

Description of Methods and Issues of RSV Excess Mortality Estimation

Goals

This is a summary of the methods used for influenza excess mortality estimation using mortality data; these methods were originally also proposed for use in excess mortality estimations for RSV.

Excess mortality estimation

Excess mortality estimation represents a suite methods/statistical models that may be described as *ecological* as the unit of analysis is not *individuals*, but *population group-weeks*. The population groups can be certain age groups, typically <5, 5-17, 18-49, 50-64, 65+, nationally or by HHS region.

Serfling-type model for excess mortality due to influenza

Traditionally, weekly influenza excess mortality was estimating using the following model. This model Thompson et al. (2009) is sometimes referred to as “Serfling-type” model, in reference to Robert Serfling who used a Fourier series to estimate the baseline mortality function Serfling (1963):

$$\log \mu_k = \beta_0 + \beta_1 w_k + \beta_2 w_k^2 + \beta_3 w_k^3 + \beta_4 \sin t_k + \beta_5 \cos t_k + \beta_6 AH1P + \beta_7 AH3 + \beta_8 B \quad (1)$$

$$y_k \sim \text{Poi}(\mu_k N_k) \quad (2)$$

where $t_k = \frac{w_k}{52.25} 2\pi$, with w_k being the index week, and $AH1P, AH3$ and B are the viral surveillance indicators for influenza, typically the percent of samples tested with positive test results for influenza A(H3N2), A(H1N1)pdm09 and B, respectively, y_k is the number of deaths (e.g. respiratory and circulatory deaths, usually by age group) in week k and N_k is the population denominator. In R, this model could be implemented as follows. We look at data over six influenza seasons; the influenza indicators are actually the proportion positive multiplied by the ILI%. This type of indicator is sometimes referred to as Goldstein indicator (see Goldstein et al. (2011)):

```
filename <- 'data_National_2010_16_5ag.dat'
setwd('../') ## change to root directory
setwd('./RSVdata')
datarr <- data.frame(read.table(file = filename, header = T))
#####
### Serfling-type model, virologic indicators (Thompson et al., 2009) #####
#####
ag <- 5 ### Setting age group to 65+
agdata <- datarr[which(datarr$age==ag),]

AH1P <- agdata$AH1P
AH3 <- agdata$AH3
B <- agdata$B

mort <- agdata$rcu
pop <- agdata$pop
```

```

N <- length(mort)
time <- (1:N)/N
t <- time * N * 2* pi/52.25
#####
data <- data.frame(mort,time,t,AH1P,AH3,B,pop)
#####
Poimod <- glm(mort ~ time + time^2 + time^3 + sin(t) + cos(t) + AH1P + AH3 +
              B + offset(log(pop)), data = data,family = poisson(link = 'log'))

dataflu0 <- data.frame(mort,time,t,AH1P=AH1P*0,AH3=AH3*0,B=B*0,pop)
mortflu0 <- predict(Poimod, newdata = dataflu0, type = 'response')

EM <- sum(mort - mortflu0)
#####

```

In this case, the total number of excess deaths in people 65 and older due to influenza over six seasons is estimated at 1.5034559×10^5 . Confidence intervals for this estimate can be obtained using bootstrapping or implementing the model in a Bayesian framework (using flat priors).

Updated models for RSV or influenza/RSV mortality

To account for a baseline mortality function that is not exactly sinusoidal, regression splines can be used. To model influenza mortality, we have been including, in addition to the influenza indicators, also an RSV indicator. We have been using influenza and RSV indicators that are lagged by 0, 1 and 2 weeks. Here is the resulting Poisson model:

$$\log(\mu_k) = \beta_0 + \sum_{m=1}^M \alpha_m B_{k,m} + \sum_{l=0}^2 \beta_{1l} AH1P_l + \sum_{l=0}^2 \beta_{2l} AH3_l + \sum_{l=0}^2 \beta_{3l} B_l + \sum_{l=0}^2 \beta_{4l} RSV_l \quad (3)$$

,

where M is the number of knots, B is the basis matrix and $AH1P_l$, $AH3_l$, B_l and RSV_l are the influenza and RSV indicators lagged by l week(s). Instead of a Poisson model, a negative binomial model can be used. We fit this model using Markov chain Monte Carlo (MCMC) methods (Bayesian framework). Here, the goal is influenza excess mortality estimation:

```

#####
setwd('..')
setwd('./RSVcode')
source('Set_environmental_variables.R')
#####
variables7 <- c('EMRSV1tot', 'EMRSV2tot', 'EMRSV3tot', 'EMRSV4tot', 'EMRSV5tot', 'EMRSV6tot',
               'EMRSVtot', 'EMflu1tot', 'EMflu2tot', 'EMflu3tot', 'EMflu4tot', 'EMflu5tot',
               'EMflu6tot', 'EMflutot')
#####
nadapt <- 1000
niter <- 5000

for (ag in rev(1:5)){
  agdata <- datarr[which(datarr$age==ag),]
  RSV <- agdata$RSV
  flumort <- agdata$flu
  mort <- agdata$rcu - flumort
  pop <- agdata$pop
}

```

```

N <- length(mort)
time <- (1:N)/N
#####
ndf <- round((nknots + 1) * nseas/(nseas*52.5)*N)
nsarr <- ns(time,df = ndf)[,] ## defining basis matrix
#####
mod <- lm(log(mort/pop) ~ ns(time, df = ndf))
smod <- summary(mod)
coeffls <- as.numeric(smod$coefficients[,1])
nsinit <- coeffls[2:(ndf + 1)]
b0init <- coeffls[1]
#####
data <- list('N'=N,'ndf'=ndf,'ns'=nsarr,'y'=mort,'RSV'=RSV,'pop'=pop, 'flumort'=flumort,
            'seas1'=seas1,'seas2'=seas2,'seas3'=seas3,'seas4'=seas4,'seas5'=seas5,
            'seas6'=seas6)
#####

inits <- function() {
  list(
    b0=b0init,

    b10=0,
    b11=0,
    b12=0,

    b20=0,
    b21=0,
    b22=0,

    b=nsinit
  )}

setwd('..')
setwd('./RSVmodels')

j.model <- jags.model(file=model3.file,data=data, inits=inits, n.adapt=nadapt, n.chains=3)
j.samples<-coda.samples(j.model, variable.names=variables7, n.iter=niter, thin = 5)

codaarr <- rbind(j.samples[[1]],j.samples[[2]],j.samples[[3]],deparse.level=0)
assign(paste0("codaarr",ag),codaarr)

setwd('..')
setwd('./RSVMCMCOutput')
fname <- paste0('codaarr',ag,' RSV ',qual3,' ',nknots,' knots ps ',nadapt,
               ' ',round(niter/5*3),'.RData')
save(codaarr,file = fname)

cat(paste0('//nAge group ',ag,': done//n'))
}
#####

```

Instead of virologic surveillance indicators for influenza circulation we have used coded influenza mortality of the age group the analysis is conducted in. In that case, coded influenza mortality has to be subtracted from

the outcome count and added to the resulting excess mortality estimate. We usually use identity-link models, because of the the component causes of mortality are additive (total mortality = baseline mortality + excess mortality) The resulting model:

```
model {
  for(l in 1:2){
    flumort[l] ~ dpois(sigma[l]*pop[l])
  }

  for(k in 3:N) {
    lambdaBase[k] <- b0 + inprod(b,ns[k,])
    lambdaB[k] <- b0 + inprod(b,ns[k,]) + sigma[k]*b10 + sigma[k-1]*b11 + sigma[k-2]*b12
    muB[k] <- lambdaB[k]*r[k-2]*pop[k]
    muB0[k] <- lambdaB[k]*pop[k]
    muBase0[k] <- lambdaBase[k]*pop[k]

    lambdaRSV[k] <- RSV[k]*b20 + RSV[k-1]*b21 + RSV[k-2]*b22
    lambdaflu[k] <- sigma[k]*b10 + sigma[k-1]*b11 + sigma[k-2]*b12

    EMRSV[k] <- lambdaRSV[k]*r[k-2]*pop[k]
    EMflu[k] <- lambdaflu[k]*r[k-2]*pop[k] + flumort[k]

    mu0[k] <- (lambdaB[k] + lambdaRSV[k])*pop[k]

    mu[k] <- muB[k] + EMRSV[k]
    r[k-2] ~ dgamma(alpha,alpha)

    y[k] ~ dpois(mu[k])
    flumort[k] ~ dpois(sigma[k]*pop[k])
  }

  b0~dnorm(0,1.0E-6)I(0,)

  b10~dnorm(0,1.0E-6)
  b11~dnorm(0,1.0E-6)
  b12~dnorm(0,1.0E-6)

  b20~dnorm(0,1.0E-6)
  b21~dnorm(0,1.0E-6)
  b22~dnorm(0,1.0E-6)

  logalpha~dnorm(0,1.0E-6)
  alpha <- exp(logalpha)

  for (k in 1:ndf){b[k]~dnorm(0,1.0E-6)}
  for (l in 1:N){sigma[l] ~ dunif(0,1)}

  EMRSV1tot <- sum(EMRSV[seas1[1]:seas1[2]])
  EMRSV2tot <- sum(EMRSV[seas2[1]:seas2[2]])
  EMRSV3tot <- sum(EMRSV[seas3[1]:seas3[2]])
  EMRSV4tot <- sum(EMRSV[seas4[1]:seas4[2]])
  EMRSV5tot <- sum(EMRSV[seas5[1]:seas5[2]])
  EMRSV6tot <- sum(EMRSV[seas6[1]:seas6[2]])
}
```

```

EMRSVtot <- sum(EMRSV[3:N])

EMflu1tot <- sum(EMflu[seas1[1]:seas1[2]])
EMflu2tot <- sum(EMflu[seas2[1]:seas2[2]])
EMflu3tot <- sum(EMflu[seas3[1]:seas3[2]])
EMflu4tot <- sum(EMflu[seas4[1]:seas4[2]])
EMflu5tot <- sum(EMflu[seas5[1]:seas5[2]])
EMflu6tot <- sum(EMflu[seas6[1]:seas6[2]])

EMflutot <- sum(EMflu[3:N])
}

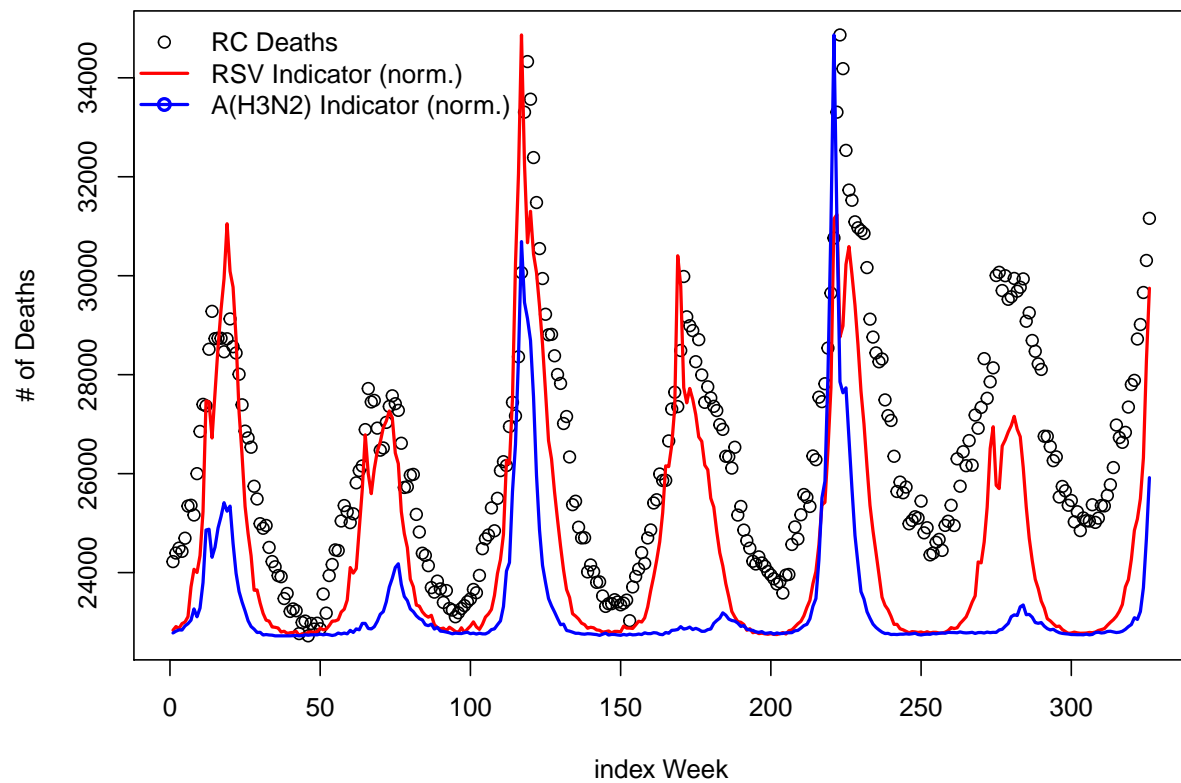
```

RSV excess mortality

As excess mortality is a function of the model parameters, it appears that this model could be used, exchangeably, for influenza and RSV excess mortality estimation.

Using the identity-link version of the above model, for age group 65+, from the season 2010/11 through 2015/16, the number of excess deaths due to influenza are estimated at 65966 (52561,79053), while deaths due to RSV are estimated at 207606 (162712,250552), roughly 3 times as many.

Even though the results from these ecological analyses are guesses, which rely on certain assumptions, such high numbers of RSV excess mortality are quite implausible. I believe that this is a result of the strong colinearity between the RSV indicator and RC mortality (correlation coefficient , p-value < 0.00001)



References

- Goldstein, Edward, Sarah Cobey, Saki Takahashi, Joel C Miller, and Marc Lipsitch. 2011. "Predicting the Epidemic Sizes of Influenza A/H1N1, A/H3N2, and B: A Statistical Method." *PLoS Medicine* 8 (7). Public Library of Science: e1001051.
- Serfling, Robert E. 1963. "Methods for Current Statistical Analysis of Excess Pneumonia-Influenza Deaths." *Public Health Reports* 78 (6). Association of Schools of Public Health: 494.
- Thompson, William W, Eric Weintraub, Praveen Dhankhar, Po-Yung Cheng, Lynnette Brammer, Martin I Meltzer, Joseph S Bresee, and David K Shay. 2009. "Estimates of Us Influenza-Associated Deaths Made Using Four Different Methods." *Influenza and Other Respiratory Viruses* 3 (1). Wiley Online Library: 37–49.