**Supplementary Material**

# S1 Key R code for analysis

#

require(readxl)

library(survival)

#

SI.data = read\_excel(path = 'comb\_data.xlsx', sheet = 'comb\_data', na = 'NA')

SI.data = as.data.frame(SI.data)

SI.data = SI.data[,1:9]

SI.data$Infector.date.lwr <- as.Date(SI.data$Infector.date.lwr,format ='%m/%d/%Y')

SI.data$Infector.date.upr <- as.Date(SI.data$Infector.date.upr,format ='%m/%d/%Y')

SI.data$Infectee.date <- as.Date(SI.data$Infectee.date,format ='%m/%d/%Y')

SI.data$mid.Infectee.date.index = (as.numeric(SI.data$Infector.date.lwr -as.Date('2019-12-31')) + as.numeric(SI.data$Infector.date.upr -as.Date('2019-12-31'))) /2

lwr.time.array = NULL

upr.time.array = NULL

surv.code.array = NULL

for (i in 1:nrow(SI.data)) {# i = 1

temp.data = SI.data[i,]

temp.lwr = as.numeric(temp.data$Infectee.date - temp.data$Infector.date.upr)

temp.upr = as.numeric(temp.data$Infectee.date - temp.data$Infector.date.lwr)

temp.lwr = ifelse(temp.lwr <=0, 0.5, temp.lwr)

temp.upr = ifelse(temp.upr <temp.lwr, temp.lwr, temp.upr)

temp.code = ifelse(temp.upr > temp.lwr, 3, 1)

lwr.time.array = c(lwr.time.array, temp.lwr)

upr.time.array = c(upr.time.array, temp.upr)

surv.code.array = c(surv.code.array, temp.code)

}

SI.data$lwr.SI = lwr.time.array

SI.data$upr.SI = upr.time.array

SI.data$mid.SI = (SI.data$lwr.SI + SI.data$upr.SI) /2

SI.data$surv.code = surv.code.array

#

sel.data = SI.data

#

cor.test(sel.data$mid.Infectee.date.index, sel.data$mid.SI, method = 'p')

#

simple.lm = lm(log(mid.SI) ~ mid.Infectee.date.index +Infector.gender, data = sel.data)

simple.lm = lm(c(mid.SI) ~ mid.Infectee.date.index +Infector.gender, data = sel.data)

#

surv.obj = Surv(time = sel.data$mid.SI, event = rep(1, nrow(sel.data)))

cx.mod = coxph(surv.obj ~ mid.Infectee.date.index +Infector.gender, data = sel.data)