

Report: initial spatial-disease model

Mengru Yuan

August 20, 2018

Long-term objective

Identify disease patterns, estimate the strength of the relationship between chronic diseases, and provide evidence for implementing potential public health interventions.

Objective for this report

Get to know the data by fitting the initial model using INLA package.
The initial model is:

$$Y_{i,j} \sim \text{Poisson}(\lambda_{i,j})$$
$$\log(\lambda_{i,j}) = \log(n_{i,j}) + \beta_0 + \gamma_i + \delta_j$$

$Y_{i,j}$: number of incidence cases in 2014 for chronic disease j in region i

$n_{i,j}$: person-year at risk in 2014 for chronic disease j in region i

i : i_{th} CLSC region, from 1:57

j : j_{th} chronic disease, from 1:8

$$\gamma_i \sim \text{Normal}(0, \sigma^2)$$

$$\delta_j \sim \text{Normal}(0, \tau^2)$$

Data

The data include the incidence cases, person-year at risk for eight chronic diseases in 57 CLSC regions in Montreal CMA in 2014.

The eight chronic diseases are ami(Acute Myocardial Infarction), asthma, chf(Congestive Heart Failure), copd(Chronic Obstructive Pulmonary Disease), diabetes, hypertension, ihd(Ischemic heart disease), stroke

Dataset

```
clsc.dis = read.table(file = "Data/disease_clsc_incidence_2014.csv", head=TRUE, sep = ',')
clsc.dis = clsc.dis[, -1]
clsc.dis = clsc.dis %>%
  group_by(disease, clsc) %>%
  summarise(cases = sum(num), pop = sum(denom))

str(clsc.dis)
```

```
## Classes 'grouped_df', 'tbl_df', 'tbl' and 'data.frame':  456 obs. of  4 variables:
## $ disease: Factor w/ 8 levels "ami","asthma",...: 1 1 1 1 1 1 1 1 1 1 ...
## $ clsc   : int  6011 6012 6013 6021 6022 6031 6032 6033 6034 6041 ...
## $ cases  : int   24 25 12 37 46 14 38 31 16 23 ...
## $ pop     : int 14346 15571 9637 10979 15168 2896 13927 6606 6715 10948 ...
```

```
## - attr(*, "vars")= chr "disease"
## - attr(*, "drop")= logi TRUE

clsc.dis$dis_fct = as.numeric(clsc.dis$disease)
clsc.dis$clsc_fct = as.numeric(as.factor(clsc.dis$clsc))
```

Built the model with random effect on disease and clsc region

```
formula.0 = cases ~ 1 +
  f(dis_fct, model="iid") +
  f(clsc_fct, model="iid")
inla.model.0 = inla(formula.0, family="poisson", data=clsc.dis, offset=log(pop),
  control.compute=list(cpo=TRUE,dic=TRUE),
  control.predictor=list(compute=TRUE))
```

function for plot

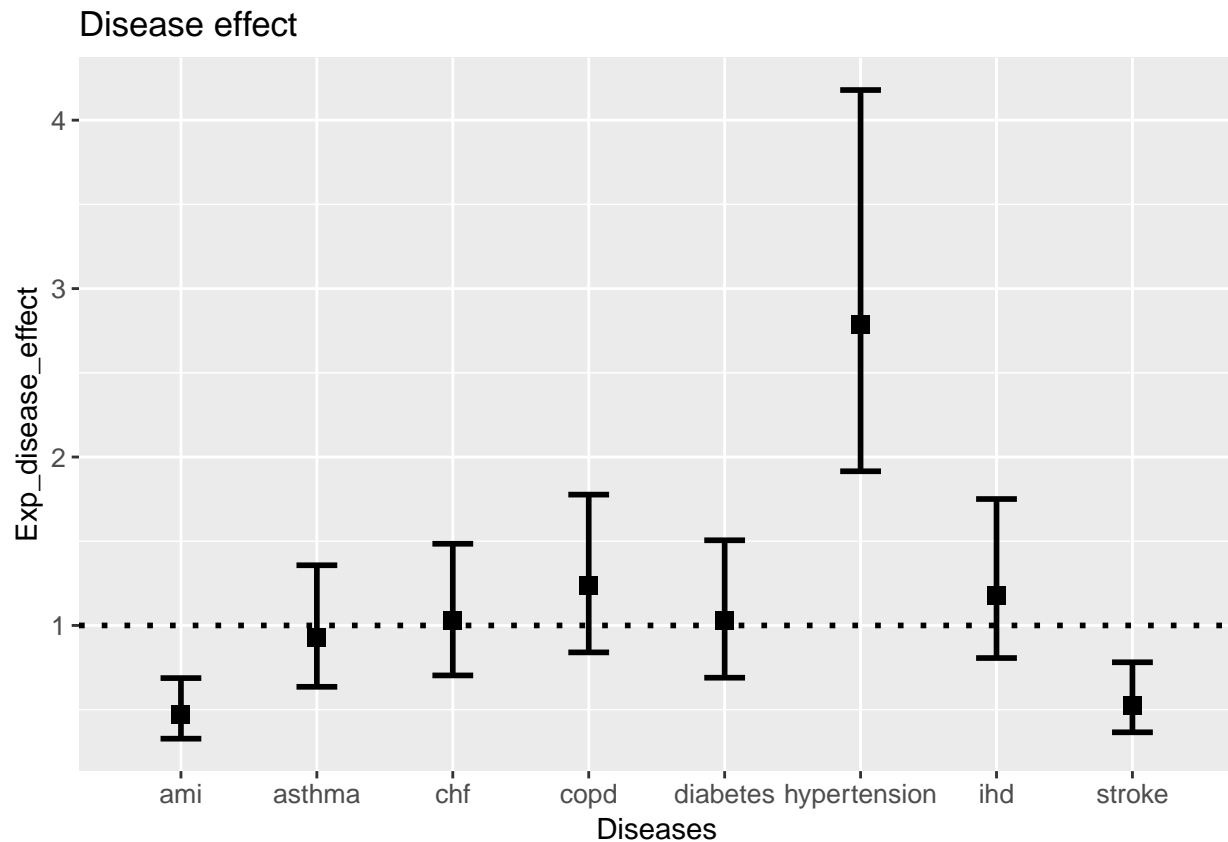
```
tbl_for_plot = function(inla_model, n, n_var){
  df = matrix(NA, 1000, n)
  for(i in 1:n){
    df[,i] <- inla.rmarginal(1000,marg = inla_model$marginals.random[[n_var]][[i]])
  }
  df.quartiles <- exp(t(apply(df, MARGIN=2,
    function(x) quantile(x, probs= c(0.025,0.5,0.975)))))

  data.test <- data.frame(varb = c(1:n),
    value = df.quartiles[,2],
    ui=df.quartiles[,3],
    li=df.quartiles[,1])

  return(data.test)
}
```

Results:

Plot the CI for posterior distribution of the random effect of each disease



plot the CI for posterior distribution of the random effect of each clsc

```
data = tbl_for_plot(inla.model.0, 57, 2)
data$clsc = unique(clsc.dis$clsc)
data = data[order(data$value), ]
data$y = c(1:57)

ggplot(data, aes(y, value)) +
  geom_point(size = 3, shape = 15) +
  geom_abline(intercept = 1, slope = 0, linetype = "dotted", size = 1) +
  geom_errorbar(aes(ymin = li, ymax = ui), size = 1, width = 0.3) +
  labs(x = "CLSC", y = "Exp_CLSC_effect", title = "Spatial effect") +
  scale_x_discrete(name = "CLSC", limits = c(1:57), labels = as.character(data$clsc)) +
  theme(axis.text = element_text(size = 5),
        axis.text.x = element_text(angle = 45, hjust = 1))
```

Spatial effect

