```
clear all

cd "I:\DOR\CNR\STATA"

capture log close
log using "CNR_HAPU_NC_Linnen.smcl", replace

set more off

version 14

/*      KPNC Division of Research
        CNR KPNC Pressure Injury Study: "Synch 'n' Swim"
        PI: Dr. Gabriel Escobar (senior author)
        Co-I: Daniel Linnen (first author)
        January-September 2017

        RQ: Which hospital-level characteristics explain the residual
variation
        in HAPU incidence rates observed across KP California
hospitals
        after adjustment for patient-level risk factors?


*/

use "J:\SRI\PROJECTS\CNR\cnr_episodes_final_d.dta", clear
preserve
import delimited I:\DOR\CNR\fac_capacity.csv, clear
save "I:\DOR\CNR\fac_capacity.dta", replace
restore
merge m:1 start_fac_id using "I:\DOR\CNR\fac_capacity.dta"

describe, short

// Step 1: Generate hospital-level variables
// ----------------------------------------


// Remove episodes meeting exclusion criteria:
// Keep: Episodes with a one or two KFH stays
// Censor: KFH to nonKFH transport out episodes
// Exclude: Episodes with 3 or more stays or nonKFH to KFH transport
in episodes

generate include = 0
replace include = 1 if episode_inclusion == "Keep" | episode_inclusion
== "Censor"
drop if include == 0
```

```
generate female = 0
replace female = 1 if sex =="F"

generate admit_cat2 = .
replace admit_cat2 = 1 if admit_cat == "1 ED, SURGICAL"
replace admit_cat2 = 2 if admit_cat == "2 NON-ED, SURGICAL"
replace admit_cat2 = 3 if admit_cat == "3 ED, MEDICAL"
replace admit_cat2 = 4 if admit_cat == "4 NON-ED, MEDICAL"


//Combine Hayward and San Leandro into San Leandro
replace start_fac = "SLN" if start_fac == "HAY"
replace end_fac_vendor = "SLN" if end_fac_vendor == "HAY"

// All hospital-level variables have fac_ prefix
rename popdensity fac_popdens
label variable fac_popdens "Facility County Population Density"

// generate new binary popdens var (<750), there appear to main groups
of density
generate fac_popdens_750 = 0
replace fac_popdens_750 = 1 if fac_popdens >750

// Hospital mean daily admission volume
bysort start_fac_id: egen fac_volume = count (main_enc_id)

// Hospital occupancy: mean monthly volume as % of maximum monthly
admission volume per hospitals
// from merged dataset "fac_capacity"
rename pc_diff_from_max fac_occupancy

// Mean hospital patient age
bysort start_fac_id: egen fac_age= mean(age_at_admit)

// Mean hospital COPS2
bysort start_fac_id: egen fac_cops2 = mean(cops2_prior)

// %COPS2  >= 65
generate fac_cops2_65 = (cops2_prior >= 65)
by start_fac_id, sort: egen fac_cops2_65_pc = mean(100 * fac_cops2_65)

//double checking percentage is calculated correctly (percent of
hospital episodes)
/// percentage syntax  "mean (100*var)" produces aggregate percentage
of var=1/all vars
/// http://www.stata.com/support/faqs/data-management/creating-
percent-summary-variables/

by start_fac_id, sort: generate fac_freq = sum(fac_cops2_65) // sums
all 1's per hospital
```

```
by start_fac_id: generate fac_pc_cops65 = fac_freq / _N // proportion
of above over all hospital episodes
bysort start_fac_id: egen fac_cops_pc = max(fac_pc_cops65) //the max
value of proportion over all hospitals

// Mean hospital LAPS2 at hospital entry time (HET)
bysort start_fac_id: egen fac_laps2 = mean(laps2_het)

// % LAPS2 => 110
generate fac_laps2_110 = (laps2_het >= 110)
by start_fac_id, sort: egen fac_laps2_110_pc = mean(100 *
fac_laps2_110)

// Mean LOS
bysort start_fac_id: egen fac_los = mean(episode_los)

// % Hospital Full Code (1/0)
by start_fac_id, sort: egen fac_full_code_pc = mean(100 * full_code)

// HAPU Stage (POA=0)
encode hapu_stage, generate (facility_hapu_stage)
codebook facility_hapu_stage
bysort start_fac_id: egen facility_hapu_stage_totals =
count(facility_hapu_stage)
// (this counts all hapus per facility)
// create hapu stage variables:
generate facility_hapu_stage_dtiunst = facility_hapu_stage == 1 |
facility_hapu_stage ==6
generate facility_hapu_stage_1 = facility_hapu_stage == 2 if
pu_poa_flag ==0
generate facility_hapu_stage_2 = facility_hapu_stage == 3 if
pu_poa_flag ==0
generate facility_hapu_stage_3 = facility_hapu_stage ==4 if
pu_poa_flag ==0
generate facility_hapu_stage_4 = facility_hapu_stage == 5 if
pu_poa_flag ==0

// create facility totals for each stage
bysort start_fac_id: egen fac_hapu_stage_dtiunst_count =
total(facility_hapu_stage_dtiunst)
bysort start_fac_id: egen fac_hapu_stage_1_count =
total(facility_hapu_stage_1)
bysort start_fac_id: egen fac_hapu_stage_2_count =
total(facility_hapu_stage_2)
bysort start_fac_id: egen fac_hapu_stage_3_count =
total(facility_hapu_stage_3)
bysort start_fac_id: egen fac_hapu_stage_4_count =
total(facility_hapu_stage_4)

// Discharges to SNF (%)
```

```
gen facility_dc_snf = 1 ///
if dschrg_disp_ds_tx == "TO LONG TERM CARE"
replace facility_dc_snf = 0 ///
if dschrg_disp_ds_tx != "TO LONG TERM CARE"

bysort start_fac_id: egen fac_dc_snf_pc = mean(100 * facility_dc_snf)

// Hospital patients with Diabetes (%)
bysort start_fac_id: egen fac_dm_pc = mean(100 * hx_diabetes_flag)

// Hospital patients with Stroke (%)
bysort start_fac_id: egen fac_stroke_pc = mean(100 * hx_stroke_flag)

// Admission Type (categories)
encode admit_cat, generate(facility_admit_cat)
codebook facility_admit_cat

// create admit cat variables:
generate facility_admit_1 = facility_admit_cat == 1
generate facility_admit_2 = facility_admit_cat == 2
generate facility_admit_3 = facility_admit_cat == 3
generate facility_admit_4 = facility_admit_cat == 4

// create facility totals for each admit cat
bysort start_fac_id: egen fac_admit_1_count = total(facility_admit_1)
bysort start_fac_id: egen fac_admit_2_count = total(facility_admit_2)
bysort start_fac_id: egen fac_admit_3_count = total(facility_admit_3)
bysort start_fac_id: egen fac_admit_4_count = total(facility_admit_4)

// admit categories as % of total
bysort start_fac_id: egen fac_admit_1_pc = mean(100 *
facility_admit_1)
bysort start_fac_id: egen fac_admit_2_pc = mean(100 *
facility_admit_2)
bysort start_fac_id: egen fac_admit_3_pc = mean(100 *
facility_admit_3)
bysort start_fac_id: egen fac_admit_4_pc = mean(100 *
facility_admit_4)

// Mean hospital Braden
destring braden_total, replace
bysort start_fac_id: egen fac_braden = mean(braden_total)

// Mean hospital mortality %
bysort start_fac_id: egen fac_died_pc = mean(100 * death_in_cd)

// Mean facility Sepsis Primary admission DRG's (ICD-9)
// 995.91=Sepsis 995.92=Severe Sepsis     785.52=Septic Shock
tab admsn_prncpl_icd_diag_cd if admsn_prncpl_icd_diag_cd =="995.91" |
admsn_prncpl_icd_diag_cd =="995.92"|admsn_prncpl_icd_diag_cd ==
```

```
"785.52"

generate fac_sepsis = 0
replace fac_sepsis = 1 if admsn_prncpl_icd_diag_cd == "995.91"
replace fac_sepsis = 1 if admsn_prncpl_icd_diag_cd == "995.92"
replace fac_sepsis = 1 if admsn_prncpl_icd_diag_cd == "785.52"
tab fac_sepsis

bysort start_fac_id: egen fac_sepsis_pc = mean(100 * fac_sepsis)

// Hospital ICU stays
// (% of hospital total admitted to the ICU at any point in time
during the episode)
bysort start_fac_id: egen fac_icu_pc = mean(100 * icu_flag)

// generate outcome
generate hapu = 0
replace hapu = 1 if hapu_midas_flag == 1
tab hapu

// Hospital-level outcome variable (MIDAS HAPU)
// to use for correlation with above predictors
bysort start_fac_id: egen fac_hapu_pc = mean(hapu * 100)

// Tabulate inital set of hospital level variables by hospital
tabstat fac_full_code fac_died_pc fac_dc_snf_pc fac_age
fac_laps2_110_pc fac_cops2_65_pc ///
        fac_volume fac_occupancy fac_admit_2_count fac_admit_4_count
fac_admit_1_pc fac_admit_2_pc fac_admit_3_pc fac_admit_4_pc ///
        fac_icu_pc fac_dm fac_stroke fac_sepsis_pc fac_braden
fac_popdens fac_hapu_pc, by(start_fac_id)


// Step 2: Drop variables
// ----------------------
ds, skip(8)
drop prio* ed_* rgn_cd brth_dt linked_epi* stays_in_e* episode_st*
drop er_in_cd_ep obs_in_cd_ep inp_in_cd_ep kphc_admsn* ptnt_cls_d*
drop linked_enc* end_fac_*  icu_flag episode_in* braden_rec*
drop braden_sen* braden_moi* braden_act* braden_mob* braden_nut*
braden_fri*
drop hapu_locat* hapu_later* hapu_medic* hapu_brade* multiple_h*
drop hapu_histo*
ds, skip(8)


// Correlation matrix of all hospital-level variables
// --------------------------------------------------

//Standardize variables
```

```
egen z_fac_popdens = std(fac_popdens)
egen z_fac_occupancy = std(fac_occupancy)
egen z_fac_volume = std(fac_volume)
egen z_fac_age = std(fac_age)
egen z_fac_cops2 = std(fac_cops2)
egen z_fac_cops2_65_pc = std(fac_cops2_65_pc)
egen z_fac_laps2 = std(fac_laps2)
egen z_fac_laps2_110_pc = std(fac_laps2_110_pc)
egen z_fac_los = std(fac_los)
egen z_fac_full_code_pc = std(fac_full_code_pc)
egen z_fac_dc_snf_pc = std(fac_dc_snf_pc)
egen z_fac_dm_pc = std(fac_dm_pc)
egen z_fac_stroke_pc = std(fac_stroke_pc)
egen z_fac_admit_1_pc = std(fac_admit_1_pc)
egen z_fac_admit_2_pc = std(fac_admit_2_pc)
egen z_fac_admit_3_pc = std(fac_admit_3_pc)
egen z_fac_admit_4_pc = std(fac_admit_4_pc)
egen z_fac_braden = std(fac_braden)
egen z_fac_died_pc = std(fac_died_pc)
egen z_fac_sepsis = std(fac_sepsis)
egen z_fac_sepsis_pc = std(fac_sepsis_pc)
egen z_fac_icu_pc = std(fac_icu_pc)
egen z_fac_hapu_pc = std(fac_hapu_pc)

sum z_*

pwcorr fac_popdens fac_occupancy fac_volume fac_age fac_cops2
fac_cops2_65_pc ///
fac_laps2 fac_laps2_110_pc fac_los fac_full_code_pc fac_dc_snf_pc ///
fac_dm_pc fac_stroke_pc fac_admit_1_pc fac_admit_2_pc fac_admit_3_pc
fac_admit_4_pc ///
fac_braden fac_died_pc fac_sepsis_pc fac_icu_pc fac_hapu_pc ///
, sig star(.05)

pwcorr z_fac_popdens z_fac_occupancy z_fac_volume z_fac_age
z_fac_cops2 ///
z_fac_cops2_65_pc z_fac_laps2 z_fac_laps2_110_pc z_fac_los
z_fac_full_code_pc ///
z_fac_dc_snf_pc z_fac_dm_pc z_fac_stroke_pc  z_fac_admit_1_pc
z_fac_admit_2_pc ///
z_fac_admit_3_pc z_fac_admit_4_pc z_fac_braden z_fac_died_pc
z_fac_sepsis_pc ///
z_fac_icu_pc z_fac_hapu_pc ///
, sig star(.05)

mkcorr fac_popdens fac_occupancy fac_volume fac_age fac_cops2
fac_cops2_65_pc ///
fac_laps2 fac_laps2_110_pc fac_los fac_full_code_pc fac_dc_snf_pc ///
fac_dm_pc fac_stroke_pc fac_admit_1_pc fac_admit_2_pc fac_admit_3_pc
fac_admit_4_pc ///
```

```
fac_braden fac_died_pc fac_sepsis_pc fac_icu_pc fac_hapu_pc ///
, log(I:\DOR\CNR\correlation matrix) replace cdec(3)

mkcorr z_fac_popdens z_fac_occupancy z_fac_volume z_fac_age
z_fac_cops2 ///
z_fac_cops2_65_pc z_fac_laps2 z_fac_laps2_110_pc z_fac_los
z_fac_full_code_pc ///
z_fac_dc_snf_pc z_fac_dm_pc z_fac_stroke_pc  z_fac_admit_1_pc
z_fac_admit_2_pc ///
z_fac_admit_3_pc z_fac_admit_4_pc z_fac_braden z_fac_died_pc
z_fac_sepsis_pc ///
z_fac_icu_pc z_fac_hapu_pc ///
, log(I:\DOR\CNR\std correlation matrix) replace cdec(3)

// Principal component analysis
pca fac_popdens fac_occupancy fac_volume fac_age fac_cops2
fac_cops2_65_pc ///
fac_laps2 fac_laps2_110_pc fac_los fac_full_code_pc fac_dc_snf_pc ///
fac_dm_pc fac_stroke_pc fac_admit_1_pc fac_admit_2_pc fac_admit_3_pc
fac_admit_4_pc ///
fac_braden fac_died_pc fac_sepsis_pc fac_icu_pc fac_hapu_pc
screeplot, yline(1)
graph save Graph "I:\DOR\CNR\pca factors.gph", replace

pca z_fac_popdens z_fac_occupancy z_fac_volume z_fac_age
z_fac_cops2 ///
z_fac_cops2_65_pc z_fac_laps2 z_fac_laps2_110_pc z_fac_los
z_fac_full_code_pc ///
z_fac_dc_snf_pc z_fac_dm_pc z_fac_stroke_pc  z_fac_admit_1_pc
z_fac_admit_2_pc ///
z_fac_admit_3_pc z_fac_admit_4_pc z_fac_braden z_fac_died_pc
z_fac_sepsis_pc ///
z_fac_icu_pc z_fac_hapu_pc
screeplot, yline(1)
graph save Graph "I:\DOR\CNR\pca factors standardized.gph", replace

pca fac_popdens fac_occupancy fac_volume fac_age fac_cops2
fac_cops2_65_pc ///
fac_laps2 fac_laps2_110_pc fac_los fac_full_code_pc fac_dc_snf_pc ///
fac_dm_pc fac_stroke_pc fac_admit_1_pc fac_admit_2_pc fac_admit_3_pc
fac_admit_4_pc ///
fac_braden fac_died_pc fac_sepsis_pc fac_icu_pc fac_hapu_pc ///
,mineigen(1) blanks (.3)
rotate, varimax blanks(.3)
// Standardized variables yield identical eigenvector values
// Using standardized variables with z_ prefix for cluster analysis


// K-means cluster analysis of hospital-level covariates
```

```
cluster kmeans z_fac_popdens z_fac_occupancy z_fac_volume z_fac_age
z_fac_cops2 ///
z_fac_cops2_65_pc z_fac_laps2 z_fac_laps2_110_pc z_fac_los
z_fac_full_code_pc ///
z_fac_dc_snf_pc z_fac_dm_pc z_fac_stroke_pc  z_fac_admit_1_pc
z_fac_admit_2_pc ///
z_fac_admit_3_pc z_fac_admit_4_pc z_fac_braden z_fac_died_pc
z_fac_sepsis_pc ///
z_fac_icu_pc z_fac_hapu_pc ///
, k(4) name(clus1) measure(L2) start(krandom)

cluster list clus1
table clus1

tabstat z_fac_popdens z_fac_occupancy z_fac_volume z_fac_age
z_fac_cops2 ///
z_fac_cops2_65_pc z_fac_laps2 z_fac_laps2_110_pc z_fac_los
z_fac_full_code_pc ///
z_fac_dc_snf_pc z_fac_dm_pc z_fac_stroke_pc  z_fac_admit_1_pc
z_fac_admit_2_pc ///
z_fac_admit_3_pc z_fac_admit_4_pc z_fac_braden z_fac_died_pc
z_fac_sepsis_pc ///
z_fac_icu_pc z_fac_hapu_pc ///
, by(clus1) stat(mean)


graph bar (mean) z_fac_popdens z_fac_occupancy z_fac_volume z_fac_age
z_fac_cops2 ///
z_fac_cops2_65_pc z_fac_laps2 z_fac_laps2_110_pc z_fac_los
z_fac_full_code_pc ///
z_fac_dc_snf_pc z_fac_dm_pc z_fac_stroke_pc  z_fac_admit_1_pc
z_fac_admit_2_pc ///
z_fac_admit_3_pc z_fac_admit_4_pc z_fac_braden z_fac_died_pc
z_fac_sepsis_pc ///
z_fac_icu_pc z_fac_hapu_pc, over(clus1)


//compare smaller/larger cluster groups
cluster kmeans z_fac_popdens z_fac_occupancy z_fac_volume z_fac_age
z_fac_cops2 ///
z_fac_cops2_65_pc z_fac_laps2 z_fac_laps2_110_pc z_fac_los
z_fac_full_code_pc ///
z_fac_dc_snf_pc z_fac_dm_pc z_fac_stroke_pc  z_fac_admit_1_pc
z_fac_admit_2_pc ///
z_fac_admit_3_pc z_fac_admit_4_pc z_fac_braden z_fac_died_pc
z_fac_sepsis_pc ///
z_fac_icu_pc z_fac_hapu_pc ///
, k(2) name(clus2) start(firstk) measure(abs)

cluster kmeans z_fac_popdens z_fac_occupancy z_fac_volume z_fac_age
```

```
z_fac_cops2 ///
z_fac_cops2_65_pc z_fac_laps2 z_fac_laps2_110_pc z_fac_los
z_fac_full_code_pc ///
z_fac_dc_snf_pc z_fac_dm_pc z_fac_stroke_pc  z_fac_admit_1_pc
z_fac_admit_2_pc ///
z_fac_admit_3_pc z_fac_admit_4_pc z_fac_braden z_fac_died_pc
z_fac_sepsis_pc ///
z_fac_icu_pc z_fac_hapu_pc ///
, k(6) name(clus3) start(random(33576))

table clus1 clus2, col
table clus1 clus3, col
table clus2 clus3, col

// at k(4) groups we appear to get similar results as in correlation
matrix and PCA

graph matrix z_fac_popdens z_fac_occupancy z_fac_volume z_fac_age
z_fac_cops2 ///
z_fac_cops2_65_pc z_fac_laps2 z_fac_laps2_110_pc z_fac_los
z_fac_full_code_pc ///
z_fac_dc_snf_pc z_fac_dm_pc z_fac_stroke_pc  z_fac_admit_1_pc
z_fac_admit_2_pc ///
z_fac_admit_3_pc z_fac_admit_4_pc z_fac_braden z_fac_died_pc
z_fac_sepsis_pc ///
z_fac_icu_pc z_fac_hapu_pc ///
, m(i) mlabel(clus1) mlabpos(0)

// After observing collinearity, pca and k-means clusters we
determined the followig hospital-level candiate variables:
tabulate fac_occupancy fac_dm_pc  fac_icu_pc fac_braden
fac_popdens ///
fac_stroke_pc fac_cops2_65_pc fac_laps2_110_pc ///
fac_age fac_dc_snf_pc fac_died_pc

// Test variables for normality, linearity, outliers/influential
points, and check distribution across hospitals
histogram fac_occupancy, bin(10) percent
graph box fac_occupancy
qnorm fac_occupancy
pnorm fac_occupancy
twoway (bar fac_occupancy start_fac_id, sort)

histogram fac_dm_pc bin(10) percent
graph box fac_dm_pc
qnorm fac_dm_pc
pnorm fac_dm_pc
twoway (bar fac_dm_pc start_fac_id, sort)

histogram fac_age bin(10) percent
```

```
graph box fac_age
qnorm fac_age
pnorm fac_age
twoway (bar fac_age start_fac_id, sort)

histogram fac_dc_snf_pc bin(10) percent
graph box fac_dc_snf_pc
qnorm fac_dc_snf_pc
pnorm fac_dc_snf_pc
twoway (bar fac_dc_snf_pc start_fac_id, sort)

histogram fac_died_pc bin(10) percent
graph box fac_died_pc
qnorm fac_died_pc
pnorm fac_died_pc
twoway (bar fac_died start_fac_id, sort)

histogram fac_icu_pc bin(10) percent
graph box fac_icu_pc
qnorm fac_icu_pc
pnorm fac_icu_pc
twoway (bar fac_icu_pc start_fac_id, sort)

histogram fac_braden bin(10) percent
graph box fac_braden
qnorm fac_braden
pnorm fac_braden
twoway (bar fac_braden start_fac_id, sort)

histogram fac_popdens bin(10) percent
graph box fac_popdens
qnorm fac_popdens
pnorm fac_popdens
twoway (bar fac_popdens start_fac_id, sort)
twoway (bar fac_popdens start_fac_id, sort) if fac_popdens < 5000

histogram fac_sepsis_pc bin(10) percent
graph box fac_sepsis_pc
qnorm fac_sepsis_pc
pnorm fac_sepsis_pc
twoway (bar fac_sepsis_pc start_fac_id, sort)
// drop sepsis from list, quite evenly distributed

histogram fac_stroke_pc bin(10) percent
graph box fac_stroke_pc
qnorm fac_stroke_pc
pnorm fac_stroke_pc
twoway (bar fac_stroke_pc start_fac_id, sort)

histogram fac_cops2_65_pc bin(10) percent
```

```
graph box fac_cops2_65_pc
qnorm fac_cops2_65_pc
pnorm fac_cops2_65_pc
twoway (bar fac_cops2_65_pc start_fac_id, sort)

histogram fac_laps2_110_pc bin(10) percent
graph box fac_laps2_110_pc
qnorm fac_laps2_110_pc
pnorm fac_laps2_110_pc
twoway (bar fac_laps2_110_pc start_fac_id, sort)

// Test linearity of univariate outcome/hospital var
mkspline2 occupancy_spline = fac_occupancy, cubic
regress hapu occupancy_spline1 occupancy_spline2 occupancy_spline3
occupancy_spline4
adjustrcspline, custominvlink("1/(1+exp(-1*xb()))")

mkspline2 age_spline = fac_age, cubic
regress hapu age_spline1 age_spline2 age_spline3 age_spline4
adjustrcspline, custominvlink("1/(1+exp(-1*xb()))")

mkspline2 dm_spline = fac_dm_pc, cubic
regress hapu dm_spline1 dm_spline2 dm_spline3 dm_spline4
adjustrcspline, custominvlink("1/(1+exp(-1*xb()))")

mkspline2 snf_spline = fac_dc_snf_pc, cubic
regress hapu snf_spline1 snf_spline2 snf_spline3 snf_spline4
adjustrcspline, custominvlink("1/(1+exp(-1*xb()))")

mkspline2 died_spline = fac_died_pc, cubic
regress hapu died_spline1 died_spline2 died_spline3 died_spline4
adjustrcspline, custominvlink("1/(1+exp(-1*xb()))")

mkspline2 icu_spline = fac_icu_pc, cubic
regress hapu icu_spline1 icu_spline2 icu_spline3 icu_spline4
adjustrcspline, custominvlink("1/(1+exp(-1*xb()))")

mkspline2 braden_spline = fac_braden, cubic
regress hapu braden_spline1 braden_spline2 braden_spline3
braden_spline4
adjustrcspline, custominvlink("1/(1+exp(-1*xb()))")

mkspline2 stroke_spline = fac_stroke_pc, cubic
regress hapu stroke_spline1 stroke_spline2 stroke_spline3
stroke_spline4
adjustrcspline, custominvlink("1/(1+exp(-1*xb()))")

mkspline2 cops_spline = fac_cops2_65_pc, cubic
regress hapu cops_spline1 cops_spline2 cops_spline3 cops_spline4
adjustrcspline, custominvlink("1/(1+exp(-1*xb()))")
```

```
mkspline2 laps_spline = fac_laps2_110_pc, cubic
regress hapu laps_spline1 laps_spline2 laps_spline3 laps_spline4
adjustrcspline, custominvlink("1/(1+exp(-1*xb()))")

//make 3-spline transformations for fac_snf and fac_cops2_65
mkspline2 cops_spline2 = fac_cops2_65_pc, cubic nknots(4)
mkspline2 snf_spline2 = fac_dc_snf_pc, cubic nknots(4)

// set up time-to-event variable

//Check for astronomical outliers for LOS
graph box episode_los, by(hapu)
// We noted 27 episodes >200days, plan to do sensitivity analysis of
pruned vs complete

// het is hospital entry time (Tzero)
// hapu_incident_date is event time
// Observations are censored by discharge (episode_end_ts)

// generate days-to-event and days-to-censor
gen days_to_event = hapu_incident_date - het
gen days_to_censor = episode_end_ts - het

// Missing ~5,900 het observations
// Preferably, use LOS (in hours) /24 since it has fewer missing:

gen los_days = episode_los/24
replace los_days = ceil(los_days)
// about 1,700 missing observations
replace los_days = days_to_censor if los_days ==.


summarize los_days days_to_event days_to_censor
codebook los_days days_to_event days_to_censor

//combine event days and censor days into one variable for CPH
generate time = .
replace time = days_to_event if hapu==1
replace time = los_days if hapu==0

generate pruned_time =time if los_days <= 200 // this did not end up
making a difference

save "J:\SRI\PROJECTS\CNR\cnr_full_dataset.dta",replace

// Step 3: Randomly split the dataset into derivation and validaton
and save the sets
// -------------------------------------------------------------
clear
```

```
use "J:\SRI\PROJECTS\CNR\cnr_full_dataset.dta"

describe
display 728266 * .7
//randomly split 70% (509786 observations) of dataset for k-fold
cross-validation, 30% for validation

set seed 03250506
gen random_number = uniform()
sort random_number

generate grp = .
replace grp = 0 in 1/509786
replace grp = 1 in 509787/728266

preserve
keep if grp==0
save "J:\SRI\PROJECTS\CNR\cnr_derivation_file.dta", replace

restore
keep if grp==1
save "J:\SRI\PROJECTS\CNR\cnr_validation_file.dta", replace

use "J:\SRI\PROJECTS\CNR\cnr_derivation_file.dta"



// ********** Cox Proportional Hazards Model
*******************************

// Patient-level predictors from Steve Zuniga's model:
// Age, sex, admit cat, code status, dm flag, stroke flag,
// cops2_prior, laps2_het, braden_total
stset time, failure(hapu)
sts graph, failure risktable
graph save Graph "I:\DOR\CNR\steve's KM.gph", replace
stcox age_at_admit i.female i.admit_cat2 i.full_code
i.hx_diabetes_flag i.hx_stroke_flag cops2_prior laps2_het
braden_total, vce(robust)

//Sensitivity analysis: compare with pruned time (outlier LOS removed)
stset pruned_time, failure(hapu)
sts graph, failure risktable
graph save Graph "I:\DOR\CNR\steve's KM.gph", replace
stcox age_at_admit i.female i.admit_cat2 i.full_code
i.hx_diabetes_flag i.hx_stroke_flag cops2_prior laps2_het
braden_total, vce(robust)

// Sensitivity analysis shows very little difference. Retaining full
dataset.
```

```
// Confirm that observations are equally distributed across derivation
set
// compared to full dataset

histogram hapu, discrete percent addlabel by(start_fac_id)

save "J:\SRI\PROJECTS\CNR\cnr_validation_file.dta", replace

use "J:\SRI\PROJECTS\CNR\cnr_episodes_final_d.dta", clear
histogram hapu, discrete percent addlabel by(start_fac_id)

use "J:\SRI\PROJECTS\CNR\cnr_derivation_file.dta"


// Steve's model: With start_fac_id as level 2 cluster variable
stset time, failure(hapu)
mestreg age_at_admit i.female ib2.admit_cat2 i.full_code
i.hx_diabetes_flag ///
i.hx_stroke_flag cops2_prior laps2_het braden_total || start_fac_id:,
distribution(weibull) vce(robust)

// without robust SE
mestreg age_at_admit i.female ib2.admit_cat2 i.full_code
i.hx_diabetes_flag ///
i.hx_stroke_flag cops2_prior laps2_het braden_total || start_fac_id:,
distribution(weibull)


// Model building: Multilevel hierarchical parametric survival
analysis
//stepwise backwards elimination with alpha = .2
// ------------------------------------------------------------

// Full model
mestreg age_at_admit i.female ib2.admit_cat2 i.full_code
i.hx_diabetes_flag ///
i.hx_stroke_flag cops2_prior laps2_het braden_total fac_occupancy
fac_dm_pc ///
fac_icu_pc fac_braden i.fac_popdens_750 fac_stroke_pc ///
fac_cops2_65_pc fac_laps2_110_pc fac_age fac_dc_snf_pc fac_died_pc ||
start_fac_id:, ///
distribution(weibull) vce(robust)

// See if cubic splines show better fit for
// fac_dc_snf_pc, fac_died_pc, fac_braden, fac_cops2_65
mestreg age_at_admit i.female ib2.admit_cat2 i.full_code
i.hx_diabetes_flag ///
i.hx_stroke_flag cops2_prior laps2_het braden_total fac_occupancy
fac_dm_pc ///
```

```
fac_icu_pc fac_braden i.fac_popdens_750 fac_stroke_pc ///
fac_cops2_65_pc fac_laps2_110_pc fac_age snf_spline* fac_died_pc ||
start_fac_id:, ///
distribution(weibull) vce(robust)
// better model fit with snf cubic splines (retain)

mestreg age_at_admit i.female ib2.admit_cat2 i.full_code
i.hx_diabetes_flag ///
i.hx_stroke_flag cops2_prior laps2_het braden_total fac_occupancy
fac_dm_pc ///
fac_icu_pc fac_braden i.fac_popdens_750 fac_stroke_pc ///
fac_cops2_65_pc fac_laps2_110_pc fac_age fac_dc_snf_pc died_spline* ||
start_fac_id:, ///
distribution(weibull) vce(robust)
// less of a difference with died spline

mestreg age_at_admit i.female ib2.admit_cat2 i.full_code
i.hx_diabetes_flag ///
i.hx_stroke_flag cops2_prior laps2_het braden_total fac_occupancy
fac_dm_pc ///
fac_icu_pc braden_spline* i.fac_popdens_750 fac_stroke_pc ///
fac_cops2_65_pc fac_laps2_110_pc fac_age fac_dc_snf_pc fac_died_pc ||
start_fac_id:, ///
distribution(weibull) vce(robust)
//less of a difference with braden spline

mestreg age_at_admit i.female ib2.admit_cat2 i.full_code
i.hx_diabetes_flag ///
i.hx_stroke_flag cops2_prior laps2_het braden_total fac_occupancy
fac_dm_pc ///
fac_icu_pc fac_braden i.fac_popdens_750 fac_stroke_pc ///
cops_spline* fac_laps2_110_pc fac_age fac_dc_snf_pc fac_died_pc ||
start_fac_id:, ///
distribution(weibull) vce(robust)
//better model fit with cops2 cubic spline (retain)

mestreg age_at_admit i.female ib2.admit_cat2 i.full_code
i.hx_diabetes_flag ///
i.hx_stroke_flag cops2_prior laps2_het braden_total fac_occupancy
fac_dm_pc ///
fac_icu_pc fac_braden i.fac_popdens_750 fac_stroke_pc ///
cops_spline* fac_laps2_110_pc fac_age snf_spline* fac_died_pc ||
start_fac_id:, ///
distribution(weibull) vce(robust)
//

//manual stepwise backward elimination (alpha = .2)
//STATA does not support stepwise command with mestreg command
stset time, failure(hapu)
mestreg age_at_admit i.female ib2.admit_cat2 i.full_code
```

```
i.hx_diabetes_flag ///
i.hx_stroke_flag cops2_prior laps2_het braden_total fac_occupancy
fac_dm_pc ///
fac_icu_pc fac_braden i.fac_popdens_750 fac_stroke_pc ///
cops_spline*_pc fac_laps2_110_pc fac_age snf_spline* fac_died_pc ||
start_fac_id:, ///
distribution(weibull) vce(robust)
// drop fac_laps2_110_pc

mestreg age_at_admit i.female ib2.admit_cat2 i.full_code
i.hx_diabetes_flag ///
i.hx_stroke_flag cops2_prior laps2_het braden_total fac_occupancy
fac_dm_pc ///
fac_icu_pc fac_braden i.fac_popdens_750 fac_stroke_pc ///
cops_spline* fac_age snf_spline* fac_died_pc || start_fac_id:, ///
distribution(weibull) vce(robust)
// drop fac_icu

mestreg age_at_admit i.female ib2.admit_cat2 i.full_code
i.hx_diabetes_flag ///
i.hx_stroke_flag cops2_prior laps2_het braden_total fac_occupancy
fac_dm_pc ///
fac_braden i.fac_popdens_750 fac_stroke_pc ///
cops_spline* fac_age snf_spline* fac_died_pc || start_fac_id:, ///
distribution(weibull) vce(robust)
// drop fac_occupancy

mestreg age_at_admit i.female ib2.admit_cat2 i.full_code
i.hx_diabetes_flag ///
i.hx_stroke_flag cops2_prior laps2_het braden_total fac_dm_pc ///
fac_braden i.fac_popdens_750 fac_stroke_pc ///
cops_spline* fac_age snf_spline* fac_died_pc || start_fac_id:, ///
distribution(weibull) vce(robust)
//drop fac_popdens_750

mestreg age_at_admit i.female ib2.admit_cat2 i.full_code
i.hx_diabetes_flag ///
i.hx_stroke_flag cops2_prior laps2_het braden_total fac_dm_pc ///
fac_braden fac_stroke_pc ///
cops_spline* fac_age snf_spline* fac_died_pc || start_fac_id:, ///
distribution(weibull) vce(robust)
//drop fac_stroke

//Final spline model before CV:
mestreg age_at_admit i.female ib2.admit_cat2 i.full_code
i.hx_diabetes_flag ///
i.hx_stroke_flag cops2_prior laps2_het braden_total fac_dm_pc ///
fac_braden cops_spline* fac_age snf_spline* fac_died_pc ||
start_fac_id:, ///
distribution(weibull) vce(robust)
```

```
sts graph, failure risktable ylabel(0(.05).25)
graph save Graph "I:\DOR\CNR\CNR KM.tif", replace

sts graph, cumhaz
graph save Graph "I:\DOR\CNR\CNR cumulative hazard.tif", replace

//Stepwise model building without splines (alpha for inclusion = .25)
stset time, failure(hapu)
mestreg age_at_admit i.female ib2.admit_cat2 i.full_code
i.hx_diabetes_flag ///
i.hx_stroke_flag cops2_prior laps2_het braden_total fac_occupancy
fac_dm_pc ///
fac_icu_pc fac_braden i.fac_popdens_750 fac_stroke_pc ///
fac_cops2_65_pc fac_laps2_110_pc fac_age fac_dc_snf_pc fac_died_pc ||
start_fac_id:, ///
distribution(weibull) vce(robust)

// drop fac_icu_pc
mestreg age_at_admit i.female ib2.admit_cat2 i.full_code
i.hx_diabetes_flag ///
i.hx_stroke_flag cops2_prior laps2_het braden_total fac_occupancy
fac_dm_pc ///
fac_braden i.fac_popdens_750 fac_stroke_pc ///
fac_cops2_65_pc fac_laps2_110_pc fac_age fac_dc_snf_pc fac_died_pc ||
start_fac_id:, ///
distribution(weibull) vce(robust)

// drop fac_popdens_750
mestreg age_at_admit i.female ib2.admit_cat2 i.full_code
i.hx_diabetes_flag ///
i.hx_stroke_flag cops2_prior laps2_het braden_total ///
fac_occupancy fac_dm_pc fac_braden  fac_stroke_pc ///
fac_cops2_65_pc fac_laps2_110_pc fac_age fac_dc_snf_pc fac_died_pc ||
start_fac_id:, ///
distribution(weibull) vce(robust)

// drop fac_cops2_65_pc
mestreg age_at_admit i.female ib2.admit_cat2 i.full_code
i.hx_diabetes_flag ///
i.hx_stroke_flag cops2_prior laps2_het braden_total ///
fac_occupancy fac_dm_pc fac_braden  fac_stroke_pc ///
fac_laps2_110_pc fac_age fac_dc_snf_pc fac_died_pc ||
start_fac_id:, ///
distribution(weibull) vce(robust)

// drop fac_age
mestreg age_at_admit i.female ib2.admit_cat2 i.full_code
```

```
i.hx_diabetes_flag ///
i.hx_stroke_flag cops2_prior laps2_het braden_total ///
fac_occupancy fac_dm_pc fac_braden  fac_stroke_pc ///
fac_laps2_110_pc fac_dc_snf_pc fac_died_pc || start_fac_id:, ///
distribution(weibull) vce(robust)

// drop fac_laps2_110_pc
mestreg age_at_admit i.female ib2.admit_cat2 i.full_code
i.hx_diabetes_flag ///
i.hx_stroke_flag cops2_prior laps2_het braden_total ///
fac_occupancy fac_dm_pc fac_braden  fac_stroke_pc ///
fac_dc_snf_pc fac_died_pc || start_fac_id:, ///
distribution(weibull) vce(robust)

// drop fac_dc_snf_pc
mestreg age_at_admit i.female ib2.admit_cat2 i.full_code
i.hx_diabetes_flag ///
i.hx_stroke_flag cops2_prior laps2_het braden_total ///
fac_occupancy fac_dm_pc fac_braden  fac_stroke_pc fac_died_pc ||
start_fac_id:, ///
distribution(weibull) vce(robust)

// drop fac_braden
mestreg age_at_admit i.female ib2.admit_cat2 i.full_code
i.hx_diabetes_flag ///
i.hx_stroke_flag cops2_prior laps2_het braden_total ///
fac_occupancy fac_dm_pc fac_stroke_pc fac_died_pc ||
start_fac_id:, ///
distribution(weibull) vce(robust)

// drop fac_occupancy
mestreg age_at_admit i.female ib2.admit_cat2 i.full_code
i.hx_diabetes_flag ///
i.hx_stroke_flag cops2_prior laps2_het braden_total ///
fac_dm_pc fac_stroke_pc fac_died_pc || start_fac_id:, ///
distribution(weibull) vce(robust)

//drop fac_stroke
mestreg age_at_admit i.female ib2.admit_cat2 i.full_code
i.hx_diabetes_flag ///
i.hx_stroke_flag cops2_prior laps2_het braden_total ///
fac_dm_pc fac_died_pc || start_fac_id:, distribution(weibull)
vce(robust)

//sensitivity analysis for model building with sepsis and popdens
inlcuded as binary variable
//Stepwise model building without splines and sepsis_grp(alpha for
inclusion = .25)
//sepsis_grp above/below median % .02
//popdens_grp above/below median 1266
```

```
stset time, failure(hapu)
mestreg age_at_admit i.female ib2.admit_cat2 i.full_code
i.hx_diabetes_flag ///
i.hx_stroke_flag cops2_prior laps2_het braden_total ///
fac_occupancy fac_dm_pc fac_icu_pc fac_braden i.popdens_grp
i.sepsis_grp fac_stroke_pc ///
fac_cops2_65_pc fac_laps2_110_pc fac_age fac_dc_snf_pc fac_died_pc ||
start_fac_id:, ///
distribution(weibull) vce(robust)

//drop fac_cops2_65_pc
mestreg age_at_admit i.female ib2.admit_cat2 i.full_code
i.hx_diabetes_flag ///
i.hx_stroke_flag cops2_prior laps2_het braden_total ///
fac_occupancy fac_dm_pc fac_icu_pc fac_braden i.popdens_grp
i.sepsis_grp fac_stroke_pc ///
fac_laps2_110_pc fac_age fac_dc_snf_pc fac_died_pc ||
start_fac_id:, ///
distribution(weibull) vce(robust)

//drop fac_age
mestreg age_at_admit i.female ib2.admit_cat2 i.full_code
i.hx_diabetes_flag ///
i.hx_stroke_flag cops2_prior laps2_het braden_total ///
fac_occupancy fac_dm_pc fac_icu_pc fac_braden i.popdens_grp
i.sepsis_grp fac_stroke_pc ///
fac_laps2_110_pc fac_dc_snf_pc fac_died_pc || start_fac_id:, ///
distribution(weibull) vce(robust)

//drop popdens_grp
mestreg age_at_admit i.female ib2.admit_cat2 i.full_code
i.hx_diabetes_flag ///
i.hx_stroke_flag cops2_prior laps2_het braden_total ///
fac_occupancy fac_dm_pc fac_icu_pc fac_braden i.sepsis_grp
fac_stroke_pc ///
fac_laps2_110_pc fac_dc_snf_pc fac_died_pc || start_fac_id:, ///
distribution(weibull) vce(robust)

//drop fac_icu_pc
mestreg age_at_admit i.female ib2.admit_cat2 i.full_code
i.hx_diabetes_flag ///
i.hx_stroke_flag cops2_prior laps2_het braden_total ///
fac_occupancy fac_dm_pc fac_braden i.sepsis_grp fac_stroke_pc ///
fac_laps2_110_pc fac_dc_snf_pc fac_died_pc || start_fac_id:, ///
distribution(weibull) vce(robust)

//drop fac_laps2_110_pc
mestreg age_at_admit i.female ib2.admit_cat2 i.full_code
i.hx_diabetes_flag ///
```

```
i.hx_stroke_flag cops2_prior laps2_het braden_total ///
fac_occupancy fac_dm_pc fac_braden i.sepsis_grp fac_stroke_pc ///
fac_dc_snf_pc fac_died_pc || start_fac_id:, ///
distribution(weibull) vce(robust)

//drop sepsis_grp
mestreg age_at_admit i.female ib2.admit_cat2 i.full_code
i.hx_diabetes_flag ///
i.hx_stroke_flag cops2_prior laps2_het braden_total ///
fac_occupancy fac_dm_pc fac_braden fac_stroke_pc ///
fac_dc_snf_pc fac_died_pc || start_fac_id:, ///
distribution(weibull) vce(robust)

//drop fac_dc_snf_pc
mestreg age_at_admit i.female ib2.admit_cat2 i.full_code
i.hx_diabetes_flag ///
i.hx_stroke_flag cops2_prior laps2_het braden_total ///
fac_occupancy fac_dm_pc fac_braden fac_stroke_pc fac_died_pc ||
start_fac_id:, distribution(weibull) vce(robust)

//steps omitted, model identical to above stepwise (whhich did not
have sepsis_grp variable)



// Model testing (with splines)
// ------------------------------------------------------------

// usig stmixed for comparison: "factor variables and time-series
operators not allowed"
// can't do lrtest m1 m2, ... (log likelihood test) with models using
robust SE

// Wald test checks out: Do the predictors make indiviudal
contributions to the model?
test fac_dm_pc fac_popdens_750 fac_died_pc

// AIC/BIC
// Model 1 Patient level (Steve)
stset time, failure(hapu)
mestreg age_at_admit i.female ib2.admit_cat2 i.full_code
i.hx_diabetes_flag ///
i.hx_stroke_flag cops2_prior laps2_het braden_total || start_fac_id:,
distribution(weibull) vce(robust)
estimates store m1
estat ic // AIC and BIC
// estat gof: Hosmer Lemeshow Goodness of fit "not valid"


// Model 2: Full model
```

```
mestreg age_at_admit i.female ib2.admit_cat2 i.full_code
i.hx_diabetes_flag ///
i.hx_stroke_flag cops2_prior laps2_het braden_total fac_dm_pc ///
fac_braden cops_spline* fac_age snf_spline* fac_died_pc ||
start_fac_id:, ///
distribution(weibull) vce(robust)
estimates store m2
estat ic


// Model 3: Full model no splines
mestreg age_at_admit i.female ib2.admit_cat2 i.full_code
i.hx_diabetes_flag///
i.hx_stroke_flag cops2_prior laps2_het braden_total fac_dm_pc ///
fac_died_pc || start_fac_id:, ///
distribution(weibull) vce(robust)


// test proportional hazards by sex and region
stphplot, by(sex)
stphplot, by(region)

// Cross-validation/bootstrapping to assess final spline model fit
//---------------------------------------------------------------

// Cross-validation: checking model stability across k-folds
// not available for mestreg in STATA, instead, using logistic
regression
// (Patricia did postestimation in SAS for the survival model)
// (5 partitions, 4:1 validation across partitions)

crossfold logistic hapu age_at_admit i.female ib2.admit_cat2
i.full_code i.hx_diabetes_flag ///
i.hx_stroke_flag cops2_prior laps2_het braden_total fac_dm_pc ///
fac_braden cops_spline* fac_age snf_spline* fac_died_pc, stub(kfold)
k(5)loud

// without splines
crossfold logistic hapu age_at_admit i.female ib2.admit_cat2
i.full_code i.hx_diabetes_flag ///
i.hx_stroke_flag cops2_prior laps2_het braden_total fac_dm_pc ///
fac_braden fac_age fac_died_pc, stub(kfold_nospline) k(5)loud


// With splines and R2 as the gof metric
crossfold logistic hapu age_at_admit i.female ib2.admit_cat2
i.full_code i.hx_diabetes_flag ///
i.hx_stroke_flag cops2_prior laps2_het braden_total fac_dm_pc ///
fac_braden cops_spline* fac_age snf_spline* fac_died_pc, r2
stub(kfold) k(5)
```

```
logistic hapu age_at_admit i.female ib2.admit_cat2 i.full_code
i.hx_diabetes_flag ///
i.hx_stroke_flag cops2_prior laps2_het braden_total fac_dm_pc ///
fac_braden cops_spline* fac_age snf_spline* fac_died_pc

// CV calibration test prodecure adapted from Vittinghoff, Glidden,
Shiboski, McCulloch
save "J:\SRI\PROJECTS\CNR\cv.dta", replace
clear all
use "J:\SRI\PROJECTS\CNR\cv.dta"

quietly logistic hapu age_at_admit i.female ib2.admit_cat2 i.full_code
i.hx_diabetes_flag ///
i.hx_stroke_flag cops2_prior laps2_het braden_total ///
fac_dm_pc fac_braden fac_age fac_died_pc ///
cops_spline21 cops_spline22 cops_spline23 snf_spline21 snf_spline22
snf_spline23

predict cvpredict, pr
predict cvresid, residuals
// Naive estimate of calibration
hl hapu cvpredict, plot

// Step 1: Divide data into 5 mututally exclusive subsets

xtile cvgroup = uniform(), nq(5)
gen cv_fitted = .

forvalues i = 1/5 {
        //Step 2: estimate  model omitting each subset
        qui logistic hapu age_at_admit i.female ib2.admit_cat2
i.full_code i.hx_diabetes_flag ///
        i.hx_stroke_flag cops2_prior laps2_het braden_total ///
        fac_dm_pc fac_braden fac_age fac_died_pc ///
        cops_spline21 cops_spline22 cops_spline23 snf_spline21
snf_spline22 snf_spline23 if cvgroup~=`i'
        qui predict cv_fittedi, pr
        hl hapu cv_fittedi, plot

        //Step 3: save cross-validated statistic for each omitted
subset
        qui replace cv_fitted = cv_fittedi if cvgroup==`i'
        qui drop cv_fittedi

        }

        // Assess Goodness of fit (Hosmer-Lemeshow)in derivation model
with three splines for snf and cops2
logistic hapu age_at_admit i.female ib2.admit_cat2 i.full_code
```

```
i.hx_diabetes_flag ///
i.hx_stroke_flag cops2_prior laps2_het braden_total ///
fac_dm_pc fac_braden fac_age fac_died_pc ///
cops_spline21 cops_spline22 cops_spline23 ///
snf_spline21 snf_spline22 snf_spline23

predict p, pr
hl hapu p, plot

predict r, residuals
predict h, hat //leverage


//Residuals vs fitted values
twoway (scatter r p, xline(0)) , plotregion(style(none))
xtitle("Predicted") ytitle("residual")

// Outlier check
twoway (scatter h p, xline(0)) , plotregion(style(none))
xtitle("Predicted") ytitle("leverage")

estat gof, table group(5)
//The Hosmer-Lemeshow test rejects the null hypothesis that estimated
and observed
// probabilities agree (P=0.008)

lroc
// c-stat = .835

// compare effects by facility
set more off
bysort start_fac: tabstat fac_dm_pc fac_braden fac_age fac_died_pc ///
cops_spline21 cops_spline22 cops_spline23 ///
snf_spline21 snf_spline22 snf_spline23


// No spline model: Assess gof in derivation set
logistic hapu age_at_admit i.female ib2.admit_cat2 i.full_code
i.hx_diabetes_flag ///
i.hx_stroke_flag cops2_prior laps2_het braden_total ///
fac_dm_pc fac_died_pc

predict p2, pr
hl hapu p2, plot

predict r2, residuals

lroc

estat gof, table group(10)
```

```
////////////////////////////////////////////////////////////////
// Final model validation using logistic regression (Patricia did
postestimation for survival model)
// Assess Goodness of fit in validation dataset

clear all
use "J:\SRI\PROJECTS\CNR\cnr_validation_file.dta"

logistic hapu age_at_admit i.female ib2.admit_cat2 i.full_code
i.hx_diabetes_flag ///
i.hx_stroke_flag cops2_prior laps2_het braden_total ///
fac_dm_pc fac_died_pc

lroc
ovfplot
predict p3, pr
hl hapu p3, plot
estat gof, table group(10)
estimates stats
// sensitivty analysis for inlcuding los in logistic regression model
for postestimation
generate los == 0
replace los = episode_los if episode_los != .

logistic hapu age_at_admit i.female ib2.admit_cat2 i.full_code
i.hx_diabetes_flag ///
i.hx_stroke_flag cops2_prior laps2_het braden_total ///
fac_dm_pc fac_died_pc

logistic hapu age_at_admit i.female ib2.admit_cat2 i.full_code
i.hx_diabetes_flag ///
i.hx_stroke_flag cops2_prior laps2_het braden_total fac_dm_pc
fac_died_pc

stset time, failure(hapu)
stcox age_at_admit i.female ib2.admit_cat2 i.full_code
i.hx_diabetes_flag ///
i.hx_stroke_flag cops2_prior laps2_het braden_total fac_dm_pc
fac_died_pc, mgale(martinresid)

predict pval, hr
predict rval, csnell
twoway (scatter rval pval)
stcoxgof, arjas(4)
ovfplot

////////////////////////////////////////////////////////////////
// Sensitivity analysis for model with and without patient-level dm
```

```
and patient-level
// cops2. Does the effect of facility-level dm change?

//Full model
mestreg age_at_admit i.female ib2.admit_cat2 i.full_code
i.hx_diabetes_flag ///
i.hx_stroke_flag cops2_prior laps2_het braden_total fac_dm_pc ///
fac_died_pc || start_fac_id:, distribution(weibull) vce(robust)

//Model without hx_diabetes_flag
mestreg age_at_admit i.female ib2.admit_cat2 i.full_code ///
i.hx_stroke_flag cops2_prior laps2_het braden_total fac_dm_pc ///
fac_died_pc || start_fac_id:, distribution(weibull) vce(robust)

//Model without cops2_prior
mestreg age_at_admit i.female ib2.admit_cat2 i.full_code
i.hx_diabetes_flag ///
i.hx_stroke_flag laps2_het braden_total fac_dm_pc ///
fac_died_pc || start_fac_id:, distribution(weibull) vce(robust)

// Removing either patient-level covariate does not change the effect
of hospital-level
// dm covariate substantially

////////////////////////////////////////////////////////////
//frailties R code adapted from Steve Zuniga

# CNR Hospital-level residual variances pre/post adding facility co-
variates

cnr1 <- "I:/DOR/CNR/R hospital level residual variance.RData" ##
Derivation dataset

library(coxme)
hapu_cox <- Surv(cnr_derivation_file$time, cnr_derivation_file$hapu ==
1)   #set event time and HAPU outcome

# KPSC model (patient level mixed effects Cox proportional hazards
model)
m1  <- coxme(hapu_cox ~   age_at_admit + female + admit_cat +
                  full_code +  hx_diabetes_flag + hx_stroke_flag +
                  cops2_prior + laps2_het + braden_total +
                  (1|start_fac), data=cnr_derivation_file, x=TRUE,
y=TRUE)

fral <- data.frame(m1[[2]])
o <- order(fral, decreasing=F)
fral <- fral[o, 1, drop=FALSE]
write.table(fral, "clipboard1", sep="\t", row.names=T)
```

```
# KPNC model (patient level and hospital level mixed effects Cox PH
model)
m2  <- coxme(hapu_cox ~   age_at_admit + female + admit_cat +
               full_code +  hx_diabetes_flag + hx_stroke_flag +
               cops2_prior + laps2_het + braden_total +
               fac_dm_pc + fac_died_pc +
               (1|start_fac), data=cnr_derivation_file, x=TRUE,
y=TRUE)

fral <- data.frame(m2[[2]])
o <- order(fral, decreasing=F)
fral <- fral[o, 1, drop=FALSE]
write.table(fral, "clipboard2", sep="\t", row.names=T)
```