

Between-centre differences for COVID-19 ICU mortality from early data in England

ANALYSIS SUPPLEMENT

Zhaozhi Qian¹, Ahmed M. Alaa², Mihaela van der Schaar¹, and Ari Ercole^{3,*}

¹University of Cambridge Centre for Mathematical Sciences, Wilberforce Rd, Cambridge CB3 0WA, UK

²University of California, Los Angeles, CA 90095, USA

³University of Cambridge Division of Anaesthesia, Addenbrooke's Hospital, Hills Road, Cambridge CB2 0QQ, UK

*Corresponding author

This file contains the code from the analysis as canonical documentation of the pre-processing and modelling assumptions.

Libraries used (latest versions as of 4/2020)

```
> library(ggplot2)
```

```
> library(dplyr)
```

```
> library(reshape2)
```

```
> library(coxme)
```

```
> library(tidyr)
```

Data cleaning specification

```
> update_date = '14/4/2020'
```

```
> update_dt = as.Date('14/4/2020', format="%d/%m/%Y")
```

```
> dat = read.csv(file='data/letter_data.csv', stringsAsFactors = FALSE)
```

```
> names(dat) = tolower(names(dat))
```

```
> dat = dat %>%
```

```
+ mutate(dateupdated = gsub(".*$", "", dateupdated)) %>%
```

```
+ mutate_at(.vars=c('dateupdated', 'hospitaladmissiondate', 'dateadmittedicu', 'dateleavingicu', 'finaloutcomeate'),
```

```
+ .funs = function(x) as.Date(x, format="%m/%d/%Y"))
```

```

> dt = dat$dateupdated

> dt[is.na(dt)] = update_dt

> dat$dateupdated = dt

> dat = dat %>%

+   mutate(hospital_ttc = as.numeric(dateupdated - hospitaladmissiondate),

+     icu_ttc = as.numeric(dateupdated - dateadmittedicu),

+     outcome_time = as.numeric(finaloutcomedate - hospitaladmissiondate),

+     time_to_icu = as.numeric(dateadmittedicu - hospitaladmissiondate),

+     icu_duration = as.numeric(dateleavingicu - dateadmittedicu)

+   ) %>%

+   mutate(finaloutcome = ifelse(finaloutcome == "Transferred", "", finaloutcome)) %>%

+   mutate_at(.vars=c('dateupdated', 'hospitaladmissiondate', 'dateadmittedicu', 'finaloutcomedate', 'dateleavingicu'),

+     .funs = as.character) %>%

+   mutate_all(.funs=function(x) x = ifelse((x == 'Unknown') | (x == 'No') | (x == 'Borderline'), "", x)) %>%

+   mutate(sex = ifelse(sex == 'Male', 'Male', "")) %>%

+   mutate(

+     death_ind = finaloutcome == 'Death',

+     icu_ind = !is.na(time_to_icu),

+     discharge_ind = finaloutcome == 'Discharged',

+     death_time = coalesce(outcome_time, hospital_ttc),

+     icu_time = coalesce(time_to_icu, hospital_ttc),

+     discharge_time = coalesce(outcome_time, hospital_ttc),

+   ) %>% group_by(caseid) %>% filter(row_number(caseid) == 1) %>% ungroup() %>%

+   filter(wasthepatientadmittedtoicu == 'Yes')

> # there are 3092 icu patients

```

```

> nrow(dat)

[1] 3092

> # death 609

> dat %>%

+ filter(finaloutcome == 'Death') %>%

+ summarise(n=n())

# A tibble: 1 x 1

  n

<int>

1 609

> # ICU discharge 384

> dat %>%

+ filter(!is.na(dateleavingicu) & finaloutcome != 'Death') %>%

+ summarise(n=n())

# A tibble: 1 x 1

  n

<int>

1 384

> dat = dat %>%

+ mutate(

+   ageyear = ifelse((is.na(ageyear) | ageyear == 0), median(ageyear, na.rm=TRUE), ageyear),

+   sex = replace_na(sex, 'Male'),

+ )

> dat$age = dat$ageyear

> dat$ageyear = (dat$ageyear - mean(dat$ageyear)) / sd(dat$ageyear)

> nrow(dat)

```

[1] 3092

```
> trust_scores = read.csv('data/letter_trusts_linked.csv', stringsAsFactors = FALSE)
> apache = read.csv('data/apache.csv', stringsAsFactors = FALSE)
> apache = apache %>%
+   group_by(hospname) %>%
+   summarise(apache = median(apache))
> dat2 = dat %>%
+   inner_join(trust_scores, by='trustname')%>%
+   inner_join(apache, by='hospname')
> dat2$apache_mean = (dat2$apache - mean(dat2$apache)) / sd(dat2$apache)
> dat2$agecut = cut(dat2$age, c(0, 50, 60, 70, 80, 999))
```

Define and fit Cox model

```
> model = coxme(Surv(death_time , death_ind) ~
+   agecut +
+   sex +
+   chonicrespiratory+
+   asthmarequiring+
+   chronicheart +
+   chronicrenal +
+   chronicliver +
+   chronicneurological +
+   isdiabetes +
+   immunosuppressiontreatment +
+   immunosuppressiondisease +
```

```

+ obesityclinical +
+ hypertension +
+ apache_mean +
+ (1|trustcode), dat2 )

```

```
> summary(model)
```

Cox mixed-effects model fit by maximum likelihood

Data: dat2

events, n = 605, 3080

Iterations= 13 110

NULL Integrated Fitted

Log-likelihood -4477.025 -4161.828 -4064.469

	Chisq	df	p	AIC	BIC
Integrated loglik	630.39	18.00	0	594.39	515.10
Penalized loglik	825.11	80.34	0	664.44	310.53

Model: Surv(death_time, death_ind) ~ agecut + sex + chonicrespiratory + asthmarequiring + chronicheart + chronicrenal + chronicliver + chronicneurological + isdiabetes + immunosuppressiontreatment + immunosuppressiondisease + obesityclinical + hypertension + apache_mean + (1 | trustcode)

Fixed coefficients

	coef	exp(coef)	se(coef)	z	p
agecut(50,60]	0.426456767	1.5318203	0.1699031	2.51	1.2e-02
agecut(60,70]	0.977811765	2.6586322	0.1605738	6.09	1.1e-09
agecut(70,80]	1.622708934	5.0667974	0.1673715	9.70	0.0e+00
agecut(80,999]	1.976208090	7.2153312	0.2184487	9.05	0.0e+00
sexMale	0.207526316	1.2306301	0.1002270	2.07	3.8e-02
chonicrespiratoryYes	0.493508253	1.6380529	0.1436712	3.43	5.9e-04

asthmarequiringYes	0.003522413	1.0035286	0.1522956	0.02	9.8e-01
chronicheartYes	0.433054805	1.5419607	0.1421668	3.05	2.3e-03
chronicrenalYes	0.743962453	2.1042570	0.1816631	4.10	4.2e-05
chronicliverYes	-0.204175642	0.8153192	0.4334035	-0.47	6.4e-01
chronicneurologicalYes	0.059611816	1.0614244	0.2977174	0.20	8.4e-01
isdiabetesYes	0.295951175	1.3444045	0.1099083	2.69	7.1e-03
immunosuppressiontreatmentYes	-0.137625451	0.8714250	0.2710732	-0.51	6.1e-01
immunosuppressiondiseaseYes	0.992929805	2.6991308	0.2759074	3.60	3.2e-04
obesityclinicalYes	0.385746701	1.4707121	0.1403943	2.75	6.0e-03
hypertensionYes	0.064996412	1.0671552	0.1073305	0.61	5.4e-01
apache_mean	0.142348747	1.1529787	0.1142078	1.25	2.1e-01

Random effects

Group Variable Std Dev Variance

trustcode Intercept 0.8683330 0.7540023

>

> # fixed effects

> coeff_main = fixef(model)

> sd_main = sqrt(diag(vcov(model)))

> df_fixed = data.frame(effect = names(coeff_main),

+ coeff = coeff_main,

+ coeff_sd = sd_main,

+ hr = exp(coeff_main),

+ hr_lower = exp(coeff_main - 2 * sd_main),

+ hr_upper = exp(coeff_main + 2 * sd_main),

```

+         row.names=NULL)
> # write.csv(df_fixed, file = 'fixed_effect_apache.csv')
> # random effects
> coeff_rand = ranef(model)$trustcode
> sd_rand = sqrt(diag(model$variance)[1:length(coeff_rand)])

> df_random = data.frame(effect = names(coeff_rand),
+         coeff = coeff_rand,
+         coeff_sd = sd_rand,
+         hr = exp(coeff_rand),
+         hr_lower = exp(coeff_rand - 2 * sd_rand),
+         hr_upper = exp(coeff_rand + 2 * sd_rand),
+         row.names=NULL)
> # write.csv(df_random, file = 'df_random_apache.csv')

```

Test random effects

```

> fit1 <- coxph(Surv(death_time , death_ind) ~
+         agecut +
+         sex +
+         chonicrespiratory+
+         asthmarequiring+
+         chronicheart +
+         chronicrenal +
+         chronicliver +
+         chronicneurological +
+         isdiabetes +

```

```

+      immunosuppressiontreatment +
+      immunosuppressiondisease +
+      obesityclinical +
+      hypertension +
+      apache_mean , dat2 )
> print(anova(fit1,model))

```

Analysis of Deviance Table

Cox model: response is Surv(death_time, death_ind)

Model 1: ~agecut + sex + chonicrepiratory + asthmarequiring + chronicheart + chronicrenal + chronicliver + chronicneurological + isdiabetes + immunosuppressiontreatment + immunosuppressiondisease + obesityclinical + hypertension + apache_mean

Model 2: ~agecut + sex + chonicrepiratory + asthmarequiring + chronicheart + chronicrenal + chronicliver + chronicneurological + isdiabetes + immunosuppressiontreatment + immunosuppressiondisease + obesityclinical + hypertension + apache_mean + (1 | trustcode)

loglik Chisq Df P(>|Chi)

1 -4279.7

2 -4161.8 235.76 1 < 2.2e-16 ***

Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

> # p_value for random effect is 2.2e-16

Proportional odds assumption

> ## check ph assumption

> cox_model = coxph(Surv(death_time , death_ind) ~

+ agecut +

+ sex +

+ chonicrepiratory +

+ asthmarequiring +

+ chronicheart +


```

+      chronicrenal +
+      chronicliver +
+      chronicneurological +
+      isdiabetes +
+      immunosuppressiontreatment +
+      immunosuppressiondisease +
+      obesityclinical +
+      apache_mean, dat2)

```

```
> cox.zph(cox_model)
```

	chisq	df	p
agecut	1.11e+01	4	0.025
sex	2.01e-02	1	0.887
chronicrespiratory	2.76e-01	1	0.600
asthmarequiring	2.76e+00	1	0.097
chronicheart	2.58e-01	1	0.612
chronicrenal	4.16e-02	1	0.838
chronicliver	6.35e+00	1	0.012
chronicneurological	1.24e+00	1	0.266
isdiabetes	6.39e-01	1	0.424
immunosuppressiontreatment	3.17e-01	1	0.573
immunosuppressiondisease	9.06e-01	1	0.341
obesityclinical	9.49e-02	1	0.758
apache_mean	7.94e-04	1	0.978
GLOBAL	2.32e+01	16	0.109

Sensitivity analysis (restrict to patients with only 7 days followup)

```
> model = coxme(Surv(death_time , death_ind) ~  
+     agecut +  
+     sex +  
+     chonicrespiratory+  
+     asthmarequiring+  
+     chronicheart +  
+     chronicrenal +  
+     chronicliver +  
+     chronicneurological +  
+     isdiabetes +  
+     immunosuppressiontreatment +  
+     immunosuppressiondisease +  
+     obesityclinical +  
+     hypertension +  
+     apache_mean +  
+     (1|trustcode), dat2 %>% filter(death_time > 7))
```

```
> summary(model)
```

Cox mixed-effects model fit by maximum likelihood

Data: dat2 %>% filter(death_time > 7)

events, n = 303, 1983

Iterations= 17 125

NULL Integrated Fitted

Log-likelihood -2102.549 -1965.555 -1891.01

Chisq df p AIC BIC

Integrated loglik 273.99 18.00 0 237.99 171.14

Penalized loglik 423.08 70.18 0 282.72 22.09

Model: Surv(death_time, death_ind) ~ agecut + sex + chonicrespiratory + asthmarequiring + chronicheart + chronicrenal + chronicliver + chronicneurological + isdiabetes + immunosuppressiontreatment + immunosuppressiondisease + obesityclinical + hypertension + apache_mean + (1 | trustcode)

Fixed coefficients

	coef	exp(coef)	se(coef)	z	p
agecut[50,60]	0.26826437	1.3076928	0.2311499	1.16	2.5e-01
agecut[60,70]	0.89731735	2.4530137	0.2173021	4.13	3.6e-05
agecut[70,80]	1.19032366	3.2881453	0.2361042	5.04	4.6e-07
agecut[80,999]	1.76714363	5.8541080	0.3060082	5.77	7.7e-09
sexMale	0.32593990	1.3853321	0.1423201	2.29	2.2e-02
chonicrespiratoryYes	0.20071042	1.2222708	0.2306903	0.87	3.8e-01
asthmarequiringYes	0.37435721	1.4540565	0.2040478	1.83	6.7e-02
chronicheartYes	0.60362269	1.8287317	0.2058135	2.93	3.4e-03
chronicrenalYes	0.60084187	1.8236534	0.3002572	2.00	4.5e-02
chronicliverYes	0.17025366	1.1856056	0.4818020	0.35	7.2e-01
chronicneurologicalYes	0.22627541	1.2539210	0.3962750	0.57	5.7e-01
isdiabetesYes	0.28228305	1.3261540	0.1600019	1.76	7.8e-02
immunosuppressiontreatmentYes	-0.03971823	0.9610602	0.3780080	-0.11	9.2e-01
immunosuppressiondiseaseYes	0.72958801	2.0742259	0.4256785	1.71	8.7e-02
obesityclinicalYes	0.23764107	1.2682539	0.1993900	1.19	2.3e-01
hypertensionYes	0.17030782	1.1856698	0.1528342	1.11	2.7e-01
apache_mean	0.11008774	1.1163760	0.1283131	0.86	3.9e-01

Random effects

Group	Variable	Std Dev	Variance
trustcode	Intercept	0.8988636	0.8079557