<1> Chapter 7: Practicing techniques in context 1: Examples from famine management

In this Chapter, we will take a brief pause from learning new techniques to reinforce several methods of modeling public health and healthcare systems that we've learned so far. There are at least three major types of modeling that we've detailed in the preceding chapters. First, we used *decision trees* to help us perform cost-benefit and cost-effectiveness. We extended the method to deal with complex disease dynamics using *Markov models*, which help us estimate the incidence and prevalence of disease in a population given complex probabilities of moving between different stages of disease. Finally, we complicated the Markov model to address infectious disease epidemics using *SIR models,* which simulate the process of disease transmission. In this Chapter, we'll practice all three of these key modeling techniques in a common context: designing and evaluating public health programs to address famine.

<2>Practice with cost-benefit analysis

Suppose we work for an international organization that helps provide famine relief internationally. Also suppose our problem is to decide how to best allocate our resources to reduce the health impact of an ongoing famine in India.

One persistent problem in famine management is the challenge of diagnosing children with severe acute malnutrition. Children with the condition can be successfully treated for the disease, but need to be caught and sent to dedicated medical care facilities because simple nutritional supplements are typically insufficient to enable their recovery.

A standard test for severe acute malnutrition is a paper tape measure (a "band test") that is color-coded and can be wrapped around the bicep of a child. The smaller the bicep, the more likely

a child is to have severe acute malnutrition. The test is typically performed by mobile community health workers who travel house to house in villages to test children and refer children with very small biceps to local health clinics for further diagnosis and treatment.

Suppose that for each group of 100 children, about 20 of them in your high-risk villages are truly malnourished and require $50 of treatment each. The band will pick up all 20 of them (suppose it has perfect sensitivity in this population), but the problem is that it will also pick up an additional 10 ("false positives", because of imperfect specificity), wasting an additional $50 of treatment resources per each additional child falsely diagnosed as having severe acute malnutrition.

A $5 prealbumin test has been proposed to be done after the band test to tell us whether the child was a false positive. The test is 90% correct at telling us whether the child was a false positive (10% chance of incorrectly telling us the child was a true positive on their band test), and does not affect true positive rates (it approves all true positive children).

Should we pay for the prealbumin test, if our objective is to minimize our total cost?

In this problem, we have to determine whether performing the prealbumin test will save costs overall, despite its imperfection. To solve the problem, we can calculate two quantities: the expected cost per child without the prealbumin test, and the expected cost per child with the prealbumin test.

We can draw a decision tree to help us estimate the expected cost per child without the prealbumin test. Figure 7.1 provides the decision tree for this example. In the tree, we have two possible situations: a child either has severe acute malnutrition (20% of the time), or not (80% of the time). If the child has the condition, there is a 100% chance that the band test detects them, and a 0% chance that the band test misses them. We then expect to pay $50 in treatment costs for each

positive child. If the child does not have the condition, there is a 10 in 80 (12.5% of the time) chance that the test falsely declares them to be positive, and a 70 in 80 (87.5% of the time) chance that the test correctly declares them to be negative. Among the false positive children, we expect to pay an additional $50 each.

[INSERT FIGURE 7.1 HERE]

We can roll-back the decision tree such that we calculate the expected cost at the root, which will be the expected cost of the top branch (0.2 x 1 x $50 + 0.2 x 0 x $0) plus the expected cost of the bottom branch (0.8 x 0.125 x $50 + 0.8 x 0.875 x $0), which equals $15 of expected cost per child.

Next, we can draw an expanded version of this decision tree to help us estimate the expected cost per child with the prealbumin test. The prealbumin test is only performed on those children testing positive, as shown in Figure 7.2, and we pay $5 for each prealbumin test. As shown in the Figure, among the true positives, the prealbumin test approves all children and we pay $50 in treatment for these children. Among the false positives, the prealbumin test approves 10% of children for treatment and we pay $50 for these false positives but not for the other 90% testing negative on the prealbumin test.

[INSERT FIGURE 7.2 HERE]

We can roll-back our decision tree such that we calculate the expected cost at the root, which will be (from top branch to bottom branch): 0.2 x 1 x 1 x ($50+$5) + 0.2 x 0 x 0 x ($0) + 0.8 x 0.125 x 0.1 x ($50+$5) + 0.8 x 0.125 x 0.9 x ($5) + 0.8 x 0.875 x $0, which equals $12 of expected cost per child.

Hence, if cost is our primary consideration, we would do better with the prealbumin test than without the prealbumin test, saving our budget $3 per child on average by purchasing the

prealbumin testing service.

<2> Practice with Markov models

Suppose we have a presentation at the United Nations next Monday and need to convince the "big wigs" that our program is making an impact on malnutrition.

We have a limited amount of data available to us to assess the impact of the program on individual children. Suppose that children can be in one of three possible states: healthy, mildly malnourished, or severely malnourished. Normally, without our program, about 5% of children per month go from being healthy to being mildly malnourished (none go straight to severe malnutrition); another 20% of children per month with mild malnutrition become healthy again through natural recovery processes, while 10% of those with mild malnutrition become severely malnourished per month; finally, 30% of children per month go from severe to mildly malnourished through natural recovery, and the rest stay in severe malnutrition (none go straight back to healthy).

Our program treats children with severe malnutrition. It increases the probability of transitioning from severely malnourished to healthy, from 0% per month to 50% per month. The other 50% of severely malnourished children in the area go from severely malnourished to mildly malnourished each month (none stay severely malnourished).

The United Nations officials will ask you: what's the impact of the program on the prevalence of severe malnutrition?

To answer this question, we can create a Markov model following the principles and strategies we detailed earlier in the book.

As with every Markov model, we can start by first drawing the model diagram, shown in Figure 7.3. Here, we draw the diagram before our intervention has been introduced.

[INSERT FIGURE 7.3 HERE]

To draw Figure 7.3, we first drew the three possible states, then the arrows connecting each possible state to each other possible state. We then filled in the probabilities per month on these arrows using the information provided in the problem. Lastly, we solved for the value of "self-loop" probabilities by reasoning that a person has probability 1 that they will either convert to another state or stay in the same state from month to month, hence 1 minus the rates of flow out of a state is the value of the probability of remaining within a given state from one month to the next. Note that we ignore mortality for our modeling example here, but can add in death as a fourth state to be more realistic.

Next, we write down our equations for this Markov model. In the pre-intervention case, we can write down a series of steady-state equations, which reflect the probability of being in a state in terms of the probability of coming into that state from one of the three states to the state being examined. A fourth equation is the equation that tells us the probability of being in any given state is equal to 1, meaning that no person can escape the model and the three states comprehensively describe all possible states of being with respect to our problem. These expressions are shown in Equations 7.5 through 7.8.

[Equation 7.5] $h = 0.95h + 0.2m + 0s$

[Equation 7.6] $m = 0.05h + 0.7m + 0.3s$

[Equation 7.7] $s = 0h + 0.1m + 0.7s$

[Equation 7.8] $h + m + s = 1$

In these equations, *h, m,* and *s* correspond to the probabilities of being healthy, or having

mild or severe malnutrition, respectively. We can solve these equations by hand, or (as we will see in a moment) solve them using *R*, giving the solution that $h = 0.75$, $m = 0.1875$, and $s = 0.0625$. Hence, before our intervention, there is about a 6% chance of being severely malnourished. In a village of 100,000 children, we would expect the long-term steady-state prevalence of severe malnutrition to be 100,000 x p_{severe} = 6,250 children.

How much difference will our program make to this prevalence? We can account for the change in the malnutrition epidemic by re-drawing our Markov model to account for the higher probability of becoming healthy from severe malnutrition, or reverting to mild malnutrition from severe malnutrition, as shown in Figure 7.4.

[INSERT FIGURE 7.4 HERE]

Note that in addition to accounting for the increased probability per month of transitioning from severe malnutrition to either the healthy or mildly malnourished state, we have also accounted for the reduction in the "self-loop" probability of staying severely malnourished, to account for the fact that the total probability of leaving or staying in the state remains at the value of 1. From the diagram, we can write the new set of expressions describing the steady state as Equations 7.9 through 7.12.

[Equation 7.9] $h = 0.95h + 0.2m + \mathbf{0.5}s$

[Equation 7.10] $m = 0.05h + 0.7m + \mathbf{0.5}s$

[Equation 7.11] $s = 0h + 0.1m + \mathbf{0}s$

[Equation 7.12] $h + m + s = 1$

We have highlighted, in bold, the key changes to our equations due to our intervention. If we solve this series of equations, we will find an estimate of $s = 0.0164$. Hence, the new long-term steady-state prevalence of severe malnutrition after our intervention would be expected to be

100,000 x s = 1,640 children. Our program has reduced the long-term prevalence of severe malnutrition by over 4,000 children.

What if we wanted to solve for periods other than the long-term steady-state? And what if we wanted to assess the cost-effectiveness of our program? Suppose we have the data available to us in Table 7.1 to detail the cost and utility, in quality-adjusted life-months, of being malnourished. What would be the incremental cost-effectiveness ratio of our intervention as compared to the baseline case of not having any intervention?

[INSERT TABLE 7.1 HERE]

We can create a program in *R* to solve this problem, can be downloaded from the textbook website. As we learned previously, we need three elements to program our Markov model: a transition matrix describing the probabilities of moving each month between each state and every other state, the number of timesteps we will simulate, and the initial states of the model that describe the initial probabilities of being in each state. Suppose we derive our pre-intervention transition matrix from Equations 7.5 through 7.8, simulate 120 months (a decade) of time, and assume all children start in the healthy state initially. Then for the pre-intervention code, we have:

```
transition = matrix(c(0.95, 0.2, 0,
                      0.05, 0.7, 0.3,
                      0, 0.1, 0.7),ncol=3,byrow=TRUE)
timesteps = 120
h = rep(0,timesteps)
m = rep(0,timesteps)
s = rep(0,timesteps)
h[1]=1
priorstate =c(h[1],m[1],s[1])
for (t in 2:timesteps)
{
  priorstate =c(h[t-1],m[t-1],s[t-1])
  newstate= transition%*%priorstate
  h[t]=newstate[1]
  m[t]=newstate[2]
  s[t]=newstate[3]
}
s[timesteps]
```

```
plot(s,xlab="Years",ylab="Severe malnutrition probability")
```

At the top of the code, we have the transition matrix for the pre-intervention period. We then provide the parameter for the number of timesteps (months) being simulated, then the starting empty vectors for each state. We specify that in the first timestep, *h* will be 1, and create the prior state vector of initial conditions. In a `for` loop, we then simulate each month by matrix multiplication of the transition matrix by the prior state vector, then update the new state vector for the duration of the simulated period. The final lines specify the probability of having severe malnutrition at the last timestep (month 120), and plots the probability of severe malnutrition over the period of the simulation, producing Figure 7.5.

[INSERT FIGURE 7.5 HERE]

We see in Figure 7.5 that the steady state probability of having severe malnutrition is about 6.25%. But we also see that it can take up to about 20 years before the steady state is reached.

Next we repeat the simulation with *_new* after each vector to indicate the new probabilities and vectors for simulating the post-intervention environment. The post-intervention transiton matrix is based on Equations 7.9 through 7.12, and produces the code:

```
transition_new = matrix(c(0.95, 0.2, 0.5,
                          0.05, 0.7, 0.5,
                          0, 0.1, 0),ncol=3,byrow=TRUE)
h_new = rep(0,timesteps)
m_new = rep(0,timesteps)
s_new = rep(0,timesteps)
h_new[1]=1
priorstate_new =c(h[1],m[1],s[1])
for (t in 2:timesteps)
{
  priorstate_new =c(h[t-1],m[t-1],s[t-1])
  newstate_new= transition_new%*%priorstate_new
  h_new[t]=newstate_new[1]
  m_new[t]=newstate_new[2]
  s_new[t]=newstate_new[3]
```

```
}
s_new[timesteps]
plot(s_new,xlab="Years",ylab="Severe malnutrition probability")
```

When implementing the post-intervention model, we see that the probability of severe malnutrition at steady state has dropped to only 1.88%. Hence, the long-term prevalence of severe malnutrition has dropped by over 4% due to the intervention. However, if we type `s - s_new` into *R*, then we see the difference made by our intervention in each year of the epidemic:

```
> s-s_new
  [1] 0.00000000 0.00000000 0.00000000 0.00350000 0.00822500 0.01309875
  [7] 0.01764656 0.02169040 0.02519401 0.02818505 0.03071639 0.03284755
 [13] 0.03463615 0.03613436 0.03738784 0.03843580 0.03931157 0.04004322
 [19] 0.04065437 0.04116482 0.04159112 0.04194715 0.04224446 0.04249275
 [25] 0.04270010 0.04287325 0.04301784 0.04313859 0.04323943 0.04332363
 [31] 0.04339395 0.04345267 0.04350171 0.04354265 0.04357685 0.04360541
 [37] 0.04362925 0.04364917 0.04366580 0.04367968 0.04369128 0.04370096
 [43] 0.04370905 0.04371580 0.04372144 0.04372615 0.04373009 0.04373337
 [49] 0.04373611 0.04373840 0.04374032 0.04374191 0.04374325 0.04374436
 [55] 0.04374529 0.04374607 0.04374672 0.04374726 0.04374771 0.04374809
 [61] 0.04374840 0.04374867 0.04374889 0.04374907 0.04374922 0.04374935
 [67] 0.04374946 0.04374955 0.04374962 0.04374968 0.04374974 0.04374978
 [73] 0.04374982 0.04374985 0.04374987 0.04374989 0.04374991 0.04374993
 [79] 0.04374994 0.04374995 0.04374996 0.04374996 0.04374997 0.04374997
 [85] 0.04374998 0.04374998 0.04374999 0.04374999 0.04374999 0.04374999
 [91] 0.04374999 0.04374999 0.04375000 0.04375000 0.04375000 0.04375000
 [97] 0.04375000 0.04375000 0.04375000 0.04375000 0.04375000 0.04375000
[103] 0.04375000 0.04375000 0.04375000 0.04375000 0.04375000 0.04375000
[109] 0.04375000 0.04375000 0.04375000 0.04375000 0.04375000 0.04375000
[115] 0.04375000 0.04375000 0.04375000 0.04375000 0.04375000 0.04375000
```

As shown, the intervention does not make any impact until about the fourth year of the malnutrition epidemic, and then accelerates in how much impact it makes until settling on about a 4% reduction in severe malnutrition prevalence.

How cost-effective is the intervention? We can calculate the incremental cost-effectiveness ratio (ICER) of our intervention as compared to the baseline condition of no intervention. We can see the input parameters in Table 7.1. Using these parameters, we can calculate the expected costs of our intervention before the intervention and after the intervention, and similarly calculate the

expected QALYs accumulated before and after the intervention. To do so, we multiply the costs or QALYs accumulated from being in each state by the probability vectors of being in each state, then sum across all states, and multiply by the population size to estimate the overall population's costs and QALYs across the simulation period.

```
costs = sum(100000*(m*10+s*(25+50)))
costs_new = sum(100000*(m_new*10+s_new*(25+50)))
qalys = sum(100000*(h*1+m*(1-0.1)+s*(1-0.3)))
qalys_new = sum(100000*(h_new*1+m_new*(1-0.1)+s_new*(1-0.3)))
icer = (costs_new-costs)/(qalys_new-qalys)
```

The resulting ICER equals -$266 per QALY. In other words, our intervention averts so many cases of severe acute malnutrition that it actually saves society money. This is one of those rare circumstances in which a public health intervention is cost-saving while also producing a gain in QALYs.

<2> Practice with infectious disease models

During periods of famine, the low nutritional status of many children unfortunately predisposes them to infections such as dysentery, which is bloody diarrhea that is often typically by infectious *Shigella* bacteria. Children who experience the illness are usually infectious for 1 week (during which the rate of death is 0.1 per week), then recover briefly and experience temporary immunity for another 130 weeks before losing that immunity. Suppose a new vaccine has been developed that also provides temporary immunity to *Shigella*, lasting 65 weeks. How many deaths among children from dysentery would a vaccination campaign prevent over a one-year period?

To solve this problem, we can first draw the state diagram illustrating the natural history of dysentery before the vaccine is introduced. As illustrated in Figure 7.6, we can extend the typical *SIR* model of disease by adding a line indicating that the recovered people in the *R* state shift back to the *S* state of susceptibility gradually, reflecting their waning immunity. We additionally label the rate of deaths among infected people *I*.

[INSERT FIGURE 7.6 HERE]

Equation 7.13 through 7.16 describe the model's transition between states of susceptibility, infection, recovery, and death from dysentery only (state *D*), with the rate of waning immunity labeled ω and the rate of death from dysentery labeled δ:

[Equation 7.13] $$\frac{dS(t)}{dt} = \mu N(t) + \omega R(t) - \beta I(t)S(t) - \mu S(t)$$

[Equation 7.14] $$\frac{dI(t)}{dt} = \beta I(t)S(t) - vI(t) - \mu I(t) - \delta I(t)$$

[Equation 7.15] $$\frac{dR(t)}{dt} = vI(t) - \mu R(t) - \omega R(t)$$

[Equation 7.16] $$\frac{dD(t)}{dt} = \delta I(t)$$

Note that the number of people per month $\omega R(t)$ who experience waning immunity transitions people from the recovered to the susceptible state. The number of people $\delta I(t)$ transitions people from the infected to the dead state.

Suppose we have a population of 100,000 people with a typical life expectancy of 75 years, among whom *Shigella* dysentery has a typical transmission rate of β = 0.001 per person per week. Then putting all the parameters in the units of weeks, we have the following code for the model:

```
N = 100000
mu = 1/(75*52)
w = 1/130
beta = 0.001
v = 1
```

```
d = 0.1

time = 52
dt = 0.01

S = 99999
I = 1
R = 0
D = 0

Svec = S
Ivec = I
Rvec = R
Dvec = D


for (i in 1:time){
  for (i in 1:(1/dt)){
    S = S + mu*N*dt + w*R*dt - beta*S*I*dt - mu*S*dt
    I = I + beta*S*I*dt - v*I*dt - mu*I*dt - d*I*dt
    R = R + v*I*dt - mu*R*dt - w*R*dt
    D = D + d*I*dt
    Svec = c(Svec, S)
    Ivec = c(Ivec, I)
    Rvec = c(Rvec, R)
    Dvec = c(Dvec, D)
    N = S+I+R
  }
}

max(Dvec)
```

The final line indicates the total number of people who have died over the course of one year, which equals 8,058 people in the absence of a vaccine.

Now we can simulate the model after the vaccine is introduced. Suppose that a vaccination campaign reaches f = 50% of the susceptible population. Equations 7.17 through 7.21 describe the updated model's transition between states of susceptibility, infection, recovery, death, and a new state called vaccinated (state M), which has to be separate from the recovered state because it has a different rate of waning immunity (let's call it φ) and should not be confused with those people who have natural immunity.

[Equation 7.17] $$\frac{dS(t)}{dt} = (1-f)\mu N(t) + \varphi M(t) + \omega R(t) - \beta I(t)S(t) - \mu S(t)$$

[Equation 7.18] $$\frac{dI(t)}{dt} = \beta I(t)S(t) - vI(t) - \mu I(t) - \delta I(t)$$

[Equation 7.19] $$\frac{dR(t)}{dt} = vI(t) - \mu R(t) - \omega R(t)$$

[Equation 7.20] $$\frac{dD(t)}{dt} = \delta I(t)$$

[Equation 7.21] $$\frac{dM(t)}{dt} = f\mu N(t) - \mu M(t) - \varphi M(t)$$

As illustrated in Figure 7.7, we can extend our previous model diagram to account for the new vaccine campaign, by adding an additional vaccinated state and drawing an arrow to indicate that there is a rate of waning immunity from that state to the susceptible state.

[INSERT FIGURE 7.7 HERE]

Noting that we'll have a fraction *f* of the initial population vaccinated, we arrive at the following code:

```
N = 100000
mu = 1/(75*52)
w = 1/130
beta = 0.001
v = 1
d = 0.1
f = 0.5
phi = 1/65

time = 52
dt = 0.01

S = 99999*(1-f)
I = 1
R = 0
D = 0
M = 99999*f

Svec = S
Ivec = I
Rvec = R
Dvec = D
Mvec = M


for (i in 1:time){
  for (i in 1:(1/dt)){
```

```
          S = S + (1-f)*mu*N*dt +phi*M*dt + w*R*dt - beta*S*I*dt -
mu*S*dt
          I = I + beta*S*I*dt - v*I*dt - mu*I*dt - d*I*dt
          R = R + v*I*dt - mu*R*dt - w*R*dt
          D = D + d*I*dt
          M = M + f*mu*N*dt - mu*M*dt - phi*M*dt
          Svec = c(Svec, S)
          Ivec = c(Ivec, I)
          Rvec = c(Rvec, R)
          Dvec = c(Dvec, D)
          Mvec = c(Mvec, M)
          N = S+I+R
        }
      }

      max(Dvec)
```

The total number of children who die over the course of one year (`max(Dvec)`), after the vaccine is introduced, now equals 7,743 children. In other words, even with an imperfect vaccine, about 314 children would be saved from death by dysentery.