Estimation of the effective reproduction number of COVID-19 outside China

February 16, 2020

What is the average R_{eff} outside of China?

At this writing (16 February 2020) a number of cases of COVID-19 have been exported outside of China. Some of these have led to secondary transmission, while others have not. Some important questions include

1. Is the current level of vigilance adequate for detection and containment?
2. What potential is there for sustained transmission in these settings?

However, because there is relatively little data from secondary transmission chains, it is not possible to estimate the effective reproduction number, R_{eff}, in these settings directly. Nonetheless, it seems reasonable to think that the relative number of secondary transmission events contains information about the effective reproduction number in these settings. The following is a calculation to estimate the average effective reproduction number of cases exported outside China.

Calculation for perfect observation

During the early phases of an SIR epidemic, the increase in the number of infected cases may be approximated by a birth-death process where the birth rate is given by the transmissibility β and the death rate is given by the recovery rate γ. It is understood from context that both β and γ are considered at their effective values and may be different from the natural transmissibility and recovery rates due to social distancing, quarantine, active case identification and isolation, etc.

A case exported to a susceptible population may give rise to secondary infections or may be isolated before such infections occur. The probability p_0 that no secondary infection occurs is given by the proportional propensity $p_0 = \frac{\gamma}{\beta+\gamma}$. Taking the reciprocal of each side of this equation, we have

$$\frac{1}{p_0} = \frac{\beta + \gamma}{\gamma} = \frac{\beta}{\gamma} + \frac{\gamma}{\gamma} = R_{\text{eff}} + 1 \tag{1}$$

A simple rearrangement yields

$$R_{\text{eff}} = \frac{1}{p_0} - 1. \tag{2}$$

The probability p_0 may be estimated from the fraction of exported cases that have not led to secondary transmission leading to

$$R_{\text{eff}} = \frac{X}{Y} - 1 \tag{3}$$

where X is the total number of exported cases and Y is the number of exported cases without secondary transmission. Note that X and Y are both observed.

The WHO Situation Report for 13 February 2020 reports that

> As of 10am CET 13 Feb 2020, a total of 170 cases of COVID-19 who had a travel history to China have been reported outside of China. The vast majority of these (151, 89%) do not appear to lead to further transmission of the virus, while the remaining 19 have been associated with onward transmission within 12 distinct groups of epidemiologically linked cases.

Thus, we estimate that the realized effective reproduction number outside China during the period of these reports was $R_{\text{eff}} = \frac{170}{151} - 1 = 0.1258278....$ By resampling from these observations with replacement, we estimate the uncertainty around this estimate with a confidence interval of $[0.07, 0.17]$.

Effective reproduction number outside China
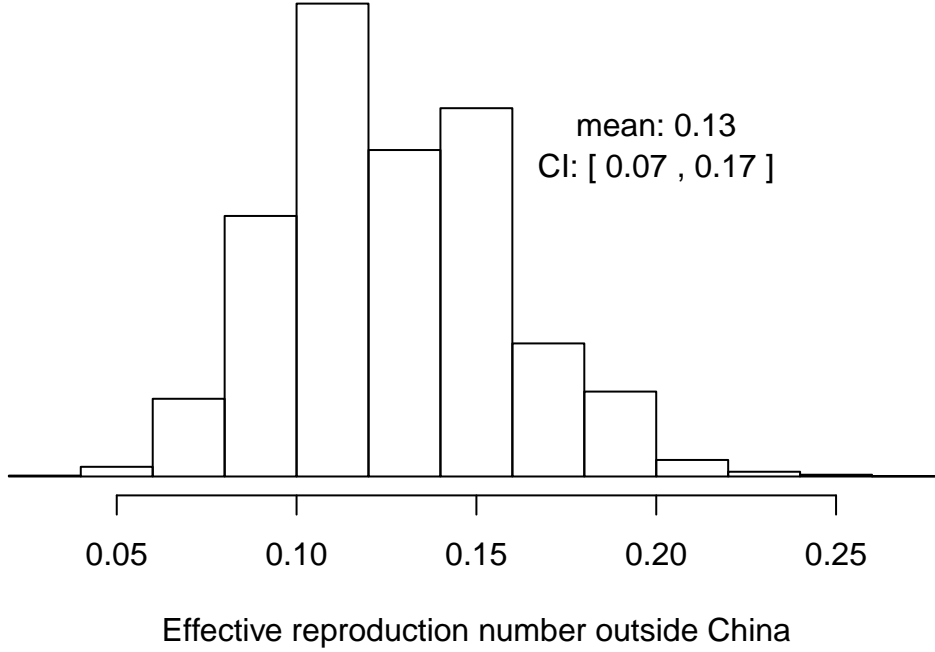
Assumptions of this calculation include:

1. That all exported transmission chains have been observed, i.e. that there are no significant ongoing, undetected chains of transmission elsewhere in the world.
2. This estimate for R_{eff} is only valid for the set of locations for which exported cases are reported. Particularly, places that differ substantially with respect to surveillance capacity, vigilance, and ability to isolate will be more vulnerable than this calculation suggests. Moreover, we assume that β and γ are similar in all reporting locations.
3. That all chains determined to have ended with no secondary transmission are correctly classified.

Calculation with imperfect observation

The first and third assumptions can be investigated further. Recall that equation 3 has the form

$$R_{\text{eff}} = \frac{\text{total exported cases}}{\text{exported cases without secondary transmission}} - 1. \tag{4}$$

Next, we allow that neither exported cases nor the secondary chains of transmission that they generate may be perfectly observed. That is, we distinguish *detected* from *undetected* exported cases and *detected* from *undetected* cases among contacts of exported cases. We denote the fraction of exported cases that are detected by $0 \le d_1 \le 1$ so d_1 is the *detection rate of exported cases*. We assume that cases among contacts of exported cases are even less likely to be detected than the exported cases themselves. This seems reasonable under current surveillance strategies that closely observe travelers from regions with active transmission. Thus, we denote the reduction in detection of transmission chains among contacts of exported cases by a factor $0 \le d_2 \le 1$ so that the *detection rate among contacts of exports* is $d_1 \times d_2$.

Equation 4 should now be written in terms of the observables $X = 170$ and $Y = 151$ in addition to unobserved quantities U_1 (the number of undetected exported cases), U_2 (detected exported cases with undetected chains of secondary transmission), and U_3 (undetected exports with no undetected chains of transmission). Now we have

$$R_{\text{eff}} = \frac{X + U_1}{Y - U_2 + U_3} - 1. \tag{5}$$

Formulas for U_1 and U_2 in terms of X, Y, d_1 and d_2 are

$$U_1 = \left(\frac{1 - d_1}{d_1}\right) \times X \tag{6}$$

and

$$U_2 = (1 - d_1 d_2) \times Y. \tag{7}$$

A formula for U_3 comes from multiplying the total number of undetected exports, U_1, by the probability p_0 of not generating any secondary chains of transmission.

$$U_3 = X \times \left(\frac{1 - d_1}{d_1}\right) \times p_0. \tag{8}$$

Writing equation 8 in terms of R_{eff} instead of p_0 we have

$$U_3 = X \times \left(\frac{1 - d_1}{d_1}\right) \times \left(\frac{1}{R_{\text{eff}} + 1}\right). \tag{9}$$

Putting it all together we have the equation

$$R_{\text{eff}} = \frac{X + \left(\frac{1 - d_1}{d_1}\right) X}{Y - (1 - d_1 d_2) Y + X \left(\frac{1 - d_1}{d_1}\right) \left(\frac{1}{R_{\text{eff}} + 1}\right)} - 1. \tag{10}$$

I haven't tried manipulating this formula. Presumably it is transcendental and R_{eff} must be found numerically by solving for the root of $g = f - R_{\text{eff}}$.

```
g <- function(Reff, d1=1, d2=1){
  X <- 170
  Y <- 151
  U1 <- X*(1-d1)/d1
  U2 <- Y*(1-d1*d2)
  U3 <- U1*(1/(Reff+1))
  g <- (X+U1)/(Y-U2+U3) - 1 - Reff
  return(g)
}
```

Here we test the function assuming perfect observation by setting $d_1 = 1$ and $d_2 = 1$. In this limit case, the numerical solution agrees with the exact solution to at least seven decimal places.

```
print(solution <- uniroot(g, c(0.0001,500), d1=1, d2=1))
```

```
## $root
## [1] 0.1258278
##
## $f.root
## [1] 0
##
```

```
## $iter
## [1] 1
##
## $init.it
## [1] NA
##
## $estim.prec
## [1] 499.8742
```

Now, we solve over a range of d_1 and d_2. Clearly, the failure to detect exported cases or cases among contacts of exported cases yields a biased estimate of R_{eff}. The heavy solid line indicates the critical boundary at $R_{\text{eff}} = 0$. Given that there have not been a singificant number of cases identified without a plausible exposure scenario (e.g. travel history to China or exposure to a known case), it seems reasonable to assume that d_1 is relatively close to 1. It is hoped that d_2 has also remained close to 1, for instance through effective contract tracing. Assuming $d_1 = d_2 = 0.9$, this estimator yields an average R_{eff} outside of China of 0.39.

```
get.root <- function(d1, d2){
  out <- tryCatch(uniroot(g, c(0.0001,50), d1, d2)$root, error=function(x) 0)
  return(out)
}


d1 <- seq(0.0, 1, length=500)
d2 <- seq(0.0, 1, length=500)

R <- array(dim=c(length(d2), length(d1)))
for(i in 1:length(d1)){
  for(j in 1:length(d2)){
    R[j,i] <- get.root(d1[i],d2[j])
  }
}
```

```
image(log(R), ylab='Detection rate among exported cases (d1)',
      xlab='Reduction of detection among contacts of exported cases (d2)',
      main='Effective reproduction number under imperfect detection',
      axes=FALSE)
axis(2, at=seq(0,1, length=11), labels=seq(0.0, 1, length=11), las=2)
axis(1, at=seq(0,1, length=11), labels=seq(0.0, 1, length=11), las=1)
box()
lev <- c(-2.5, -2, -1.5, -1, -0.5, 0.5, 1, 1.5, 2, 2.5, 3, 3.5)
contour(log(R), levels=lev, labels=round(exp(lev),1), add = TRUE, drawlabels = TRUE)
contour(log(R), levels=0, labels = 1, add = TRUE, drawlabels = TRUE, lwd=3)
```

Effective reproduction number under imperfect detection

Detection rate among exported cases (d1)

Reduction of detection among contacts of exported cases (d2)

0.2
0.4
0.6
1
1.6
2.7
4.5
7.4
12.2
20.1
33.1