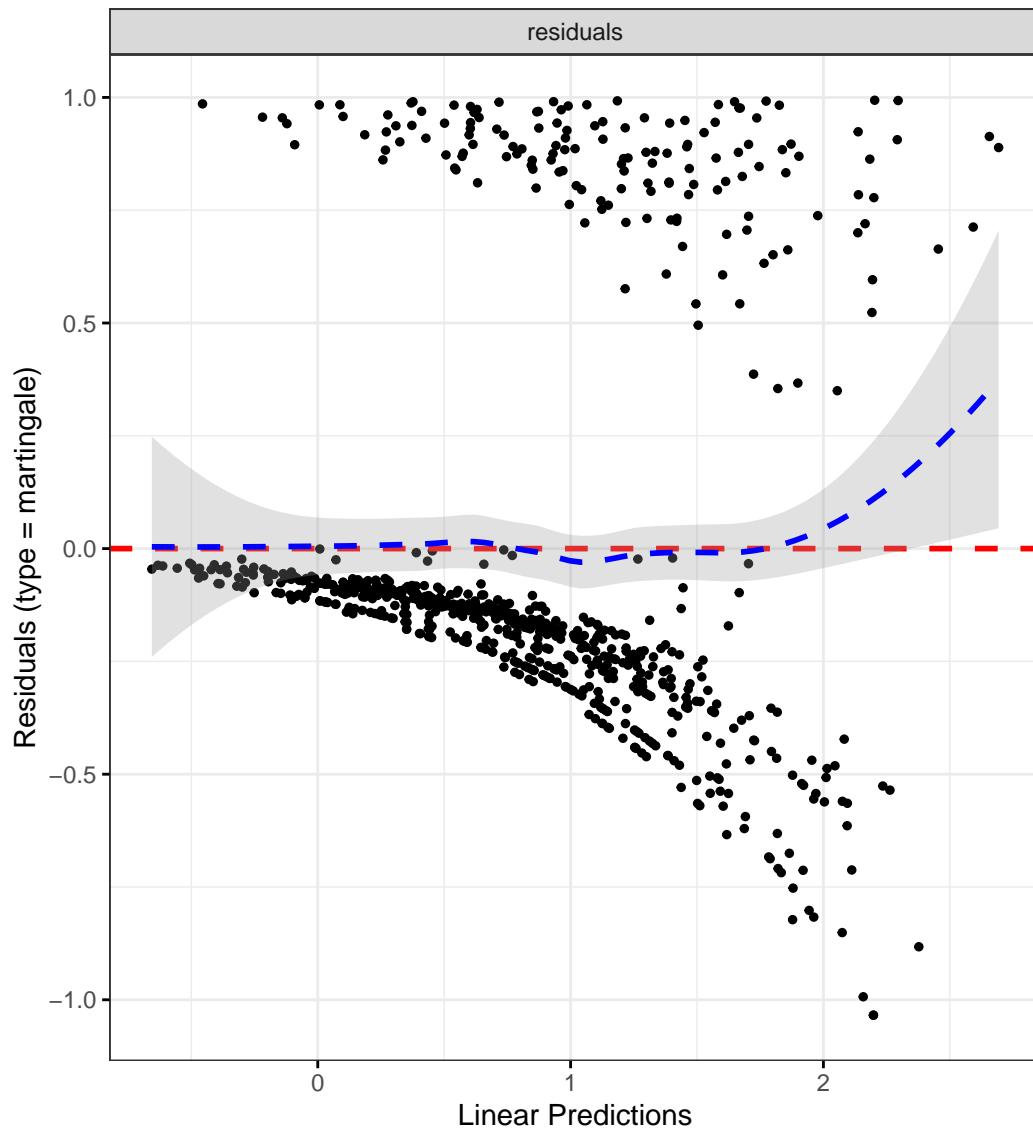
## DF betas

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



Martingale residuals

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



Haemoglobin

```
fit.me <- coxph(
  Surv(hardflare_time, hardflare) ~
    Sex +
    IMD +
    cat +
    `IBD Duration` +
    BMI +
    Treatment +
```

```

Age +
Haemoglobin +
frailty(SiteNo),
control = coxph.control(outer.max = 20),
data = flare.cd.df
)

invisible(cox_summary(fit.me))

```

Cox model summary:

Variable	HR	Lower 95%	Upper 95%	P-value
SexFemale	1.6362	1.1086	2.4150	0.0132
IMD2	1.1011	0.5521	2.1957	0.7845
IMD3	1.0211	0.4986	2.0911	0.9545
IMD4	1.1355	0.5706	2.2596	0.7175
IMD5	1.2666	0.6581	2.4377	0.4792
catFC 50-250	2.0028	1.3772	2.9125	0.0003
catFC > 250	3.4633	2.3126	5.1866	0.0000
IBD Duration	0.9831	0.9669	0.9996	0.0442
BMI	1.0186	0.9874	1.0507	0.2454
TreatmentMono biologic	1.0278	0.6616	1.5967	0.9029
TreatmentCombo therapy	0.6058	0.3435	1.0684	0.0834
Treatment5-ASA	1.3132	0.5601	3.0791	0.5309
TreatmentNone reported	0.7177	0.4661	1.1050	0.1320
Age	0.9874	0.9758	0.9991	0.0347
Haemoglobin	0.9948	0.9819	1.0079	0.4365

Diagnostics:

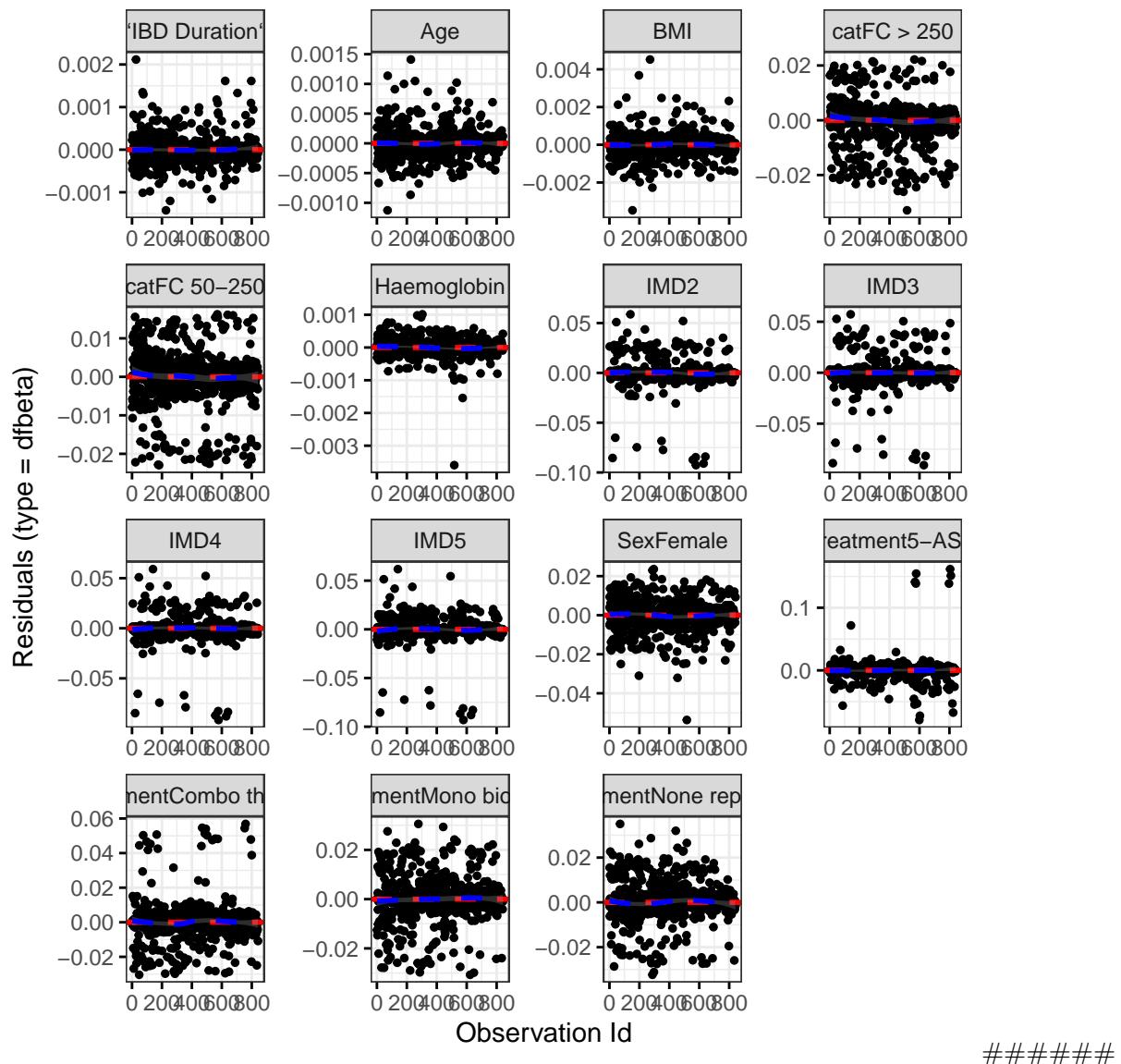
### Proportional hazards assumption test

	Chi-squared statistic	DF	P-value
Sex	2.7311	0.9825	0.0961
IMD	1.5024	3.9512	0.8206
cat	8.0023	1.9856	0.0180
IBD Duration	0.5793	0.9951	0.4447
BMI	0.6984	0.9922	0.4004
Treatment	3.7108	3.8784	0.4281

	Chi-squared statistic	DF	P-value
Age	4.1839	0.9936	0.0404
Haemoglobin	3.3939	0.9924	0.0647
GLOBAL	24.5533	20.9027	0.2622

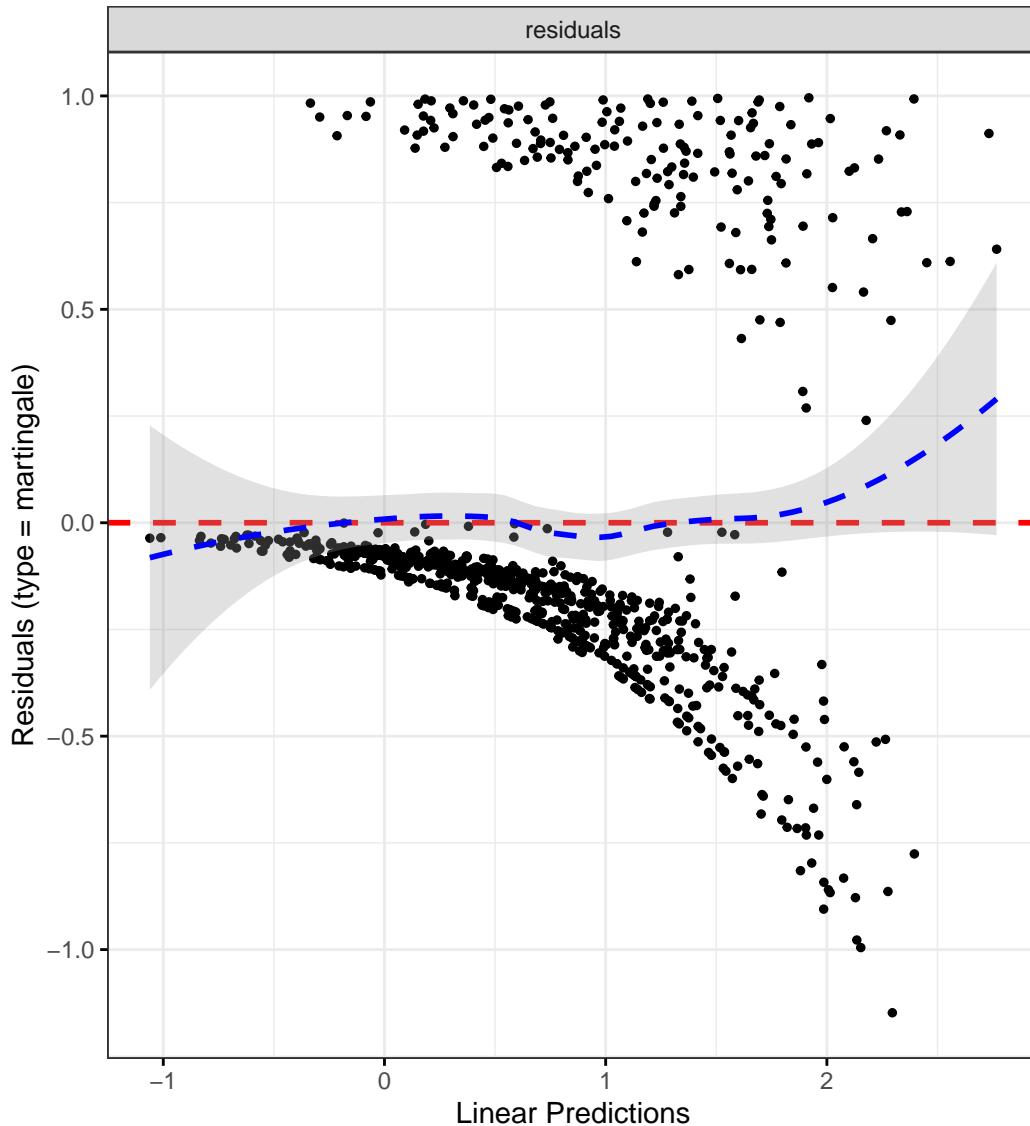
## DF betas

`geom\_smooth()` using formula = 'y ~ x'



Martingale residuals

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



White cell count

```
fit.me <- coxph(  
  Surv(hardflare_time, hardflare) ~  
    Sex +
```

```

IMD +
cat +
`IBD Duration` +
BMI +
Treatment +
Age +
WCC +
frailty(SiteNo),
control = coxph.control(outer.max = 20),
data = flare.cd.df
)

invisible(cox_summary(fit.me))

```

Cox model summary:

Variable	HR	Lower 95%	Upper 95%	P-value
SexFemale	1.7502	1.2575	2.4359	0.0009
IMD2	1.0977	0.5504	2.1895	0.7912
IMD3	0.9946	0.4858	2.0365	0.9882
IMD4	1.1361	0.5709	2.2607	0.7163
IMD5	1.2736	0.6618	2.4509	0.4690
catFC 50-250	1.9979	1.3735	2.9063	0.0003
catFC > 250	3.4956	2.3395	5.2231	0.0000
IBD Duration	0.9833	0.9671	0.9998	0.0477
BMI	1.0158	0.9851	1.0475	0.3155
TreatmentMono biologic	0.9958	0.6401	1.5493	0.9851
TreatmentCombo therapy	0.5987	0.3388	1.0581	0.0775
Treatment5-ASA	1.2362	0.5277	2.8958	0.6255
TreatmentNone reported	0.6974	0.4529	1.0740	0.1019
Age	0.9879	0.9764	0.9995	0.0414
WCC	1.0276	0.9739	1.0843	0.3197

Diagnostics:

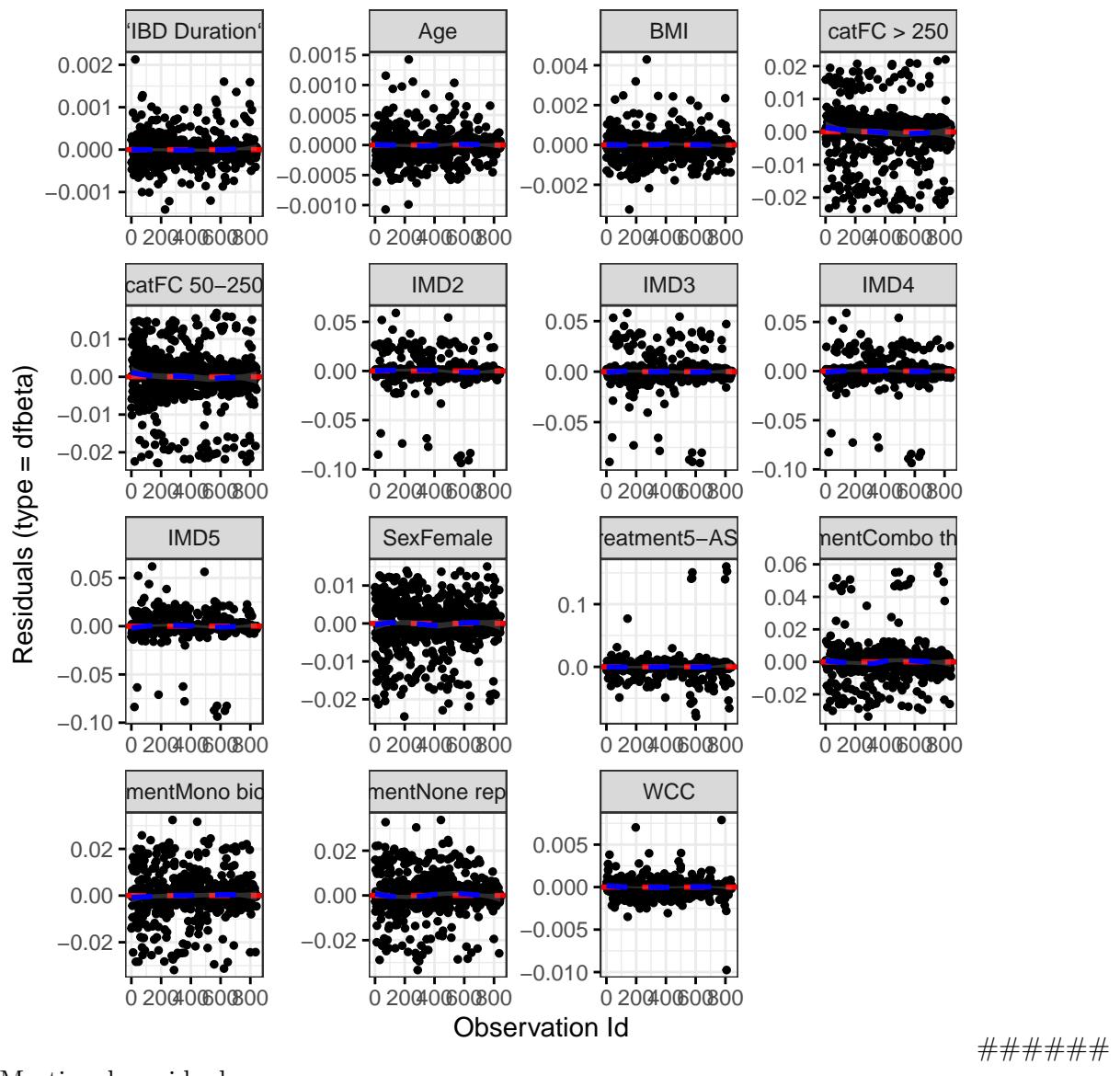
#### Proportional hazards assumption test

	Chi-squared statistic	DF	P-value
Sex	2.6902	0.9856	0.0990

	Chi-squared statistic	DF	P-value
IMD	1.4300	3.9504	0.8335
cat	8.0723	1.9857	0.0174
IBD Duration	0.5879	0.9955	0.4415
BMI	0.6962	0.9918	0.4010
Treatment	3.7116	3.8780	0.4279
Age	4.2743	0.9935	0.0383
WCC	3.7589	0.9883	0.0516
GLOBAL	26.5752	20.9009	0.1814

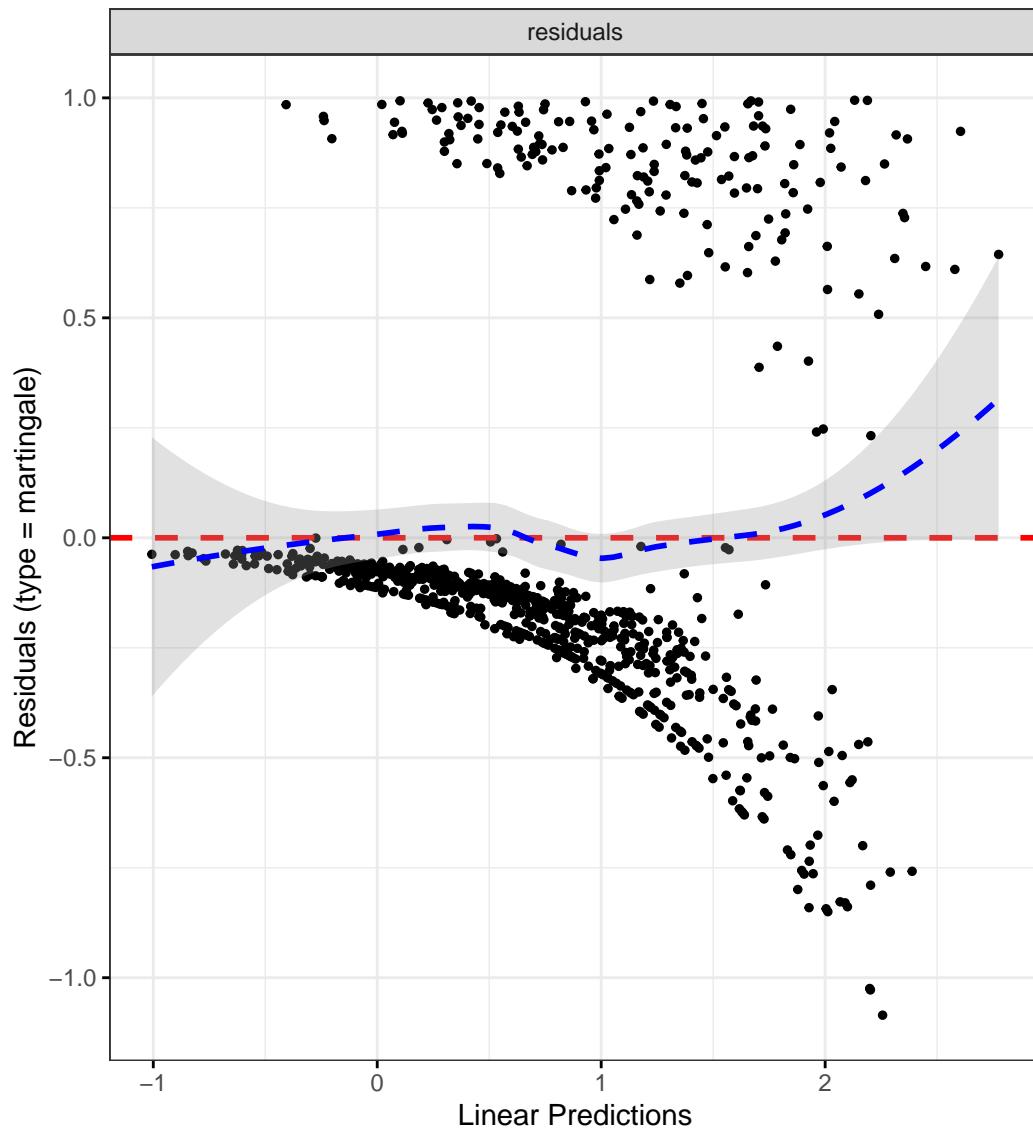
## DF betas

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



Martingale residuals

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



Platelets

```
fit.me <- coxph(
  Surv(hardflare_time, hardflare) ~
    Sex +
    IMD +
    cat +
    `IBD Duration` +
    BMI +
    Treatment +
```

```

Age +
Platelets +
frailty(SiteNo),
control = coxph.control(outer.max = 20),
data = flare.cd.df
)

invisible(cox_summary(fit.me))

```

Cox model summary:

Variable	HR	Lower 95%	Upper 95%	P-value
SexFemale	1.7384	1.2340	2.4491	0.0016
IMD2	1.0967	0.5496	2.1885	0.7933
IMD3	0.9943	0.4856	2.0362	0.9876
IMD4	1.1219	0.5637	2.2331	0.7432
IMD5	1.2584	0.6535	2.4231	0.4918
catFC 50-250	1.9963	1.3714	2.9061	0.0003
catFC > 250	3.4689	2.3069	5.2163	0.0000
IBD Duration	0.9831	0.9669	0.9996	0.0448
BMI	1.0162	0.9855	1.0478	0.3055
TreatmentMono biologic	1.0226	0.6589	1.5868	0.9207
TreatmentCombo therapy	0.6081	0.3449	1.0724	0.0857
Treatment5-ASA	1.2552	0.5363	2.9376	0.6004
TreatmentNone reported	0.7108	0.4622	1.0931	0.1200
Age	0.9881	0.9765	0.9998	0.0458
Platelets	1.0005	0.9985	1.0025	0.6412

Diagnostics:

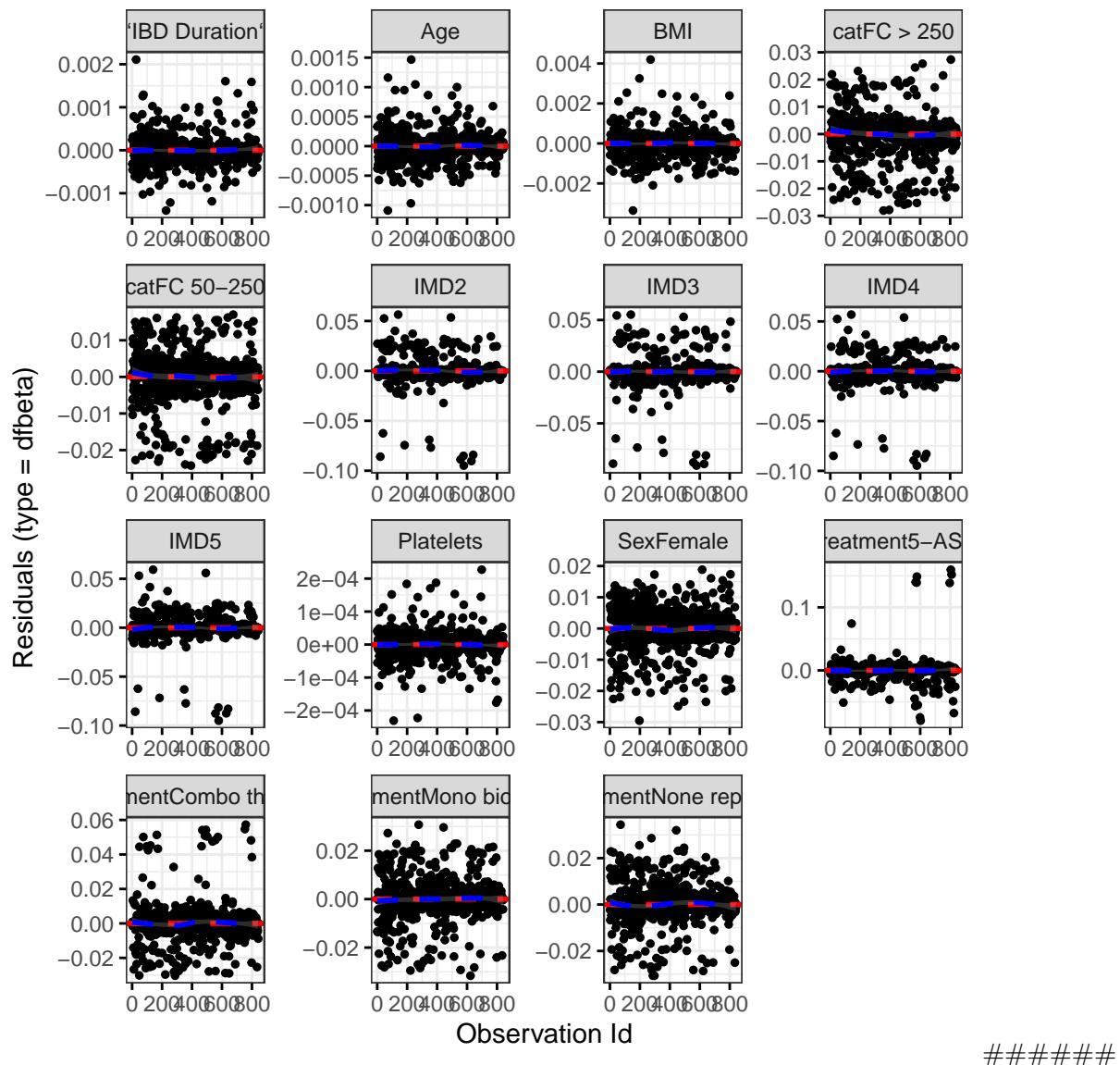
### Proportional hazards assumption test

	Chi-squared statistic	DF	P-value
Sex	2.7275	0.9878	0.0970
IMD	1.4720	3.9507	0.8260
cat	7.9710	1.9855	0.0183
IBD Duration	0.6017	0.9954	0.4361
BMI	0.6526	0.9916	0.4160
Treatment	3.7113	3.8794	0.4281

	Chi-squared statistic	DF	P-value
Age	4.1367	0.9938	0.0416
Platelets	0.4746	0.9931	0.4881
GLOBAL	23.3913	20.9178	0.3189

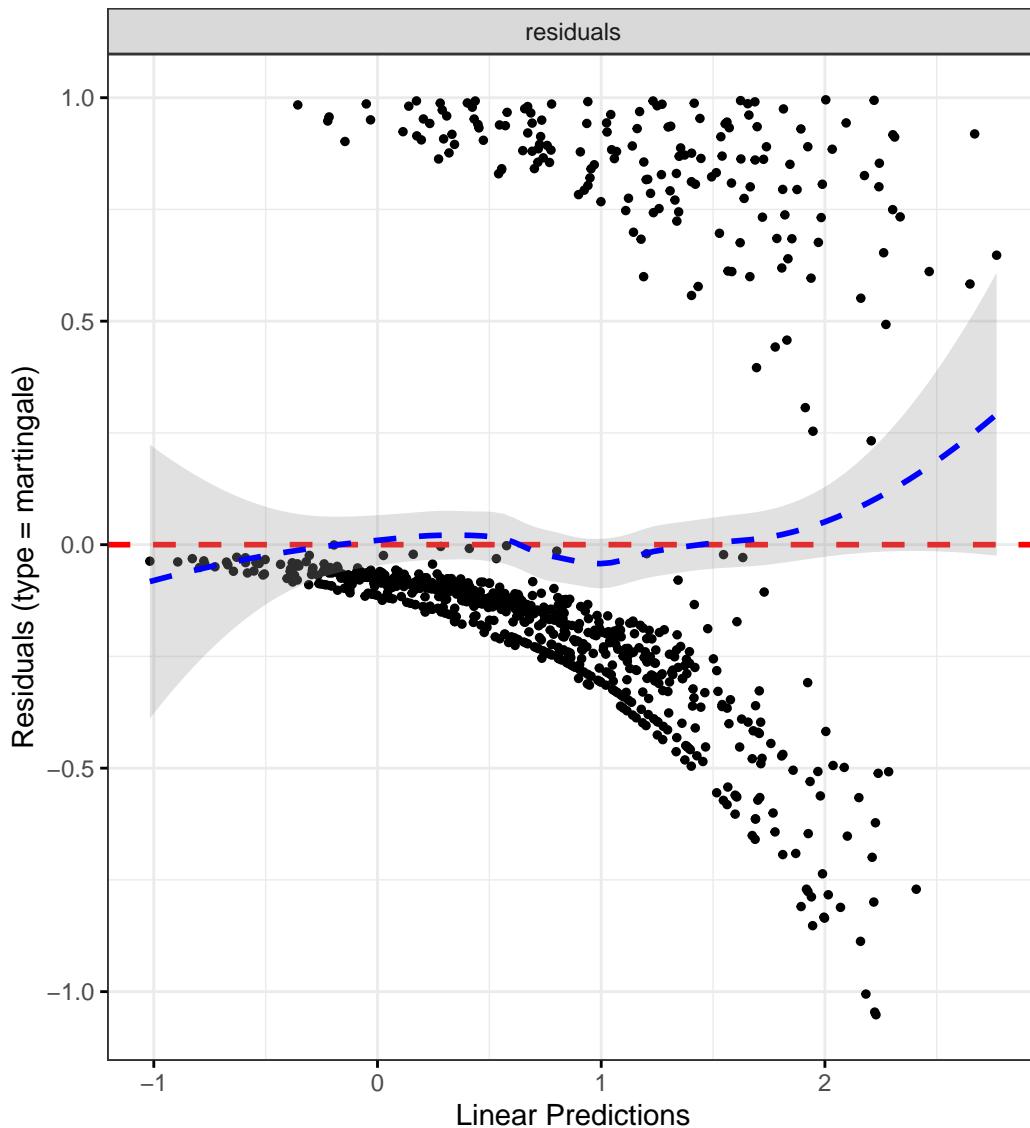
## DF betas

`geom\_smooth()` using formula = 'y ~ x'



Martingale residuals

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



Albumin

```
fit.me <- coxph(  
  Surv(hardflare_time, hardflare) ~  
    Sex +
```

```

IMD +
cat +
`IBD Duration` +
BMI +
Treatment +
Age +
Albumin +
frailty(SiteNo),
control = coxph.control(outer.max = 20),
data = flare.cd.df
)

invisible(cox_summary(fit.me))

```

Cox model summary:

Variable	HR	Lower 95%	Upper 95%	P-value
SexFemale	1.8278	1.2965	2.5769	0.0006
IMD2	1.0933	0.5464	2.1877	0.8010
IMD3	0.8936	0.4299	1.8577	0.7632
IMD4	1.0527	0.5257	2.1082	0.8847
IMD5	1.1676	0.6016	2.2662	0.6469
catFC 50-250	2.0575	1.4021	3.0193	0.0002
catFC > 250	3.2613	2.1172	5.0237	0.0000
IBD Duration	0.9857	0.9693	1.0024	0.0929
BMI	1.0151	0.9835	1.0477	0.3539
TreatmentMono biologic	1.0019	0.6377	1.5741	0.9935
TreatmentCombo therapy	0.5399	0.2978	0.9788	0.0423
Treatment5-ASA	1.1465	0.4843	2.7141	0.7558
TreatmentNone reported	0.6538	0.4192	1.0197	0.0609
Age	0.9895	0.9774	1.0018	0.0939
Albumin	1.0190	0.9740	1.0660	0.4136

Diagnostics:

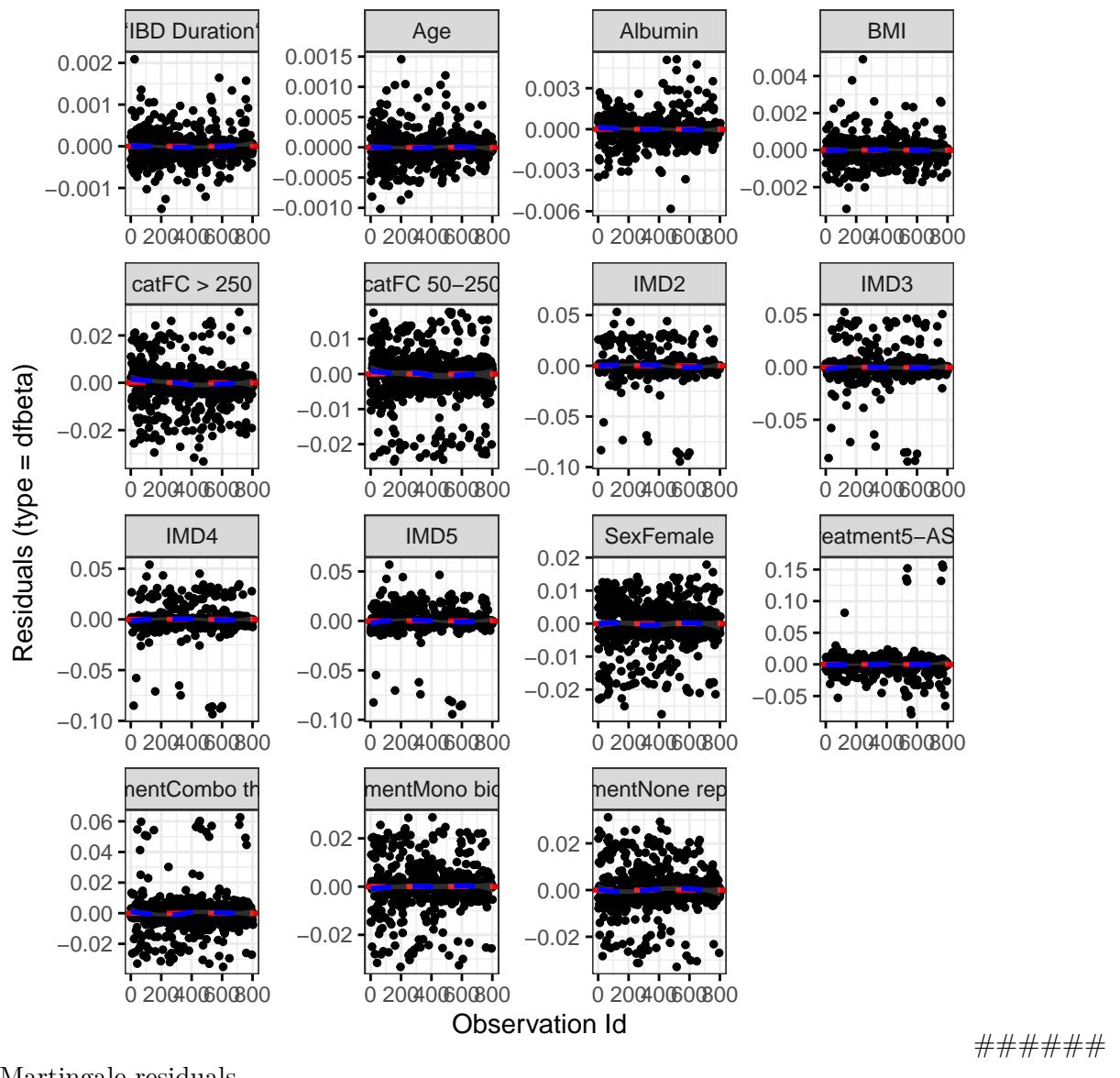
### Proportional hazards assumption test

	Chi-squared statistic	DF	P-value
Sex	3.1349	0.9861	0.0751

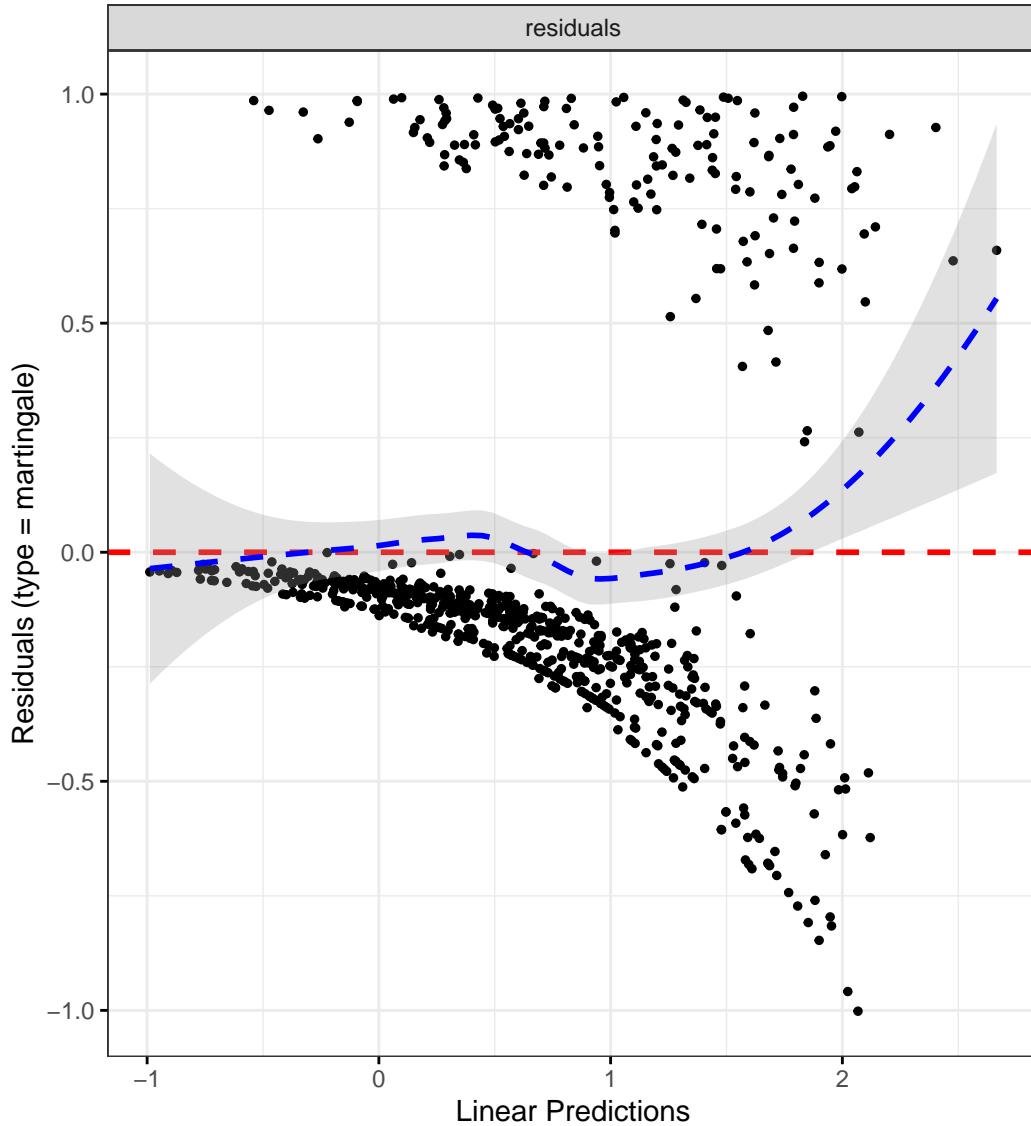
	Chi-squared statistic	DF	P-value
IMD	2.8274	3.9536	0.5798
cat	7.3567	1.9863	0.0249
IBD Duration	0.4475	0.9961	0.5019
BMI	1.0279	0.9916	0.3080
Treatment	3.6720	3.8938	0.4360
Age	2.7182	0.9894	0.0978
Albumin	2.4245	0.8954	0.1035
GLOBAL	24.6183	20.1272	0.2222

## DF betas

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



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



Ulcerative colitis

Patient-reported flare

C-reactive protein

```
fit.me <- coxph(
  Surv(softflare_time, softflare) ~
    Sex +
    IMD +
    cat +
```

```

`IBD Duration` +
BMI +
Treatment +
Age +
CReactiveProtein +
frailty(SiteNo),
control = coxph.control(outer.max = 20),
data = flare.uc.df
)

invisible(cox_summary(fit.me))

```

Cox model summary:

Variable	HR	Lower 95%	Upper 95%	P-value
SexFemale	1.5884	1.2280	2.0545	0.0004
IMD2	1.2122	0.7052	2.0835	0.4863
IMD3	0.9269	0.5411	1.5877	0.7821
IMD4	1.3318	0.8017	2.2125	0.2686
IMD5	1.0119	0.6126	1.6714	0.9633
catFC 50-250	1.6716	1.2434	2.2473	0.0007
catFC > 250	2.2180	1.6193	3.0381	0.0000
IBD Duration	0.9930	0.9790	1.0071	0.3278
BMI	0.9853	0.9607	1.0106	0.2531
TreatmentMono biologic	0.8830	0.5427	1.4365	0.6162
TreatmentCombo therapy	0.6591	0.3475	1.2501	0.2018
Treatment5-ASA	1.4834	1.0029	2.1941	0.0483
TreatmentNone reported	1.3374	0.8983	1.9911	0.1522
Age	0.9941	0.9844	1.0038	0.2312
CReactiveProtein	1.0042	0.9889	1.0198	0.5937

Diagnostics:

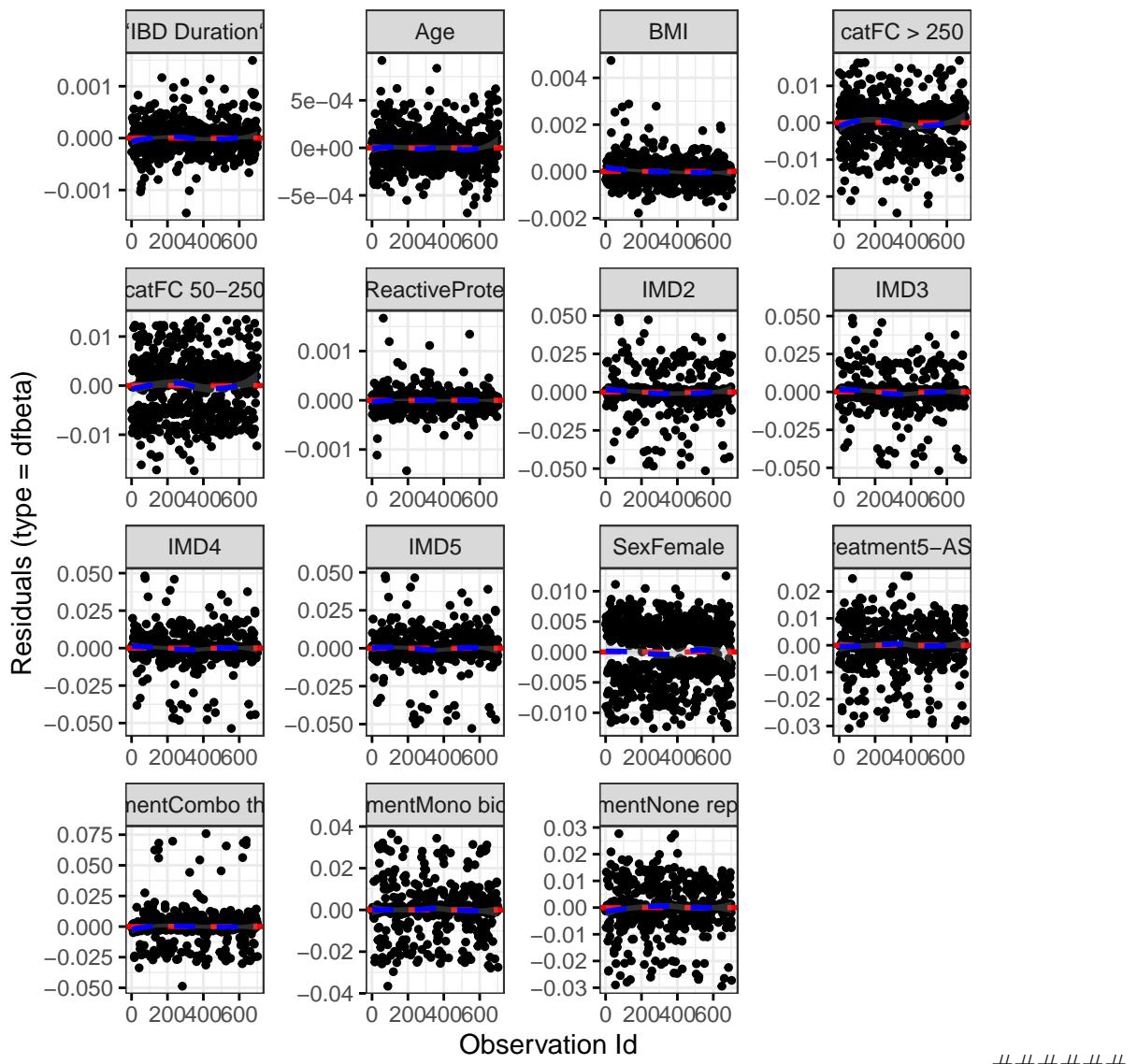
### Proportional hazards assumption test

	Chi-squared statistic	DF	P-value
Sex	2.4489	0.9872	0.1156
IMD	1.9479	3.9295	0.7356
cat	4.1711	1.9631	0.1205

	Chi-squared statistic	DF	P-value
IBD Duration	0.2523	0.9781	0.6062
BMI	1.1961	0.9835	0.2692
Treatment	7.9469	3.8538	0.0856
Age	0.0372	0.9642	0.8357
CReactiveProtein	5.8082	0.9858	0.0156
GLOBAL	24.0490	25.0397	0.5188

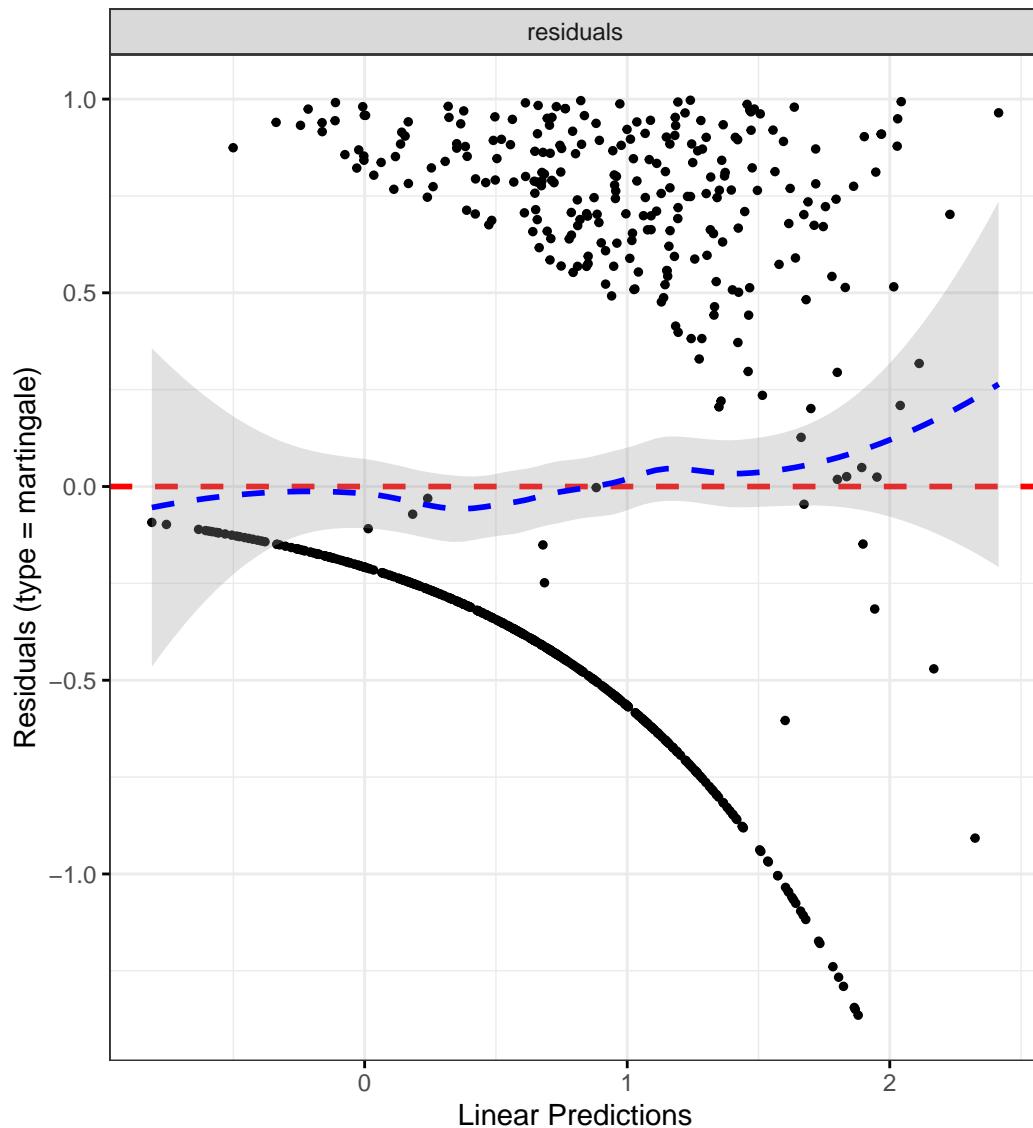
## DF betas

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



Martingale residuals

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



Haemoglobin

```
fit.me <- coxph(
  Surv(softflare_time, softflare) ~
    Sex +
    IMD +
    cat +
    `IBD Duration` +
    BMI +
    Treatment +
```

```

Age +
Haemoglobin +
frailty(SiteNo),
control = coxph.control(outer.max = 20),
data = flare.uc.df
)

invisible(cox_summary(fit.me))

```

Cox model summary:

Variable	HR	Lower 95%	Upper 95%	P-value
SexFemale	1.2212	0.8988	1.6593	0.2014
IMD2	1.2158	0.7299	2.0254	0.4529
IMD3	0.8262	0.4939	1.3822	0.4672
IMD4	1.0870	0.6679	1.7693	0.7370
IMD5	1.0451	0.6501	1.6801	0.8554
catFC 50-250	1.7626	1.3163	2.3603	0.0001
catFC > 250	2.2369	1.6398	3.0512	0.0000
IBD Duration	0.9914	0.9777	1.0053	0.2234
BMI	0.9881	0.9636	1.0132	0.3497
TreatmentMono biologic	0.9040	0.5631	1.4515	0.6762
TreatmentCombo therapy	0.6827	0.3620	1.2874	0.2382
Treatment5-ASA	1.5491	1.0584	2.2674	0.0243
TreatmentNone reported	1.3208	0.8958	1.9475	0.1601
Age	0.9922	0.9827	1.0018	0.1097
Haemoglobin	0.9841	0.9726	0.9957	0.0075

Diagnostics:

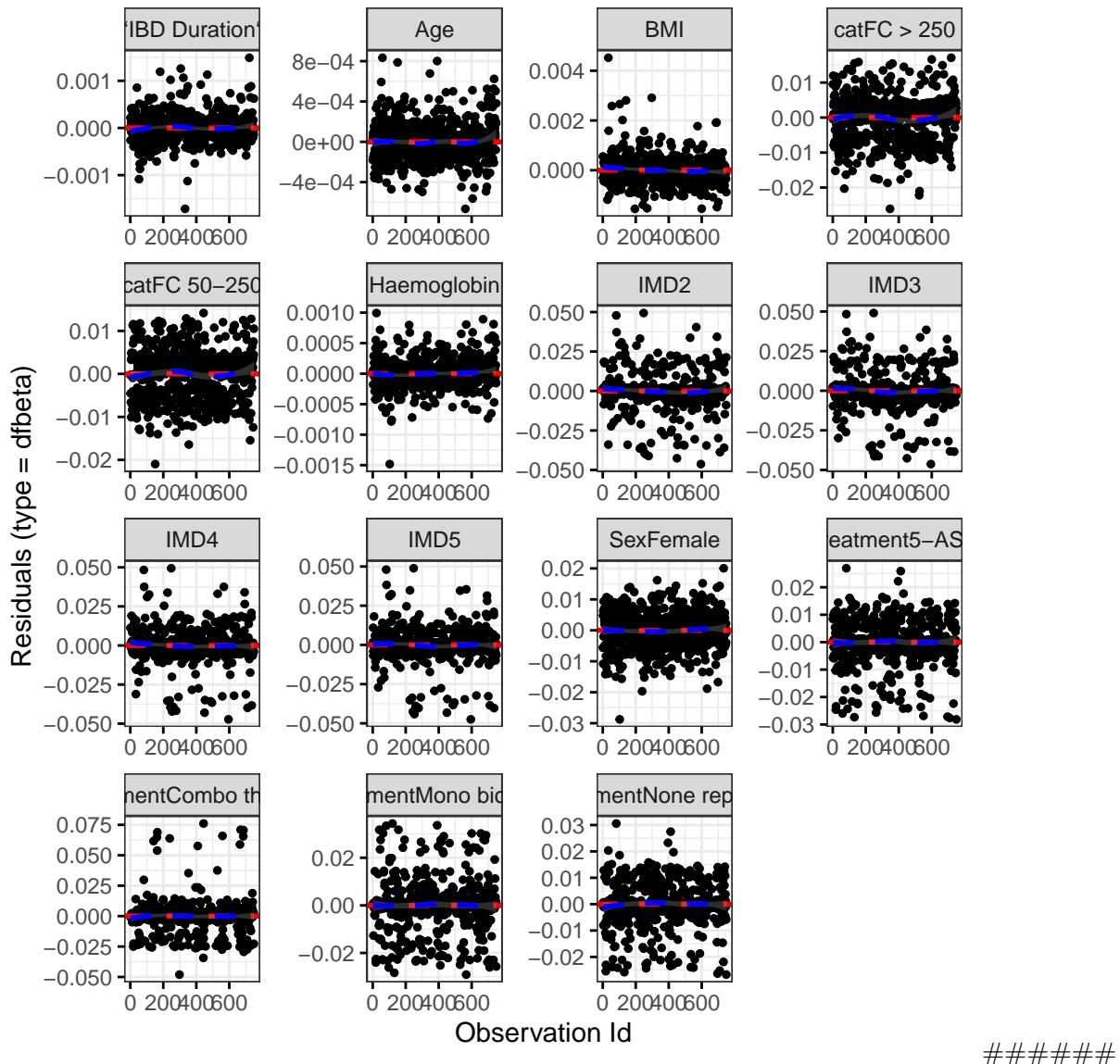
### Proportional hazards assumption test

	Chi-squared statistic	DF	P-value
Sex	1.5625	0.9825	0.2070
IMD	3.6344	3.9262	0.4464
cat	5.5826	1.9625	0.0592
IBD Duration	0.7479	0.9798	0.3798
BMI	1.5887	0.9774	0.2021
Treatment	7.3278	3.8500	0.1096

	Chi-squared statistic	DF	P-value
Age	0.0264	0.9622	0.8598
Haemoglobin	0.0066	0.9766	0.9308
GLOBAL	20.7310	25.6975	0.7417

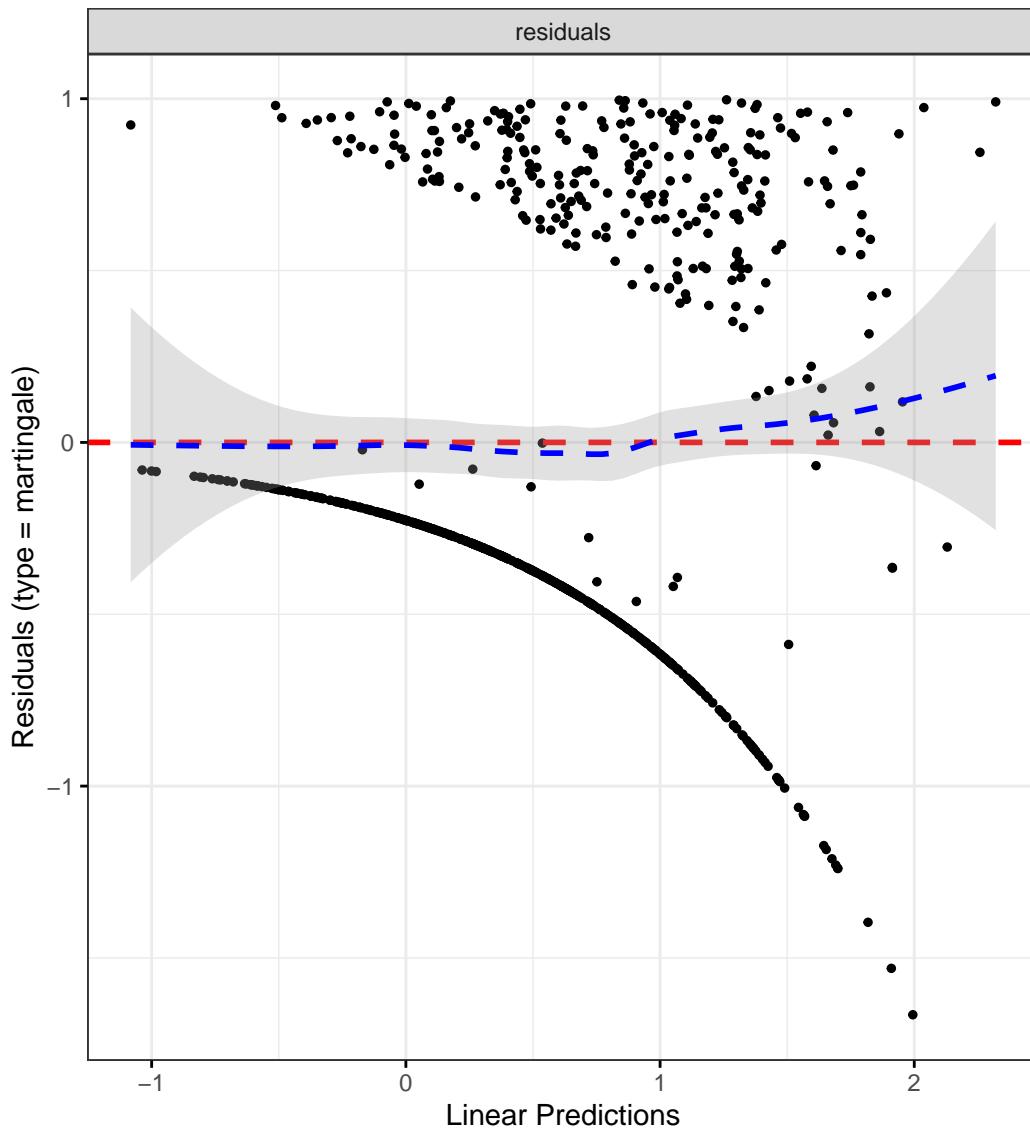
## DF betas

`geom\_smooth()` using formula = 'y ~ x'



Martingale residuals

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



White cell count

```
fit.me <- coxph(  
  Surv(softflare_time, softflare) ~  
    Sex +
```

```

IMD +
cat +
`IBD Duration` +
BMI +
Treatment +
Age +
WCC +
frailty(SiteNo),
control = coxph.control(outer.max = 20),
data = flare.uc.df
)

invisible(cox_summary(fit.me))

```

Cox model summary:

Variable	HR	Lower 95%	Upper 95%	P-value
SexFemale	1.5619	1.2129	2.0113	0.0005
IMD2	1.2368	0.7417	2.0624	0.4153
IMD3	0.8270	0.4952	1.3812	0.4679
IMD4	1.0515	0.6468	1.7093	0.8396
IMD5	1.0026	0.6251	1.6079	0.9916
catFC 50-250	1.7380	1.2958	2.3312	0.0002
catFC > 250	2.1789	1.5924	2.9815	0.0000
IBD Duration	0.9915	0.9776	1.0055	0.2319
BMI	0.9762	0.9517	1.0014	0.0639
TreatmentMono biologic	0.8061	0.5009	1.2972	0.3746
TreatmentCombo therapy	0.6372	0.3380	1.2011	0.1635
Treatment5-ASA	1.2605	0.8532	1.8624	0.2450
TreatmentNone reported	1.1027	0.7415	1.6397	0.6293
Age	0.9946	0.9849	1.0043	0.2731
WCC	1.1121	1.0404	1.1887	0.0018

Diagnostics:

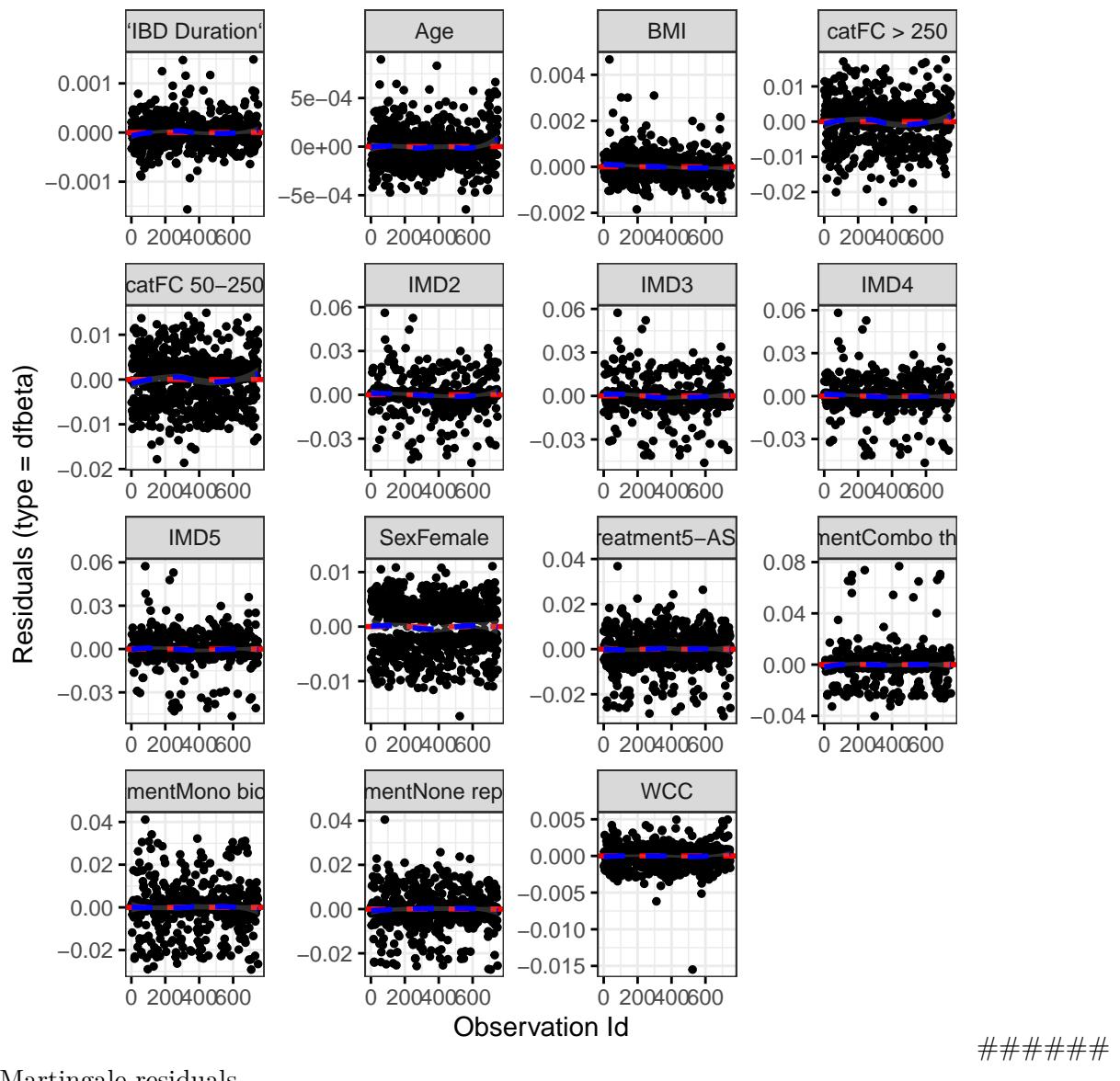
### Proportional hazards assumption test

	Chi-squared statistic	DF	P-value
Sex	1.5998	0.9861	0.2026

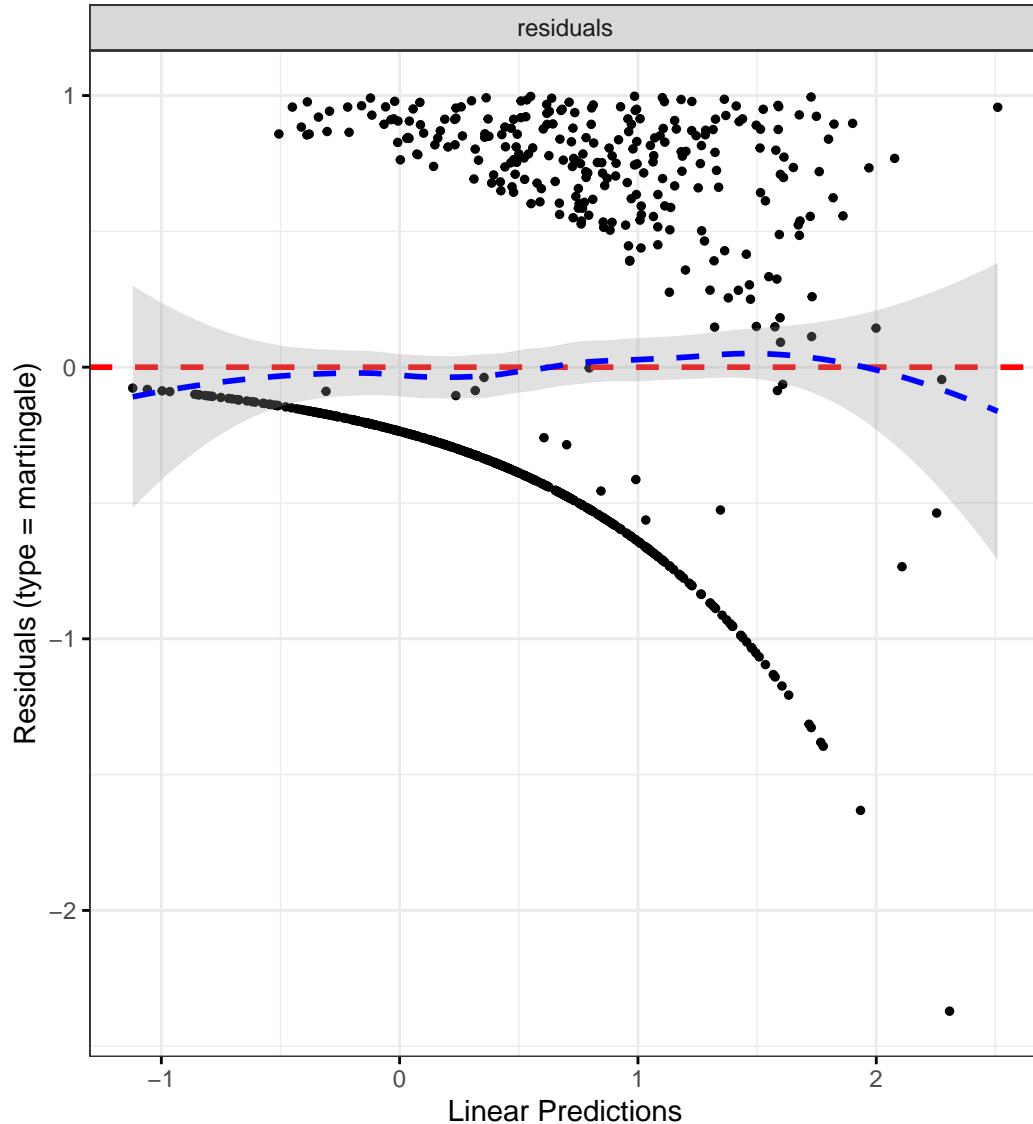
	Chi-squared statistic	DF	P-value
IMD	3.3195	3.9300	0.4949
cat	5.2289	1.9633	0.0708
IBD Duration	0.5297	0.9809	0.4592
BMI	1.3185	0.9781	0.2448
Treatment	7.1686	3.8575	0.1173
Age	0.0016	0.9689	0.9639
WCC	0.1649	0.9865	0.6792
GLOBAL	18.7212	24.9537	0.8084

## DF betas

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



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



Platelets

```
fit.me <- coxph(
  Surv(softflare_time, softflare) ~
    Sex +
    IMD +
    cat +
    `IBD Duration` +
    BMI +
    Treatment +
```

```

Age +
Platelets +
frailty(SiteNo),
control = coxph.control(outer.max = 20),
data = flare.uc.df
)

invisible(cox_summary(fit.me))

```

Cox model summary:

Variable	HR	Lower 95%	Upper 95%	P-value
SexFemale	1.5951	1.2326	2.0643	0.0004
IMD2	1.1802	0.7088	1.9653	0.5242
IMD3	0.8286	0.4956	1.3854	0.4735
IMD4	1.0665	0.6564	1.7328	0.7947
IMD5	0.9915	0.6188	1.5886	0.9715
catFC 50-250	1.8077	1.3477	2.4247	0.0001
catFC > 250	2.3952	1.7567	3.2657	0.0000
IBD Duration	0.9931	0.9794	1.0070	0.3287
BMI	0.9828	0.9586	1.0077	0.1736
TreatmentMono biologic	0.9082	0.5666	1.4558	0.6891
TreatmentCombo therapy	0.6641	0.3523	1.2518	0.2057
Treatment5-ASA	1.5019	1.0284	2.1935	0.0353
TreatmentNone reported	1.2726	0.8633	1.8758	0.2234
Age	0.9924	0.9829	1.0020	0.1218
Platelets	0.9993	0.9974	1.0012	0.4961

Diagnostics:

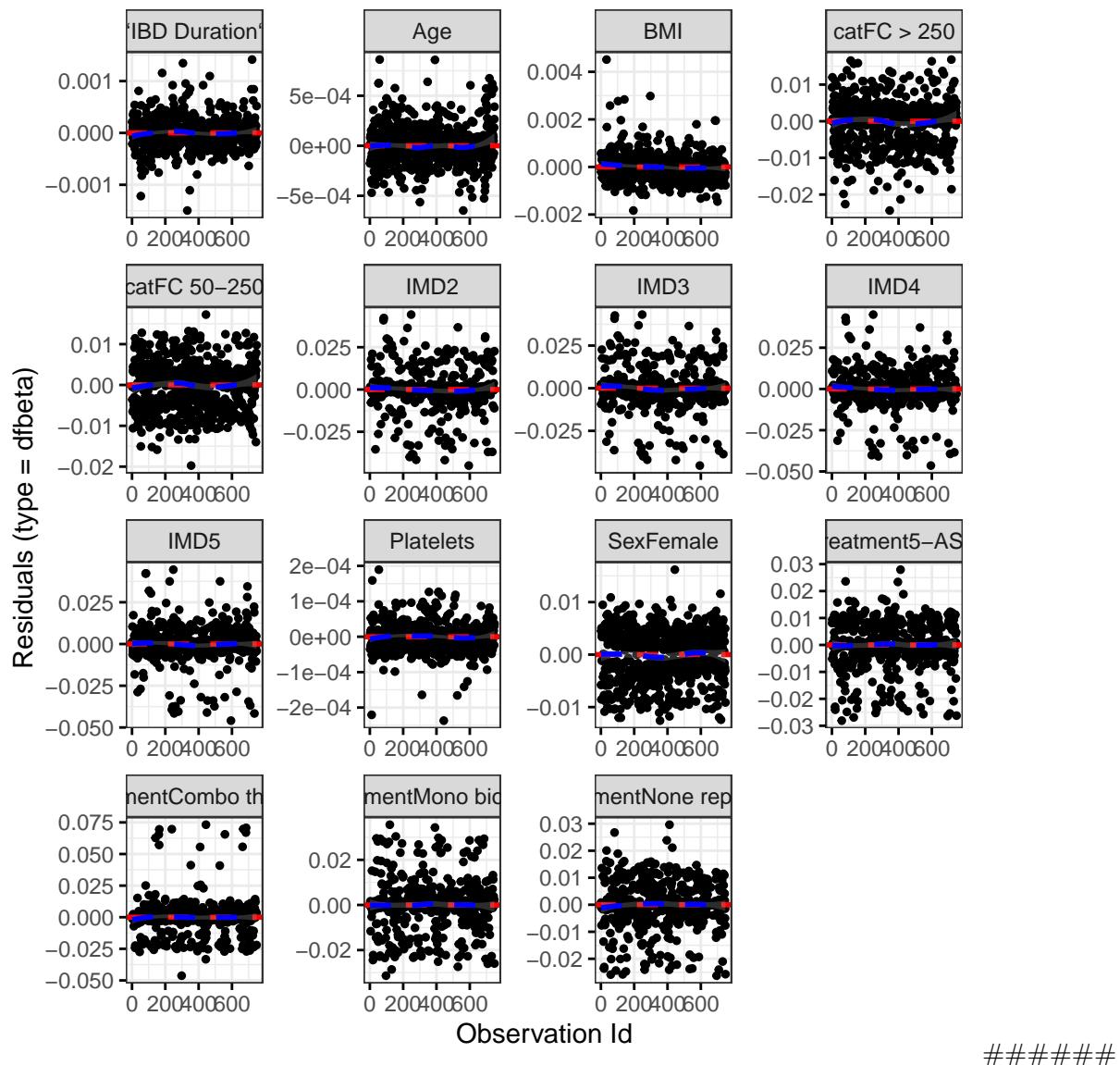
### Proportional hazards assumption test

	Chi-squared statistic	DF	P-value
Sex	1.4531	0.9846	0.2240
IMD	3.5679	3.9306	0.4570
cat	5.5454	1.9643	0.0604
IBD Duration	0.8621	0.9807	0.3465
BMI	1.7223	0.9790	0.1847
Treatment	7.3289	3.8532	0.1097

	Chi-squared statistic	DF	P-value
Age	0.0070	0.9651	0.9264
Platelets	0.6136	0.9833	0.4270
GLOBAL	20.3604	25.3193	0.7431

## DF betas

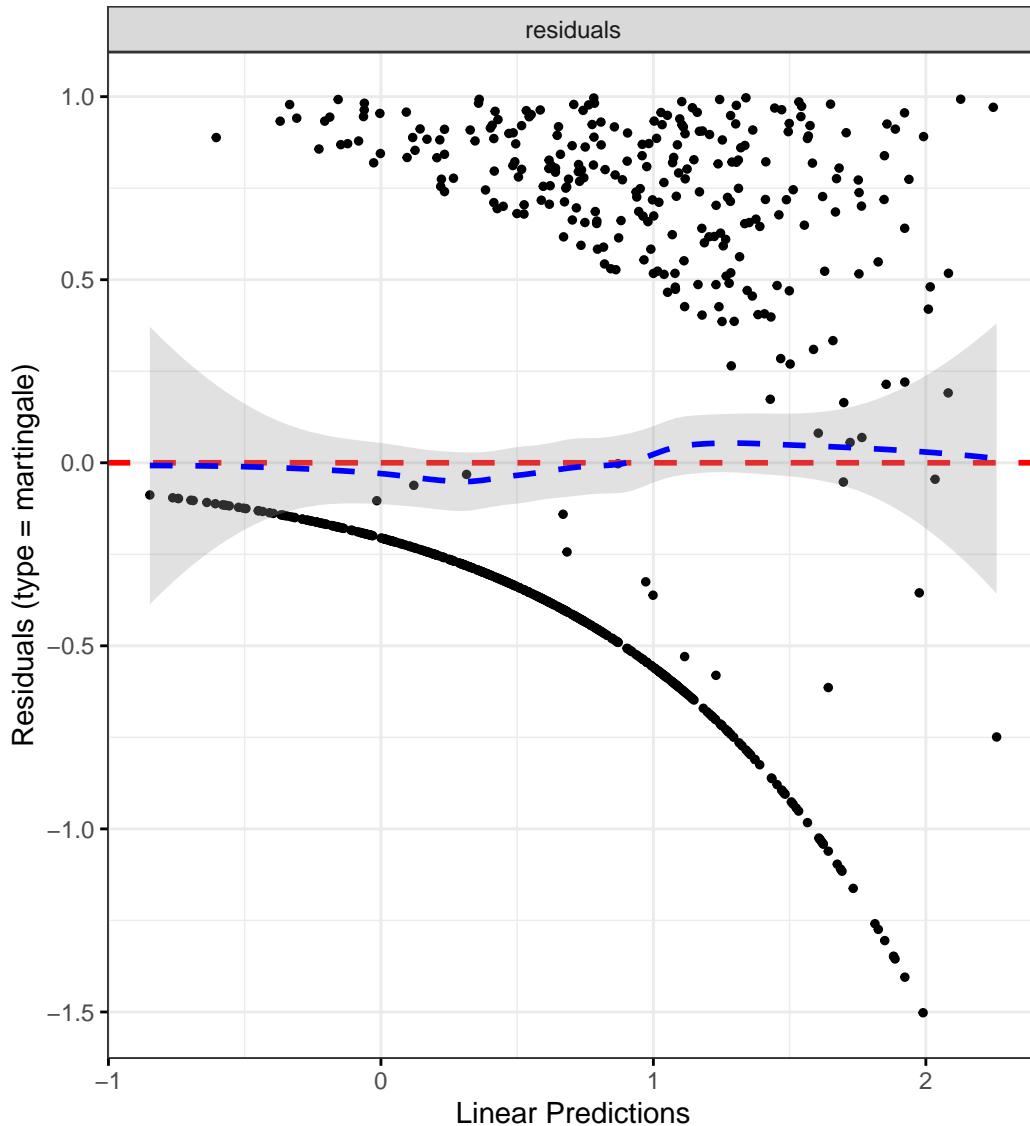
`geom\_smooth()` using formula = 'y ~ x'



#####

Martingale residuals

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



Albumin

```
fit.me <- coxph(  
  Surv(softflare_time, softflare) ~  
    Sex +
```

```

IMD +
cat +
`IBD Duration` +
BMI +
Treatment +
Age +
Albumin +
frailty(SiteNo),
control = coxph.control(outer.max = 20),
data = flare.uc.df
)

invisible(cox_summary(fit.me))

```

Cox model summary:

Variable	HR	Lower 95%	Upper 95%	P-value
SexFemale	1.4662	1.1319	1.8993	0.0037
IMD2	1.1015	0.6552	1.8518	0.7153
IMD3	0.8273	0.4951	1.3826	0.4694
IMD4	1.0695	0.6596	1.7341	0.7853
IMD5	0.9614	0.5981	1.5452	0.8707
catFC 50-250	1.8467	1.3706	2.4880	0.0001
catFC > 250	2.2264	1.6238	3.0526	0.0000
IBD Duration	0.9916	0.9777	1.0057	0.2437
BMI	0.9823	0.9582	1.0071	0.1607
TreatmentMono biologic	0.9426	0.5786	1.5356	0.8123
TreatmentCombo therapy	0.6627	0.3437	1.2777	0.2193
Treatment5-ASA	1.6021	1.0822	2.3718	0.0185
TreatmentNone reported	1.4166	0.9499	2.1126	0.0877
Age	0.9908	0.9812	1.0006	0.0646
Albumin	0.9766	0.9517	1.0022	0.0727

Diagnostics:

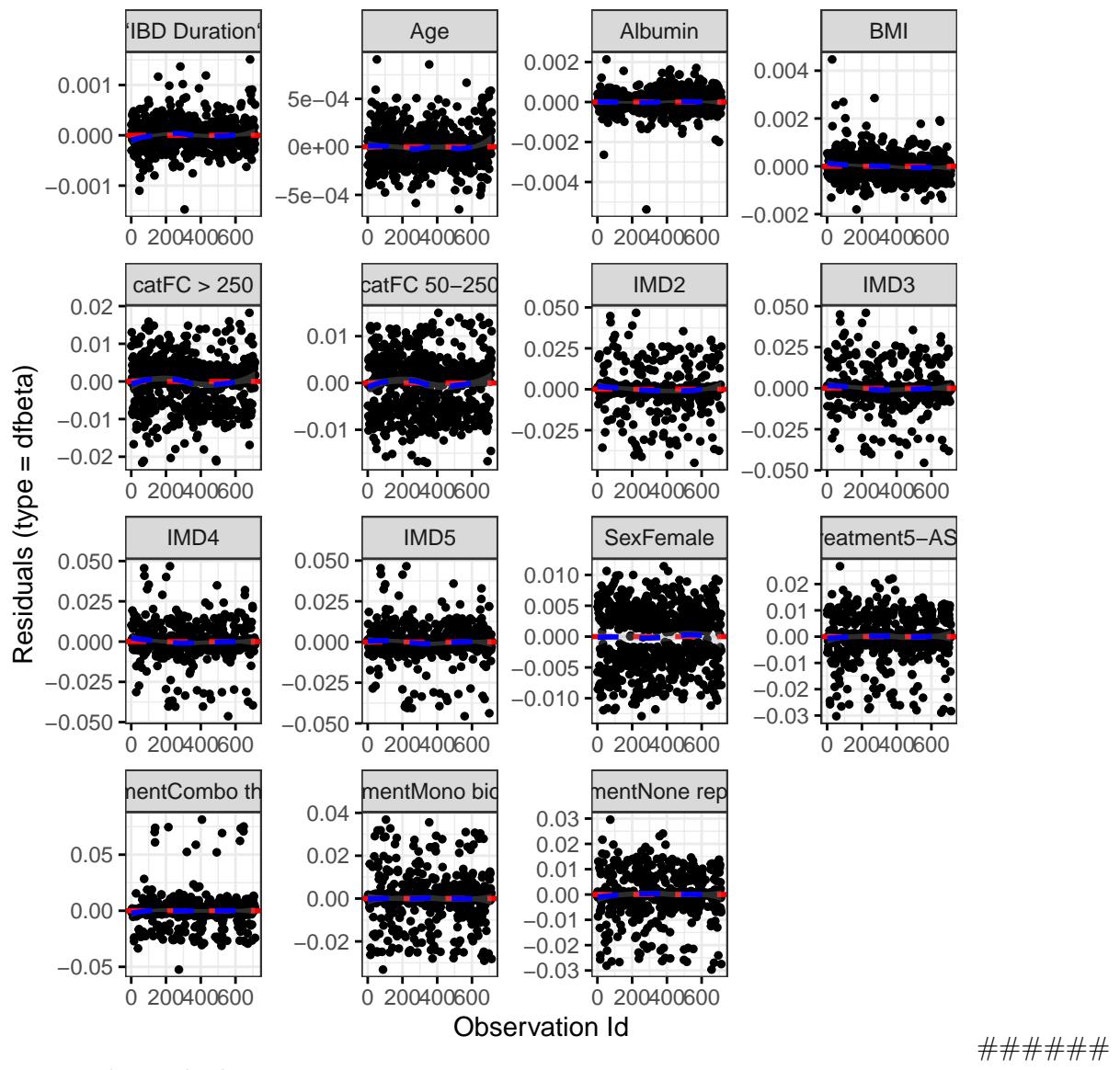
### Proportional hazards assumption test

	Chi-squared statistic	DF	P-value
Sex	1.3158	0.9827	0.2465

	Chi-squared statistic	DF	P-value
IMD	4.8742	3.9334	0.2920
cat	4.4353	1.9652	0.1057
IBD Duration	0.8682	0.9806	0.3448
BMI	2.4009	0.9786	0.1179
Treatment	6.3200	3.8541	0.1635
Age	0.1203	0.9629	0.7139
Albumin	0.3638	0.8944	0.5012
GLOBAL	20.9351	25.2380	0.7083

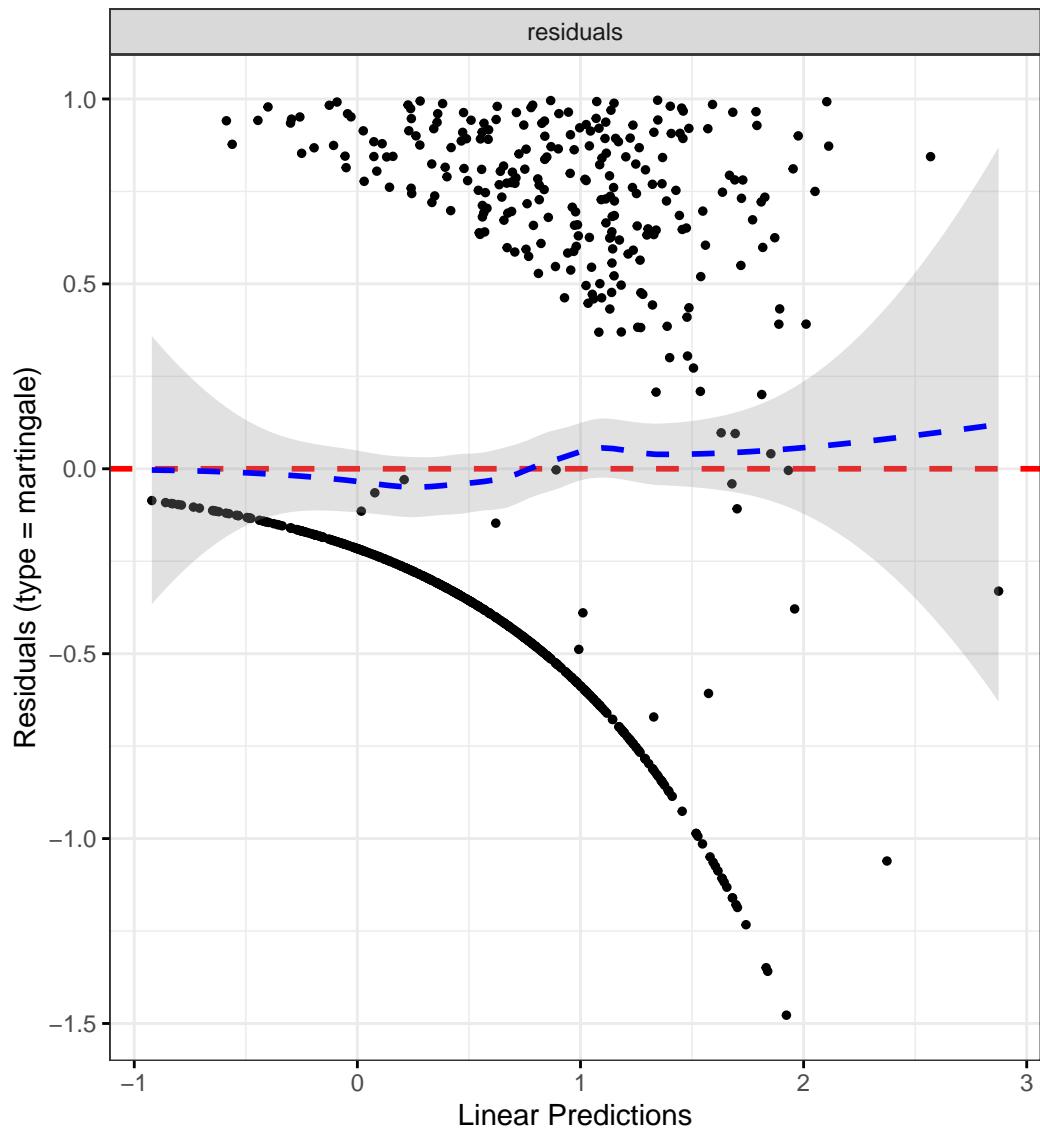
## DF betas

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



Martingale residuals

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



Objective flare

C-reactive protein

```
fit.me <- coxph(
  Surv(hardflare_time, hardflare) ~
    Sex +
    IMD +
    cat +
    `IBD Duration` +
```

```

    BMI +
Treatment +
Age +
CReactiveProtein +
frailty(SiteNo),
control = coxph.control(outer.max = 20),
data = flare.uc.df
)

invisible(cox_summary(fit.me))

```

Cox model summary:

Variable	HR	Lower 95%	Upper 95%	P-value
SexFemale	1.3266	0.9825	1.7911	0.0651
IMD2	1.6258	0.8009	3.3003	0.1785
IMD3	1.4117	0.6999	2.8476	0.3354
IMD4	2.1013	1.0693	4.1295	0.0312
IMD5	1.7372	0.8892	3.3939	0.1060
catFC 50-250	2.1419	1.4976	3.0633	0.0000
catFC > 250	2.9398	2.0222	4.2739	0.0000
IBD Duration	0.9893	0.9714	1.0075	0.2471
BMI	1.0101	0.9818	1.0393	0.4877
TreatmentMono biologic	1.1243	0.6801	1.8585	0.6479
TreatmentCombo therapy	1.0275	0.5445	1.9387	0.9333
Treatment5-ASA	0.9580	0.6161	1.4898	0.8491
TreatmentNone reported	0.8990	0.5689	1.4206	0.6484
Age	0.9880	0.9763	0.9998	0.0456
CReactiveProtein	0.9959	0.9751	1.0172	0.7067

Diagnostics:

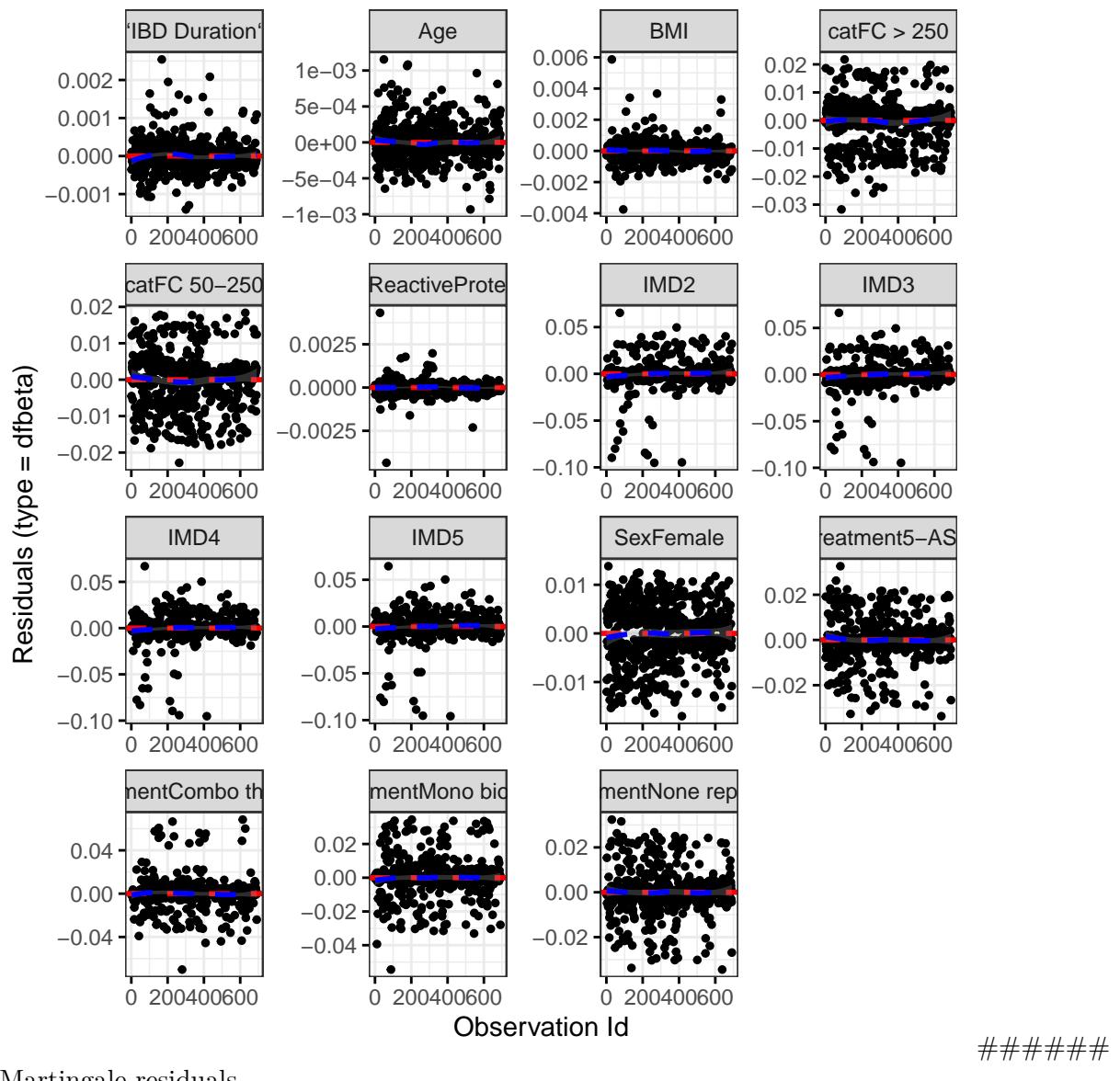
### Proportional hazards assumption test

	Chi-squared statistic	DF	P-value
Sex	0.4382	0.9869	0.5026
IMD	1.5345	3.9327	0.8127
cat	3.9063	1.9666	0.1380
IBD Duration	1.6503	0.9778	0.1937

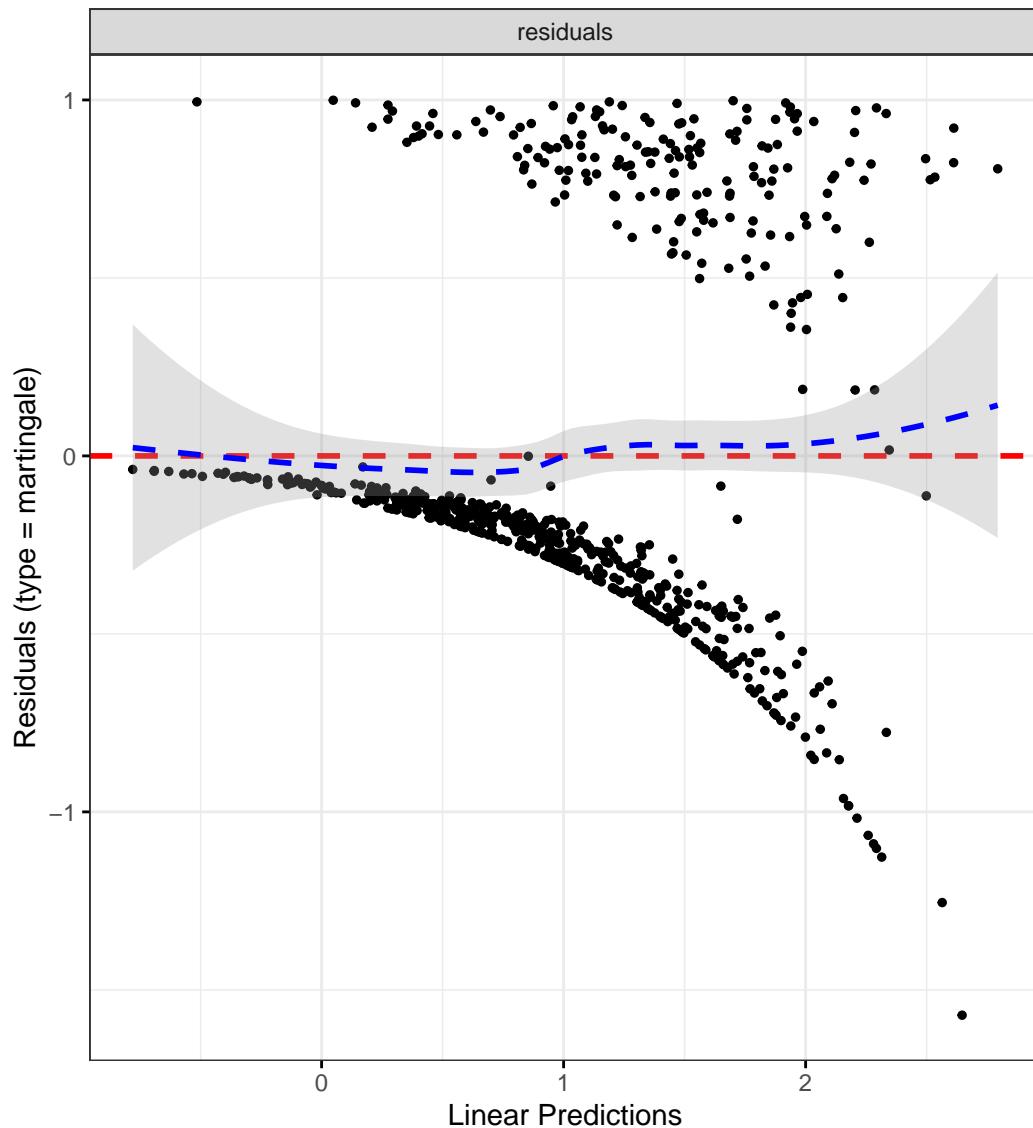
	Chi-squared statistic	DF	P-value
BMI	0.0762	0.9838	0.7765
Treatment	20.9558	3.8287	0.0003
Age	0.0766	0.9712	0.7713
CRactiveProtein	0.0002	0.9887	0.9873
GLOBAL	28.7465	25.5184	0.2992

## DF betas

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



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



Haemoglobin

```
fit.me <- coxph(
  Surv(hardflare_time, hardflare) ~
    Sex +
    IMD +
    cat +
    `IBD Duration` +
    BMI +
    Treatment +
```

```

Age +
Haemoglobin +
frailty(SiteNo),
control = coxph.control(outer.max = 20),
data = flare.uc.df
)

invisible(cox_summary(fit.me))

```

Cox model summary:

Variable	HR	Lower 95%	Upper 95%	P-value
SexFemale	1.2689	0.8942	1.8008	0.1823
IMD2	1.3100	0.6852	2.5045	0.4142
IMD3	1.1036	0.5791	2.1031	0.7645
IMD4	1.5665	0.8432	2.9101	0.1555
IMD5	1.4111	0.7667	2.5971	0.2685
catFC 50-250	2.0196	1.4142	2.8842	0.0001
catFC > 250	2.8997	2.0025	4.1989	0.0000
IBD Duration	0.9905	0.9734	1.0079	0.2818
BMI	1.0013	0.9734	1.0300	0.9288
TreatmentMono biologic	1.0846	0.6570	1.7905	0.7509
TreatmentCombo therapy	1.0638	0.5668	1.9966	0.8474
Treatment5-ASA	0.9868	0.6377	1.5271	0.9525
TreatmentNone reported	0.8806	0.5583	1.3889	0.5845
Age	0.9895	0.9780	1.0012	0.0772
Haemoglobin	0.9914	0.9783	1.0046	0.2013

Diagnostics:

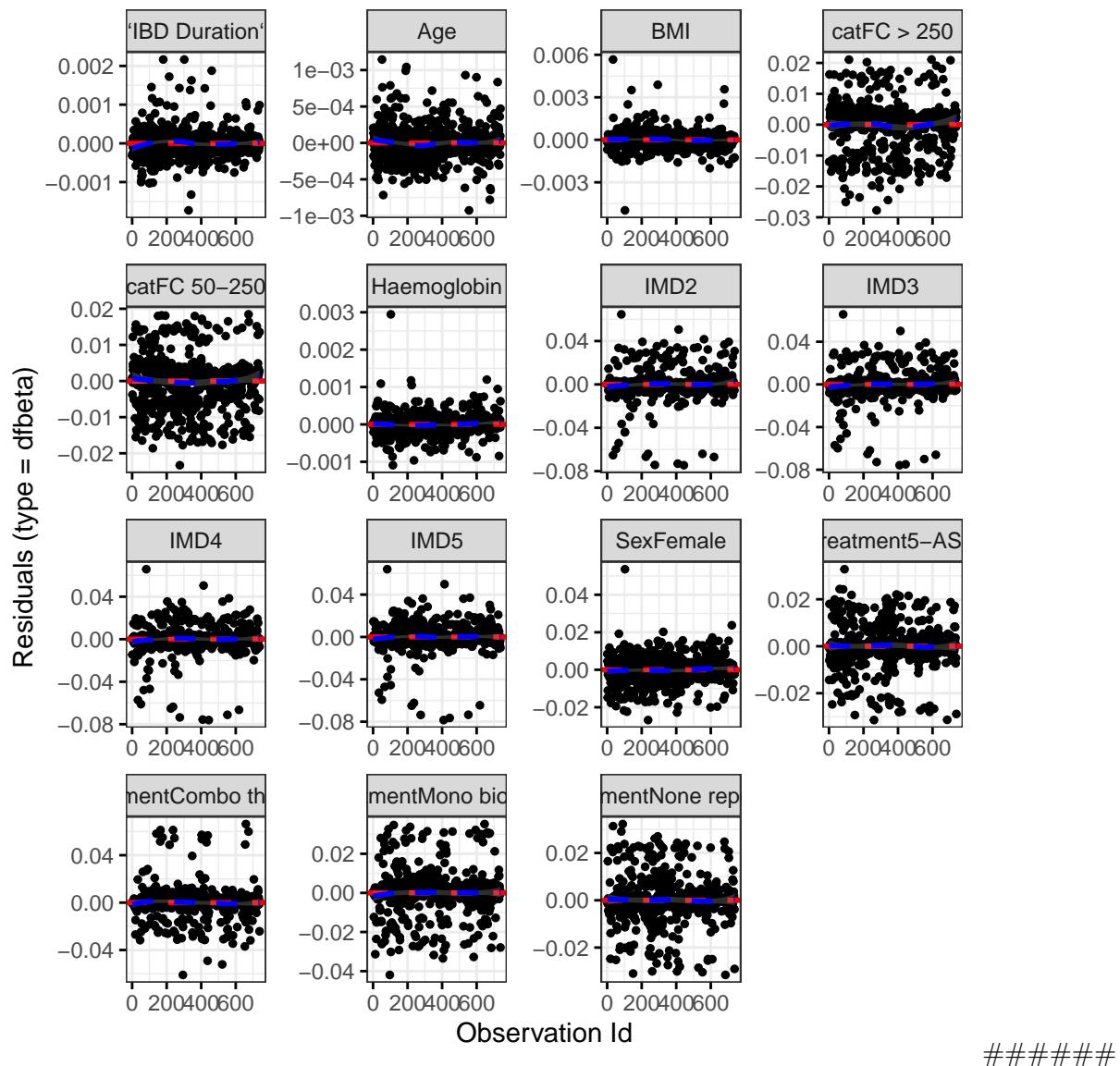
### Proportional hazards assumption test

	Chi-squared statistic	DF	P-value
Sex	0.0847	0.9863	0.7659
IMD	2.2329	3.9344	0.6833
cat	5.3762	1.9709	0.0662
IBD Duration	1.2156	0.9806	0.2646
BMI	0.0334	0.9816	0.8493
Treatment	17.1476	3.8340	0.0015

	Chi-squared statistic	DF	P-value
Age	0.0015	0.9713	0.9654
Haemoglobin	0.9658	0.9862	0.3212
GLOBAL	25.4426	24.4328	0.4060

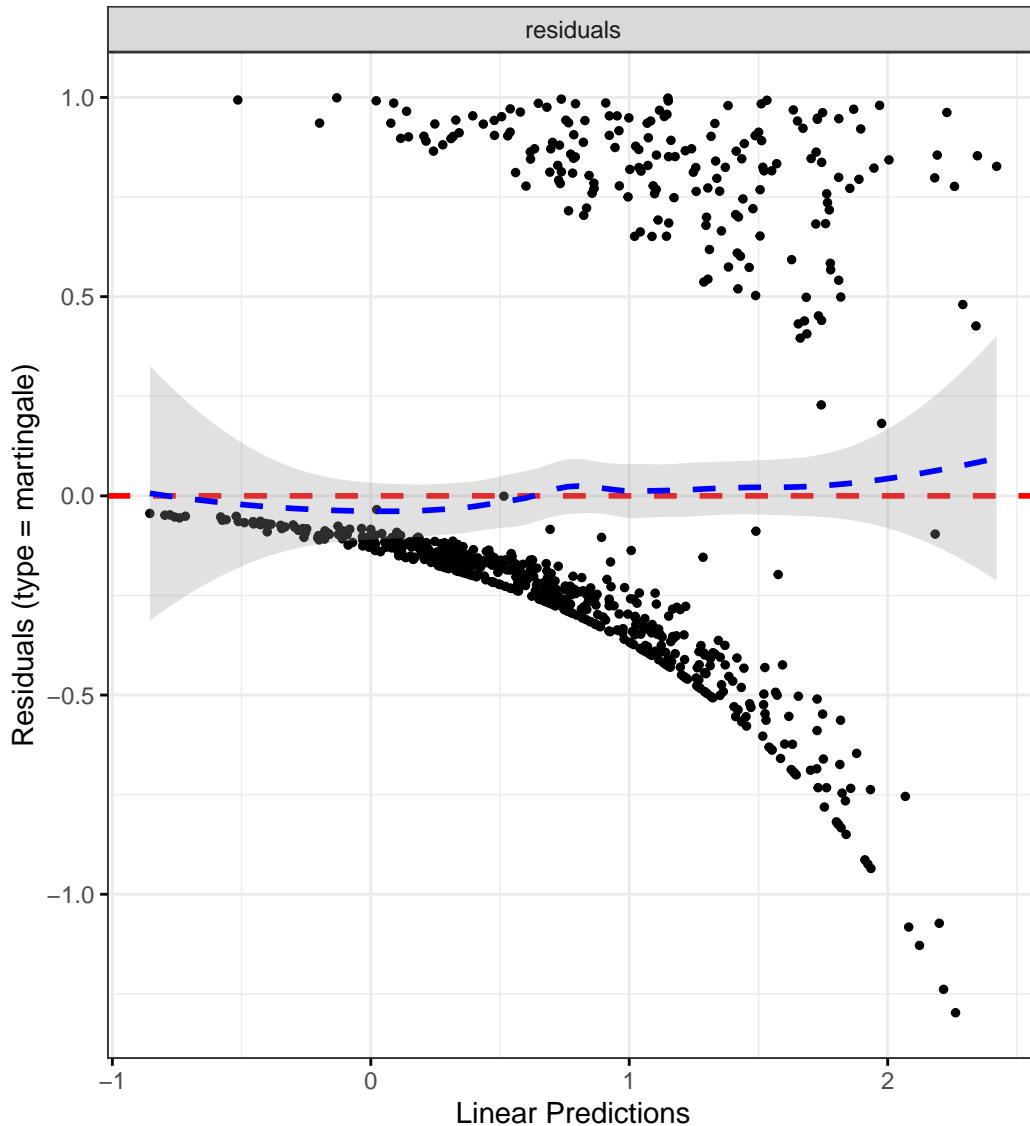
## DF betas

`geom\_smooth()` using formula = 'y ~ x'



Martingale residuals

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



White cell count

```
fit.me <- coxph(  
  Surv(hardflare_time, hardflare) ~  
    Sex +
```

```

IMD +
cat +
`IBD Duration` +
BMI +
Treatment +
Age +
WCC +
frailty(SiteNo),
control = coxph.control(outer.max = 20),
data = flare.uc.df
)

invisible(cox_summary(fit.me))

```

Cox model summary:

Variable	HR	Lower 95%	Upper 95%	P-value
SexFemale	1.4318	1.0619	1.9306	0.0186
IMD2	1.3293	0.6948	2.5434	0.3899
IMD3	1.0815	0.5683	2.0582	0.8113
IMD4	1.5146	0.8163	2.8104	0.1880
IMD5	1.3903	0.7574	2.5522	0.2876
catFC 50-250	2.0087	1.4056	2.8707	0.0001
catFC > 250	2.8801	1.9864	4.1760	0.0000
IBD Duration	0.9906	0.9735	1.0081	0.2922
BMI	0.9952	0.9665	1.0247	0.7465
TreatmentMono biologic	1.0271	0.6191	1.7042	0.9174
TreatmentCombo therapy	1.0308	0.5497	1.9330	0.9246
Treatment5-ASA	0.8807	0.5610	1.3826	0.5809
TreatmentNone reported	0.8071	0.5064	1.2863	0.3674
Age	0.9908	0.9792	1.0026	0.1264
WCC	1.0583	0.9774	1.1459	0.1626

Diagnostics:

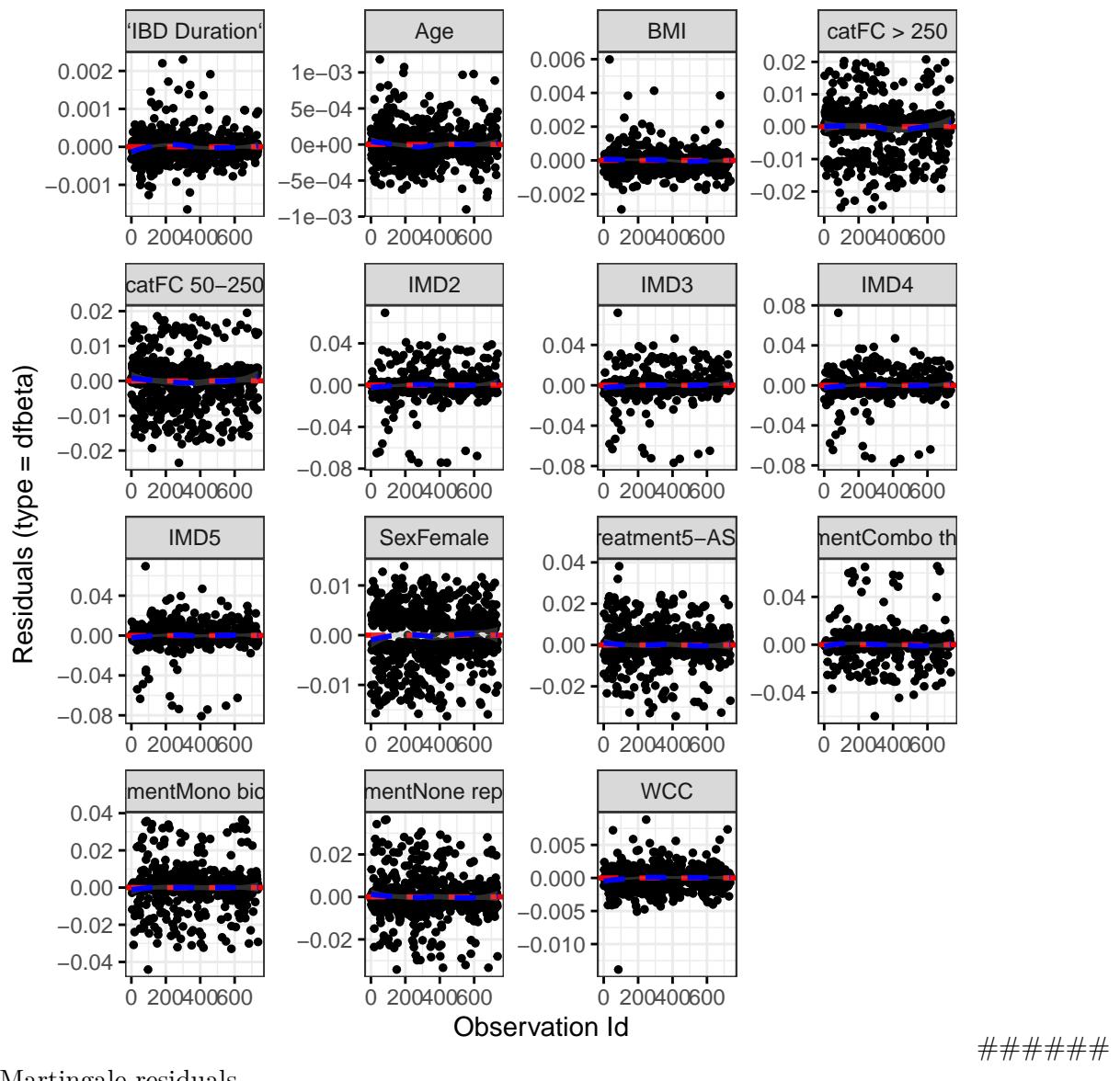
### Proportional hazards assumption test

	Chi-squared statistic	DF	P-value
Sex	0.1295	0.9865	0.7136

	Chi-squared statistic	DF	P-value
IMD	2.3079	3.9352	0.6697
cat	5.3490	1.9719	0.0672
IBD Duration	1.1250	0.9814	0.2832
BMI	0.0388	0.9817	0.8381
Treatment	17.2058	3.8387	0.0015
Age	0.0023	0.9732	0.9577
WCC	0.5024	0.9902	0.4745
GLOBAL	25.1507	24.2412	0.4111

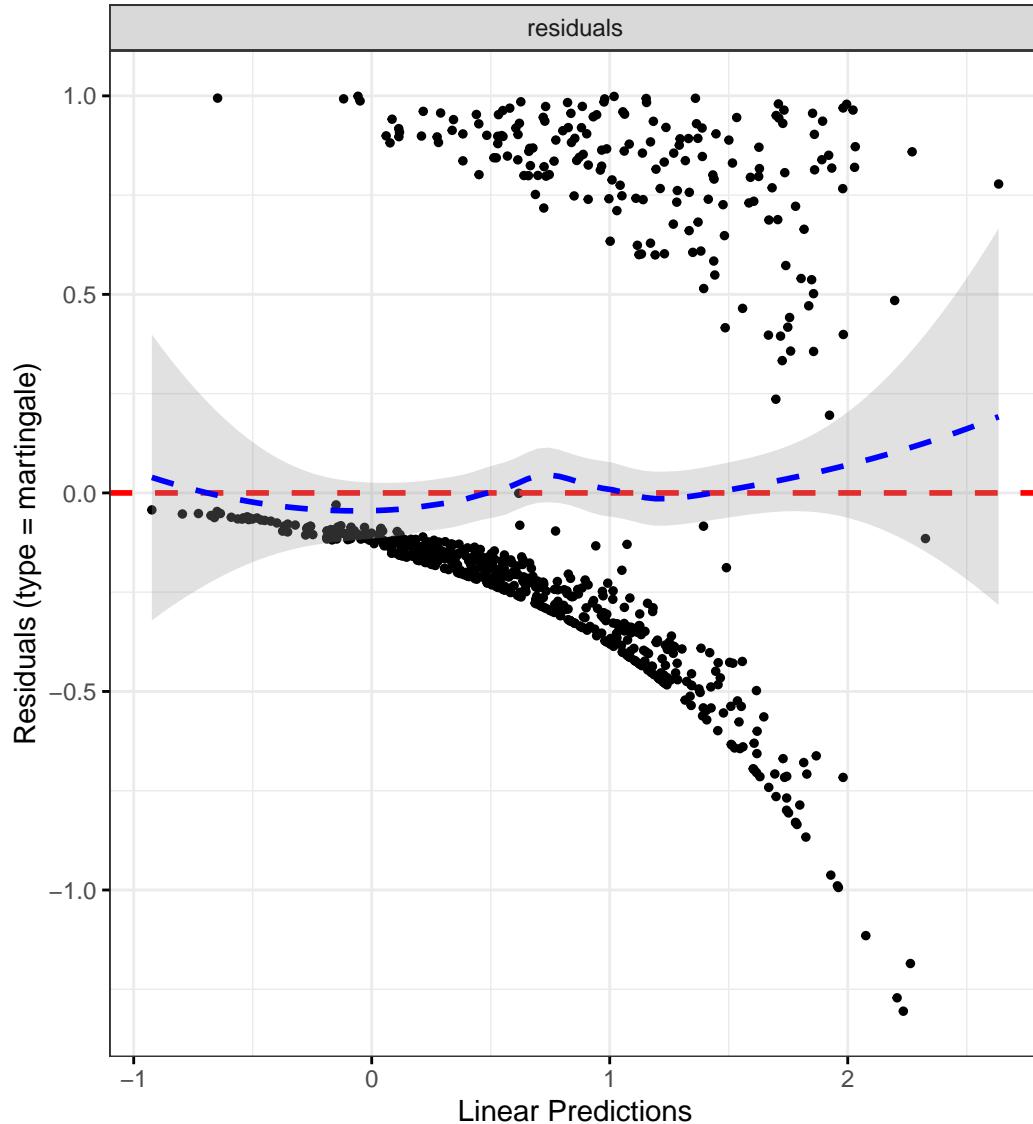
## DF betas

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



Martingale residuals

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



Platelets

```
fit.me <- coxph(
  Surv(hardflare_time, hardflare) ~
    Sex +
    IMD +
    cat +
    `IBD Duration` +
    BMI +
    Treatment +
```

```

Age +
Platelets +
frailty(SiteNo),
control = coxph.control(outer.max = 20),
data = flare.uc.df
)

invisible(cox_summary(fit.me))

```

Cox model summary:

Variable	HR	Lower 95%	Upper 95%	P-value
SexFemale	1.4532	1.0721	1.9698	0.0160
IMD2	1.2987	0.6789	2.4845	0.4297
IMD3	1.0943	0.5740	2.0864	0.7843
IMD4	1.5309	0.8254	2.8393	0.1766
IMD5	1.3676	0.7449	2.5111	0.3126
catFC 50-250	2.0389	1.4260	2.9153	0.0001
catFC > 250	3.0423	2.1047	4.3976	0.0000
IBD Duration	0.9911	0.9740	1.0085	0.3158
BMI	0.9997	0.9715	1.0288	0.9863
TreatmentMono biologic	1.0937	0.6636	1.8026	0.7254
TreatmentCombo therapy	1.0431	0.5559	1.9573	0.8954
Treatment5-ASA	0.9731	0.6298	1.5035	0.9023
TreatmentNone reported	0.8662	0.5492	1.3661	0.5366
Age	0.9896	0.9781	1.0014	0.0835
Platelets	0.9996	0.9974	1.0018	0.7241

Diagnostics:

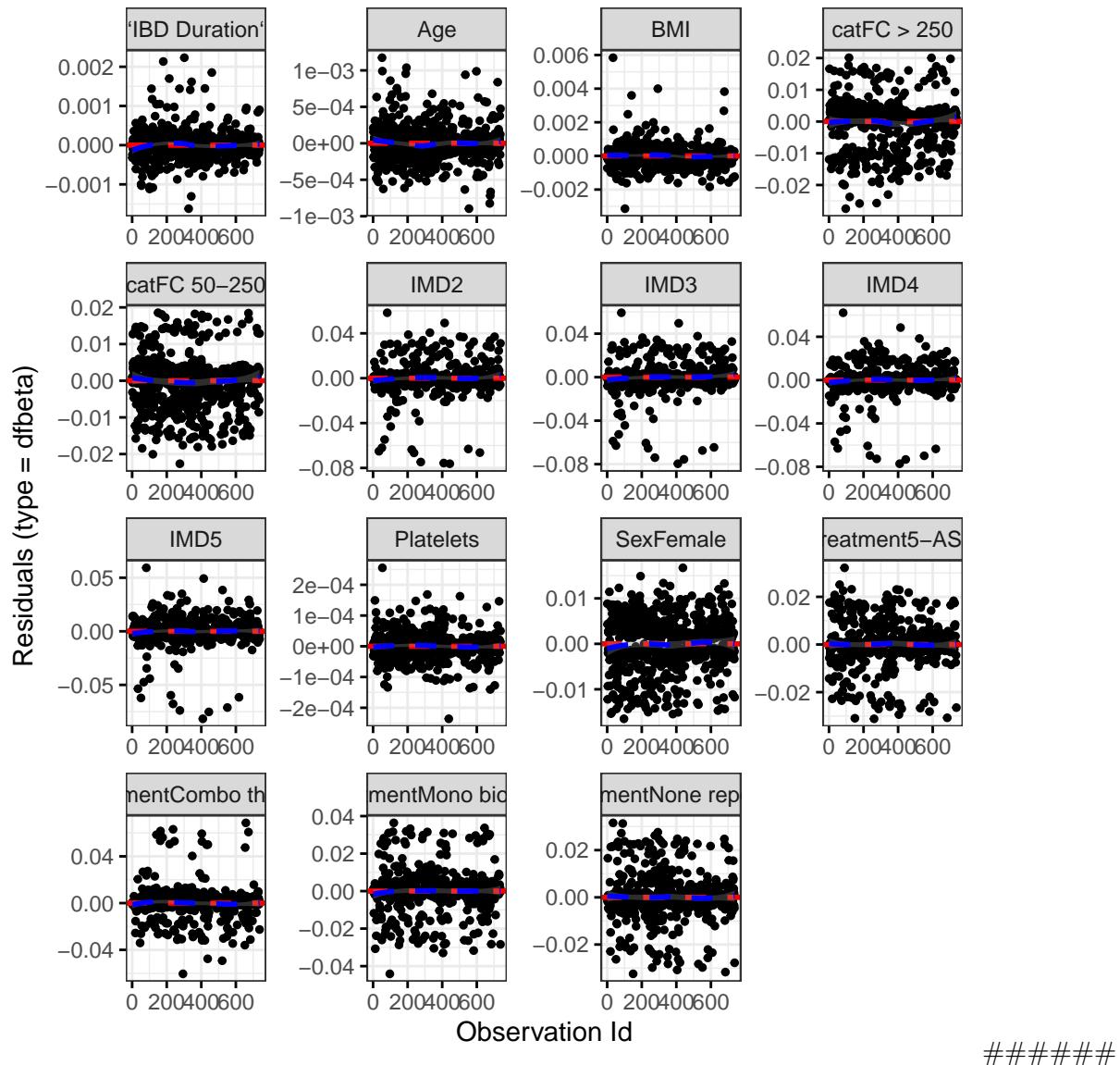
### Proportional hazards assumption test

	Chi-squared statistic	DF	P-value
Sex	0.1022	0.9860	0.7438
IMD	2.3673	3.9361	0.6589
cat	5.2906	1.9710	0.0691
IBD Duration	1.1990	0.9813	0.2680
BMI	0.0451	0.9823	0.8260
Treatment	17.1265	3.8341	0.0016

	Chi-squared statistic	DF	P-value
Age	0.0003	0.9728	0.9843
Platelets	0.0474	0.9891	0.8240
GLOBAL	24.6162	24.3181	0.4450

## DF betas

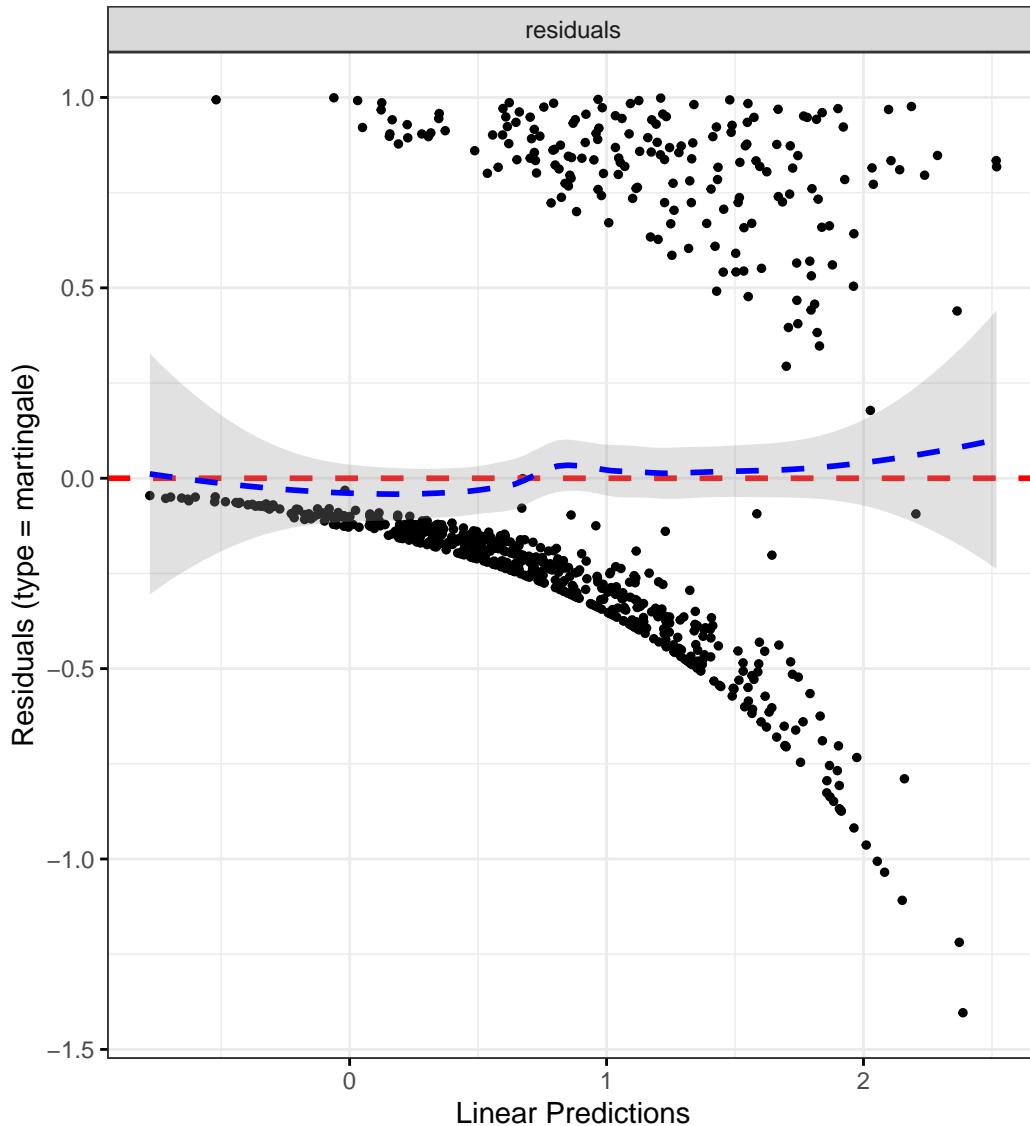
`geom\_smooth()` using formula = 'y ~ x'



#####

Martingale residuals

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



Albumin

```
fit.me <- coxph(  
  Surv(hardflare_time, hardflare) ~  
    Sex +
```

```

IMD +
cat +
`IBD Duration` +
BMI +
Treatment +
Age +
Albumin +
frailty(SiteNo),
control = coxph.control(outer.max = 20),
data = flare.uc.df
)

invisible(cox_summary(fit.me))

```

Cox model summary:

Variable	HR	Lower 95%	Upper 95%	P-value
SexFemale	1.3512	0.9926	1.8392	0.0557
IMD2	1.2052	0.6216	2.3369	0.5807
IMD3	1.0665	0.5599	2.0314	0.8447
IMD4	1.4742	0.7950	2.7335	0.2180
IMD5	1.3472	0.7332	2.4755	0.3370
catFC 50-250	2.0455	1.4137	2.9598	0.0001
catFC > 250	2.9342	2.0117	4.2799	0.0000
IBD Duration	0.9911	0.9736	1.0090	0.3271
BMI	0.9931	0.9645	1.0226	0.6435
TreatmentMono biologic	1.2102	0.7272	2.0137	0.4629
TreatmentCombo therapy	1.0523	0.5504	2.0119	0.8775
Treatment5-ASA	0.9737	0.6178	1.5346	0.9086
TreatmentNone reported	0.8889	0.5559	1.4214	0.6228
Age	0.9897	0.9779	1.0016	0.0893
Albumin	0.9859	0.9581	1.0146	0.3317

Diagnostics:

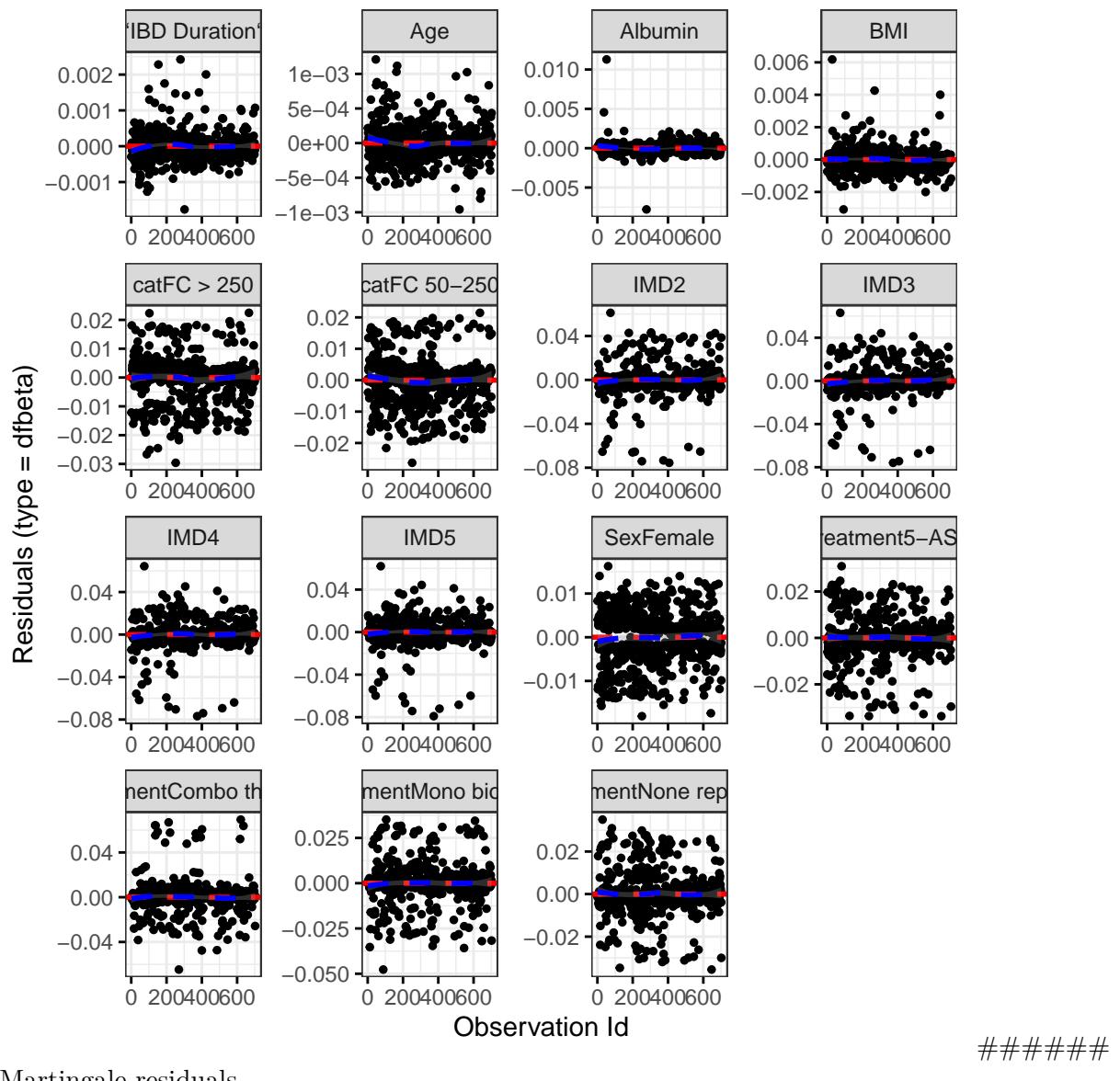
### Proportional hazards assumption test

	Chi-squared statistic	DF	P-value
Sex	0.2869	0.9845	0.5857

	Chi-squared statistic	DF	P-value
IMD	2.6727	3.9387	0.6045
cat	4.2734	1.9711	0.1152
IBD Duration	1.1940	0.9806	0.2688
BMI	0.0146	0.9805	0.8990
Treatment	18.4072	3.8363	0.0009
Age	0.0261	0.9721	0.8636
Albumin	0.2169	0.9193	0.6067
GLOBAL	26.1516	24.2769	0.3603

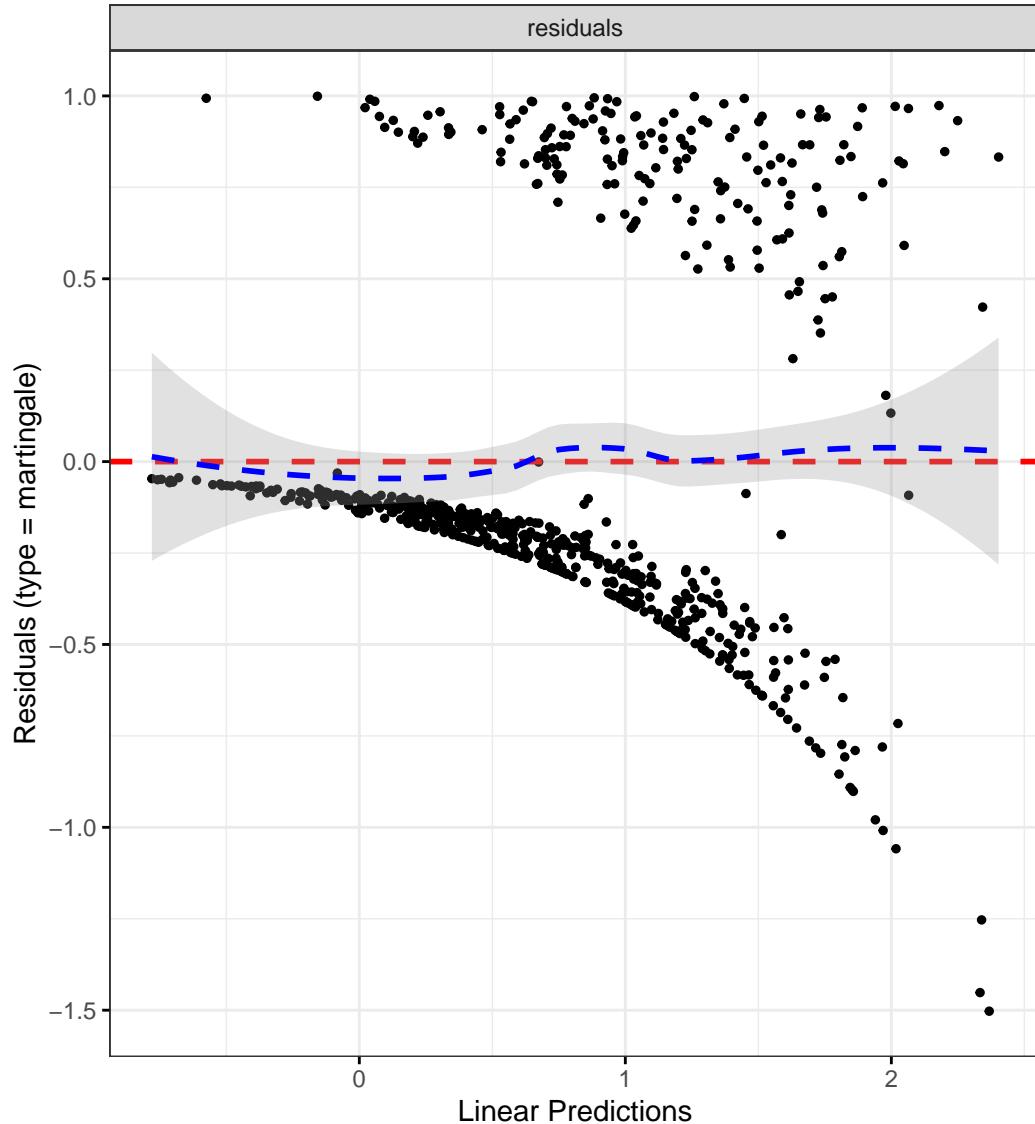
## DF betas

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



Martingale residuals

```
`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"))
```