



전염병 모델과 COVID-19

2020-26

김종우 선임연구위원

함건희 선임연구위원

아산정책연구원

2020.08.31

한국은 지난 1 월 20 일 COVID-19 첫 확진자가 발생한 뒤 현재까지 19,077 명의 누적 확진자가 발생했으며 이 중 4,210 명이 격리 중이며 316 명이 사망하였다.¹ 이후 2 월 18 일 31 번 확진자² 발생을 기점으로 확진자 수가 기하급수적으로 증가하였고 그 결과 2 월 말 기준으로 COVID-19 진원 국가인 중국 다음으로 가장 많은 2,337 명을 기록하는 오명을 남겼다. 질병관리본부와 의료진의 헌신 그리고 국민들의 적극적인 방역 참여로 확산세가 제어 불능 상태에 도달하는 최악의 상황을 피할 수 있었으나 최근 또 다시 COVID-19 가 재 확산 조짐이 나타나고 있다. 이번 이슈브리프에서는 대표적인 전염병 모델링 (Epidemic Modeling) 방법 중 하나인 *SIR* 모델을 활용해 향후 전개될 상황을 예측해보았다. *SIR* 모델과 같은 전염병 모델링 방법은 현재 해외에서도 COVID-19 와 관련해 활발하게 사용하고 있다.³

질병관리본부에서는 한 명의 감염자가 평균적으로 감염시킬 수 있는 2 차 감염자 수를 나타낸 기초감염재생산수 (Basic reproductive number)를 COVID-19 의 활동성 지표로 사용하고 있는데 이를 세 가지 상황을 가정해 *SIR* 모델로 예측한 결과는 아래 [표 1]과 같다.

기초감염재생산수	9 월	10 월	11 월	12 월
4.32	659,327	21,529,437	10,428,556	2,464,172
1.67	12,484	34,600	92,389	251,917
0.33	1,371	495	172	60

[표 1] *SIR* 모델로 예측한 시기별 격리 중 인구 수 (단위: 명, 월말 기준)

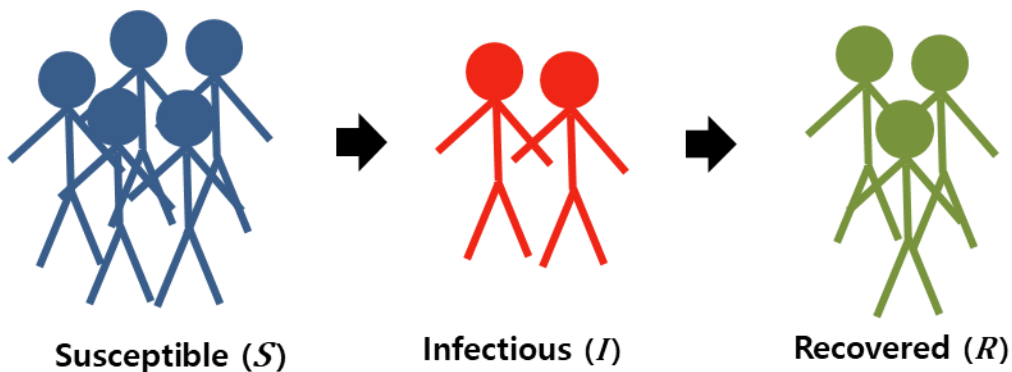
보이는 바와 같이 상황에 따라 결과의 차이가 매우 큰 것을 확인할 수 있다. 최근의 증가 추세가 계속 유지된다고 가정해 기초감염재생산수 4.32 을 적용한 경우 올해 연말 기준 격리 중 인구 수가 약 246 만명까지 증가하는 반면 '고강도 사회적 거리두기'를 시행한 지난 3 월 22 일부터 5 월 5 일까지의 기초감염재생산수 0.33 을 적용한 경우 올해 연말 기준 격리 중 인구 수가 60 명까지 줄어들게 된다. 지난 8 월 23 일 중앙방역대책본부 정례 브리핑에서 정은경 본부장이 질의응답에서 밝힌 전국 단위 기초감염재생산수 1.67 을 적용한 경우 올해 연말 기준 격리 중 인구수는 약 25 만명으로 예상된다.

이처럼 *SIR* 모델과 같은 전염병 모델을 사용하면 전염병이 확산되는 추세를 상황별로 예상해 볼 수 있다. 이제 *SIR* 모델에 대한 소개와 함께 한국의 상황에 실제 적용해 보고 그 결과를 통해 앞으로 어떤 방식으로 COVID-19 확산에 대응해야 할지 이야기해보고자 한다.

SIR 모델

SIR 모델은 1927 년 영국의 W. Kermack 과 A. McKendrick 에 의해 발표된 모델로 어떻게 전염병이 확산될 수 있는지를 수학적으로 보여준다⁴ 감염 가능한 인구(*S*, Susceptible)와 감염된 인구(*I*, Infectious), 감염 후 회복된 인구(*R*, Recovered) 총 세 개의 구획으로 이루어져 있으며,⁵ 각 개개인은 세 구획 중 한 곳에 속해 아래의 [그림 1]에서 보여주는 것과 같이 $S \rightarrow I \rightarrow R$ 순서로 이동하는 개념적으로 단순한 구조를 갖는다.

[그림 1] *SIR* 모델의 구획



SIR 모델에서는 감염된 인구 I 의 비율은 감염 가능한 인구 S 와 감염된 인구 I 사이의 접촉에 β 만큼 일정하게 비례하며 회복된 인구 R 의 비율은 감염된 인구 I 에 γ 만큼 일정하게 비례한다고 가정한다.⁶ 만약 회복되지 못하고 사망했다면 회복된 인구 R 에 속하며 한번 회복된 감염자는 다시 감염되지 않는다는 것을 전제로 한다.⁷ 이를 수식으로 표현하면 다음 세 개의 미분방정식으로 표현된다.⁸

$$\frac{dS}{dt} = -\frac{\beta}{N} \cdot S \cdot I$$

$$\frac{dI}{dt} = \frac{\beta}{N} \cdot S \cdot I - \gamma \cdot I$$

$$\frac{dR}{dt} = \gamma \cdot I$$

첫 번째 수식은 S 구획의 감염되지 않았던 인구가 얼마나 빠르게 감염돼 I 구획으로 이동하는지를, 두 번째 수식에서는 S 구획에서 감염된 인구가 I 구획으로 들어옴과 동시에 얼마나 빨리 회복되어 R 구획으로 이동하는지를, 마지막 수식에서는 I 구획 내 감염된 인구가 얼마나 빠르게 회복돼 R 구획으로 이동하는지를 나타낸다.⁹

이 수식들을 통해 전염병 확산과 관련해 다음과 같은 정보를 얻을 수 있다. 첫째, 전염병이 빠르게 확산될 것인지 아니면 줄어들 것인지에 대해 알 수 있다. 둘째, 전염병 종결 시점과 더불어 감염되지 않은 인구를 파악할 수 있다. 셋째, 감염된 인구의 최대값과 발생 시점 등을 계산할 수 있다.

앞서 언급한 COVID-19의 활동성을 측정하는 지표로 질병관리본부가 활용하고 있는 기초감염재생산수 R_0 (Basic reproductive number)를 통해 *SIR* 모델을 구체적으로 알아보자. R_0 는 감염자가 없는 인구집단에 처음으로 감염자가 발생하였을 때 한 명의 감염자가 평균적으로 감염시킬 수 있는 2차 감염자 수를 나타낸 것이다. 이는 *SIR* 모델에서는 다음과 같이 표현된다.

$$R_0 = \frac{\beta}{\gamma}$$

즉, 감염률과 회복률의 비로 R_0 가 커질수록 전염병이 확산되며, 작아질수록 전염병이 감소한다는 것을 직관적으로 해석할 수 있다. 이를 SIR 모델 안에서 자세하게 들여다보자. 전염이 시작된 시점에는 감염 초기이므로 I 구획의 인구 수는 극소수임으로 N 과 S 구획의 감염 초기 인구 수 S_0 는 거의 비슷하다 ($N \sim S_0$). 또한 회복된 인구가 없기 때문에 R 구획의 인구 수는 0 명이며 S 구획의 인구 수는 감염 초기 시점이 최대이다.¹⁰ 최종적으로 감염 초기상황에서 앞의 두 번째 수식을 다음과 같이 나타낼 수 있다.¹¹

$$\frac{dI}{dt} < I \cdot \left(\frac{\beta}{N} \cdot S_0 - \gamma \right)$$

여기서 괄호 안 식의 부호('+' 또는 '-')에 의해 전염병이 확산될 것인지 아닌지 결정된다.

① 전염병이 확산되는 경우, '+'

$$\frac{\beta}{N} \cdot S_0 - \gamma > 0 \Rightarrow \frac{\beta}{\gamma} > \frac{N}{S_0} \sim 1 \Rightarrow R_0 > 1$$

② 전염병이 줄어드는 경우, '-'

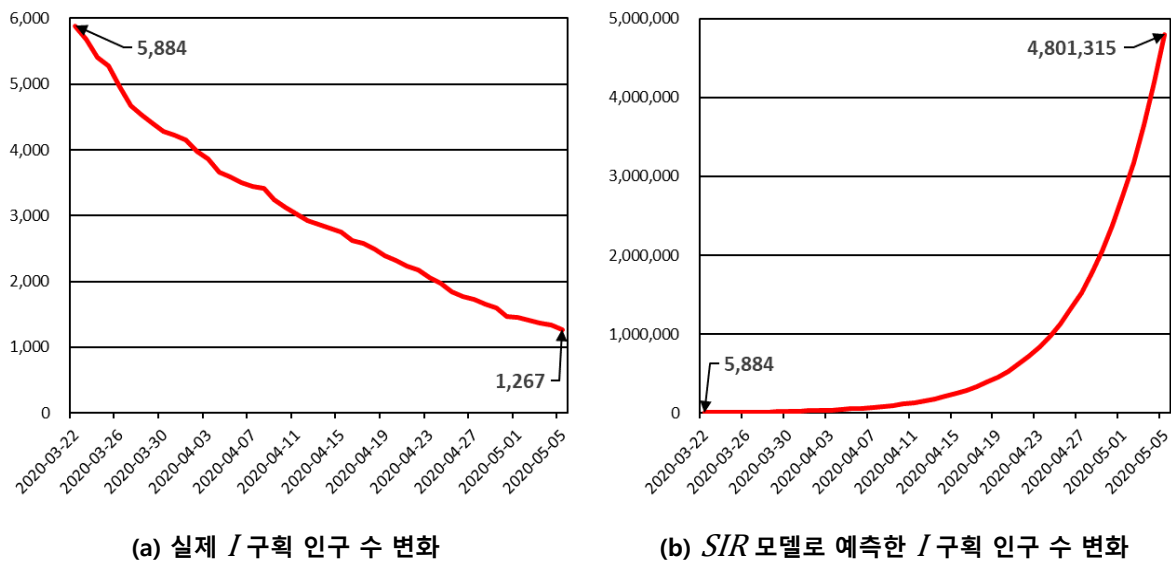
$$\frac{\beta}{N} \cdot S_0 - \gamma < 0 \Rightarrow \frac{\beta}{\gamma} < \frac{N}{S_0} \sim 1 \Rightarrow R_0 < 1$$

R_0 값이 1 을 초과할 경우 전염병은 퍼지며 미만일 경우에는 줄어든다. 수식으로 보면 β 값을 줄이거나 γ 값을 키워 R_0 값을 1 미만으로 유지시킬 수 있다. β 값을 감소시키는 방법으로는 '사회적 거리두기'를 통해 불필요한 만남을 줄여 평균 접촉자수를 감소시키거나 마스크 착용, 손 씻기, 기침할 때 옷소매로 입과 코 가리기 등의 예방수칙을 준수하는 방법 등이 있으며 γ 값을 키우는 방법으로는 COVID-19 치료제가 개발되어 회복률을 높이는 방법이 있다. 다만 효과적인 치료제가 가까운 시일 내에 나오길 기대하기는 힘든 상황이기 때문에 현실적으로는 β 값을 감소시키는 것이 현 상황에서는 최선이다. S_0 자체를 COVID-19 백신을 통한 예방으로 감염자 수를 줄이는 방법도 있으나 치료제와 마찬가지로 아직 개발 단계에 있으며 현재 검증 완료된 백신은 없다.

SIR 모델로 바라본 '사회적 거리두기'의 방역 효과

그렇다면 과연 '사회적 거리두기' 등의 조치가 얼마나 β 값을 줄일 수 있는지 확인해보자. 한국 정부는 지난 2월 COVID-19 1차 대유행 시 감염 확산을 통제하기 위해 3월 22일을 기점으로 5월 6일 '생활 속 거리두기'로 전환하기 직전까지 총 45일간 '고강도 사회적 거리두기'를 시행하였다. 이 기간 동안 실제 I구획의 인구 수 변화와 만약 '고강도 사회적 거리두기'를 시행하지 않았다고 가정했을 때 SIR 모델로 I구획의 인구 수 변화를 예측해 비교한 결과는 아래 [그림 2]와 같다.¹²

[그림 2] '고강도 사회적 거리두기' 시행 유무에 따른 I구획 인구 변화 (단위: 명)¹³

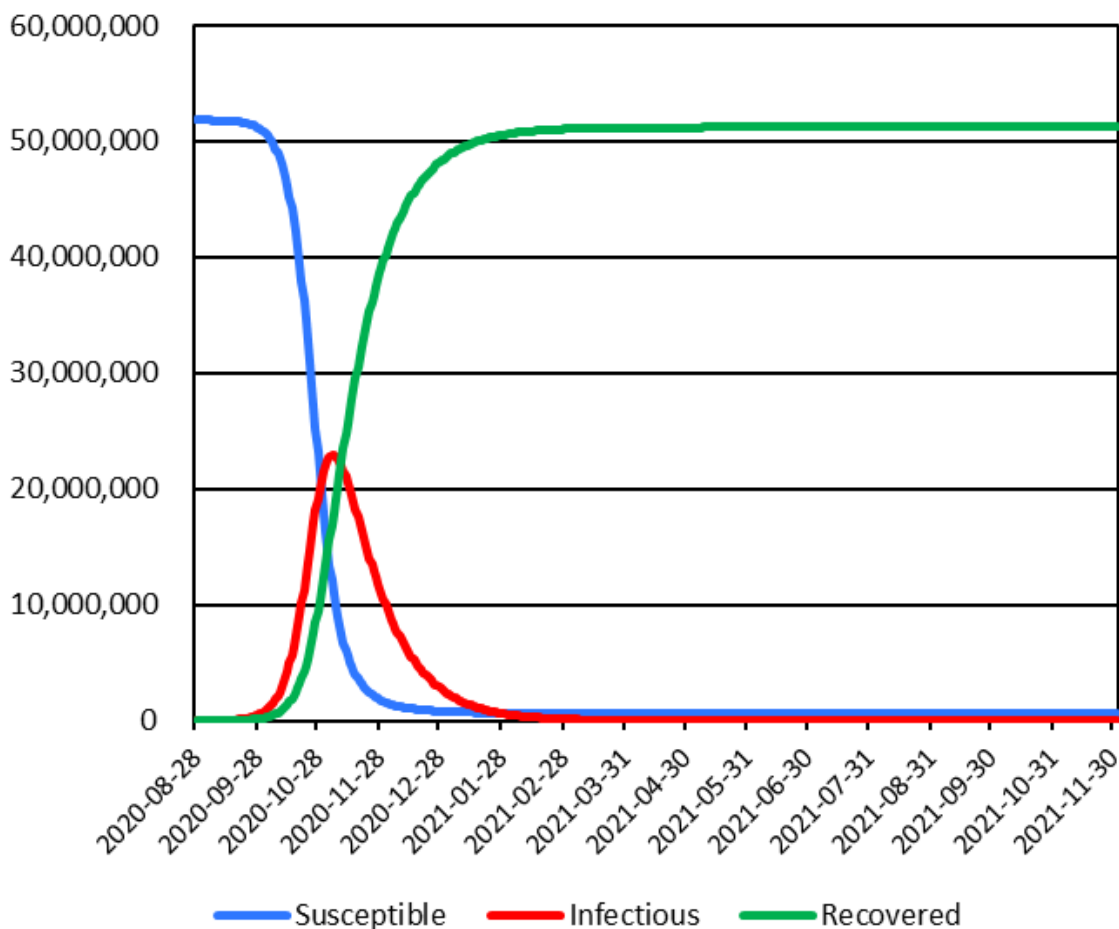


[그림 2] (a)는 3월 22일부터 '고강도 사회적 거리두기' 고강도 사회적 거리 두기가 시행된 날 이후 실제 I구획의 인구 수 변화, (b)는 '고강도 사회적 거리두기'를 시행하지 않았을 경우를 가정해 SIR 모델로 예측한 I구획의 인구 수 변화이다. 만약 '고강도 사회적 거리두기' 이전의 증가추세 그대로 COVID-19 확산이 지속되었다면 (b)의 예측 결과에서 확인할 수 있듯이 COVID-19 감염자가 기하급수적으로 늘어나 최악의 상황을 맞았을 것이다. 하지만 '고강도 사회적 거리두기'의 시행으로 실제로는 시행 초기 5,884 명에서 종료시점 기준 1,267 명으로 총 4,617 명으로 I구획의 인구 수가 크게 감소하였다. 과거의 경험에서 '사회적 거리두기'의 방역 효과는 이미 검증되었다.

SIR 모델로 예측한 향후 COVID-19 확산 추세

8 월 14 일 COVID-19 일일 확진자 수가 다시 100 명대로 진입한 이후 8 월 28 일까지 보름간 세 자리 수의 일일 확진자 수를 유지하면서 COVID-19 2 차 대유행 조짐이 나타나고 있으며 특히 8 월 27 일 일일 확진자는 441 명을 기록하면서 COVID-19 1 차 대유행 시점 이후 최대 규모의 일일 확진자가 발생했다. 심지어 기존의 코로나 바이러스에서 변이가 된 'GH'형으로 알려짐에 따라 1 차 대유행 시점에 시행된 '고강도 사회적 거리두기' 이상의 조치가 필요한 상황이다.¹⁴ 이에 현 상황을 반영해 8 월 28 일을 기점으로 세 가지 상황을 가정해 SIR 모델로 예측해 보고자 한다.¹⁵

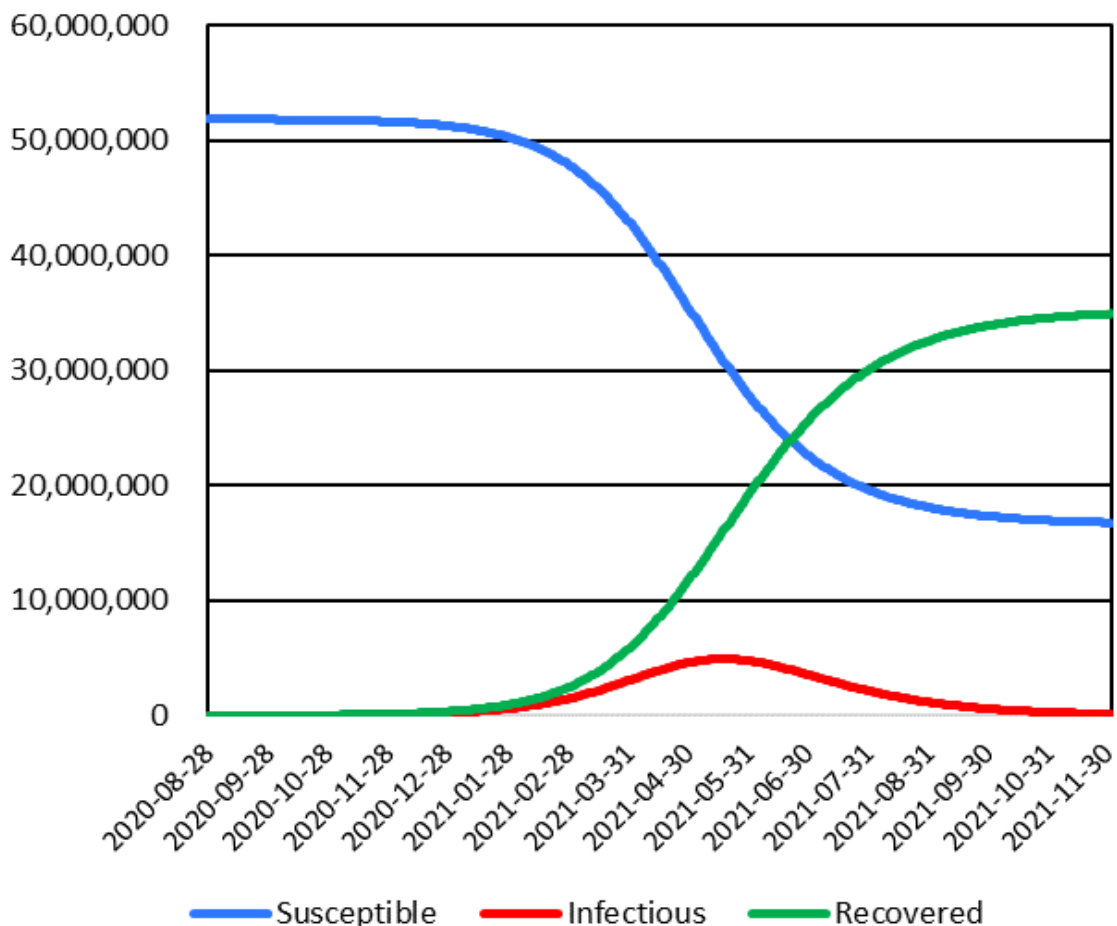
[그림 3] SIR 모델로 예측한 S, I, R 구획 인구 수 변화 ($\beta=0.2161$, $\gamma=0.05$, $R_0=4.32$, 단위: 명)



[그림 3]은 다시 세 자리 수 이상의 일일 확진자 수가 발생한 8 월 14 일부터 전국적으로 '사회적 거리두기' 2 단계를 시행하기 전날인 8 월 22 일까지의 증가추세가 이어진다는

가정하에 *SIR* 모델로 예측한 결과이다. 예측 결과 감염된 인구는 2020년 11월 초 최대치인 약 2,288만명을 기록한 후 줄어들어 2021년 11월 중순에 상황이 종료된다. 최종적으로 감염 가능한 인구 약 5,182만명 중 98.8%가 감염 후 회복되며 종료시점 감염되지 않은 인구는 1.2%에 해당되는 약 63만명에 불과하다. 8월 28일 기준 치명률은 1.66%인데 이를 확진된 인구에 적용해보면 최악의 경우 약 87만명에 이르는 사망자가 발생할 수 있다. 다소 과장된 결과로 보일 수 있지만 아무 조치가 없는 상황이 유지되는 경우 매우 심각한 상황에 직면할 수 있다. 이러한 최악의 상황이 실현될 경우 의료체계의 붕괴로 이어질 수 있기 때문에 사전에 강력한 방역조치를 취해야 한다.

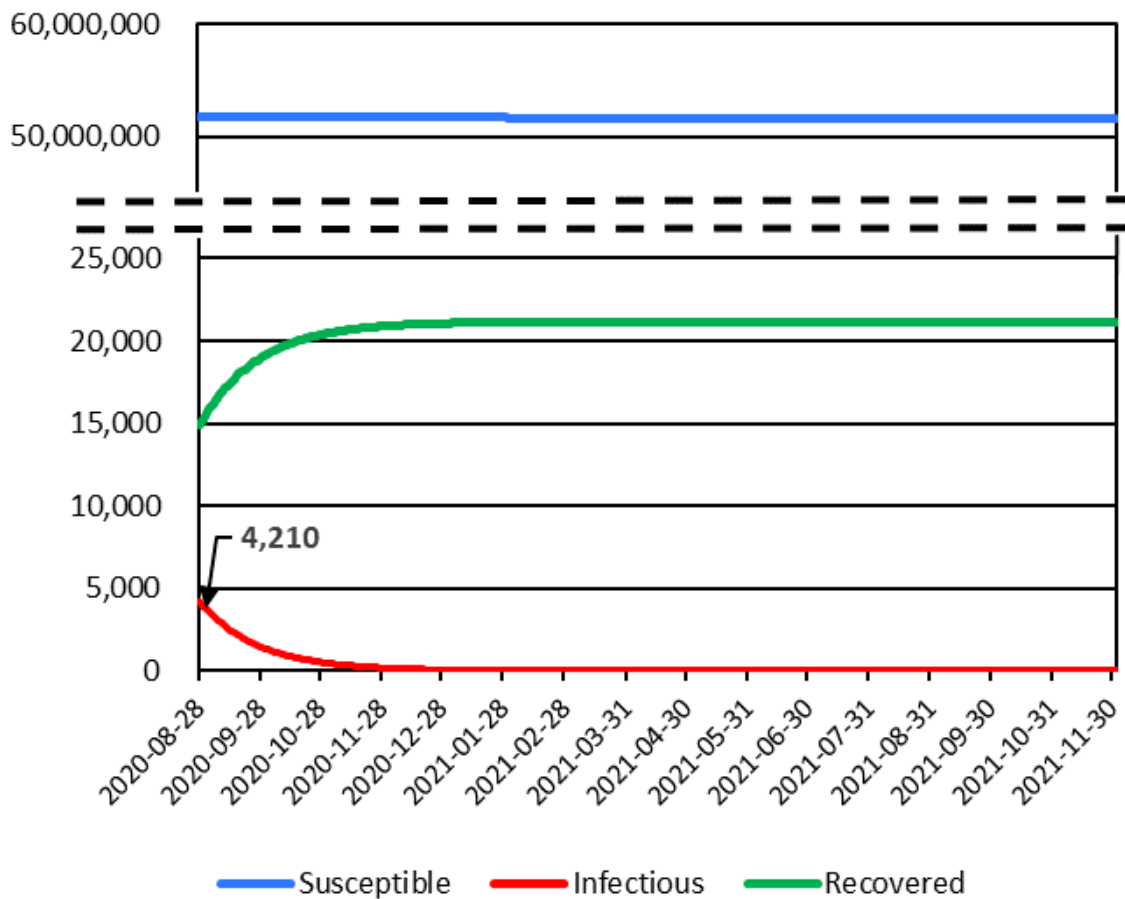
[그림 4] *SIR* 모델로 예측한 *S, I, R* 구획 인구 수 변화 ($\beta=0.0835, \gamma=0.05, R_0=1.67$, 단위: 명)



[그림 4]는 2020년 8월 23일 중앙방역대책본부 정례 브리핑에서 정은경 본부장이 질의응답에서 밝힌 전국 단위 R_0 값인 1.67을 적용해 *SIR* 모델로 예측한 결과이다. 예측 결과 감염된 인구는 2021년 5월 중순 최대치인 약 492만명을 기록한 후 줄어들어 2023년 5월

말에 상황이 종료된다. 최종적으로 감염 가능한 인구 약 5,182 만명 중 68.0%가 감염 후 회복되며 종료시점 32.0%에 해당하는 약 1,661 만명이 감염되지 않는다. 8 월 28 일 기준 치명률 1.66%으로 계산하면 약 58 만명에 이르는 사망자가 추가로 발생할 수 있다. 이러한 예측 결과는 앞의 [그림 3] 결과에 비해 확산이 많이 줄긴 했지만 여전히 의료체계가 감당할 수 없는 수준이다.

[그림 5] *SIR* 모델로 예측한 *S, I, R* 구획 인구 수 변화 ($\beta=0.0165, \gamma=0.05, R_0=0.33$, 단위: 명)



[그림 5]는 지난 2 월 COVID-19 1 차 대유행 시 시행한 45 일간의 '고강도 사회적 거리두기' 수준의 '사회적 거리두기'를 시행한다는 가정하에 *SIR* 모델로 예측한 결과이다. 예측 결과 감염된 인구는 시작점인 2020 년 8 월 28 일 최대치인 4,210 명을 기록하며 그 이후 줄어들어 2021 년 5 월 중순에 상황이 종료된다. 최종적으로 감염 가능한 인구 약 5,182 만명 중 0.004%인 2,089 명이 추가로 감염되며 8 월 28 일 기준 치명률 1.66%으로 계산하면 약 35 명에 이르는 사망자가 추가로 발생할 수 있다. 실제로 이러한 결과가 나타나게 된다면 앞선 [그림

3]의 결과와 비교했을 때 매우 성공적인 방역 결과로 이어질 수 있으며 만약 '고강도 사회적 거리두기' 수준 이상의 '사회적 거리두기'를 시행한다면 더욱 효과적인 방역이 될 수 있다.

정부의 선택은?

이번 이슈브리프에서는 COVID-19 2차 대유행 조짐이 나타나는 현 시점에서 향후 전개될 상황을 *SIR* 모델을 활용해 예상해보았다. 강력한 '사회적 거리두기'는 앞서 확인했듯이 현 시점에서 가장 효과적인 방역 조치이다. 유일한 방역조치로 표현해도 과언이 아닐 것이다. 이미 감염병 전문가 사이에서도 현재 대응 수준으로는 COVID-19 확산 추세를 꺾을 수 없으며 하루빨리 '사회적 거리두기' 3 단계로 격상해야 한다는 이야기가 나오고 있다.¹⁶ 다행히 정부는 8월 27일 중앙사고수습본부 정례 브리핑을 통해 상황의 엄중함을 인식하고 '사회적 거리두기' 3 단계 격상을 포함한 모든 가능성을 고려하고 있으며, 필요한 조치는 신속하고 과감하게 취할 것이라고 밝혔다.¹⁷ 강력한 방역 조치는 사회·경제에 큰 파장을 미치므로 매우 신중한 결정이 요구된다. 정부가 COVID-19 방역 최적의 타이밍을 놓치는 과오를 범하지 않길 바란다.

¹ 8월 28일 기준.

² 신천지 대구교회 신도.

³ <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC7270519/#CR5>

⁴ Kermack, W. O. and McKendrick, A. G., "A Contribution to the Mathematical Theory of Epidemics," Proceedings of the Royal Society, London, A115, 700-721, 1927.

⁵ SIR 모델은 뚜렷한 구획으로 나누어져 있어 구획모델(Compartmental model)로도 널리 알려져 있다.

⁶ β 는 한 사람이 주어진 시간(하루)에 만나는 평균 접촉자수와 감염자-비감염자 접촉 시 전염병이 전달될 확률(전염성)을 곱한 값이며, γ 는 회복률을 나타내며 $1/\gamma$ 는 평균 회복기간을 의미한다.

⁷ 여기서 R 은 Recovered 이외에 Removed 로도 표현된다.

⁸ 여기서 N 은 총 인구수이며, 전염병 확산과정의 시초부터 회복단계까지 걸리는 시간이 사람 평균수명과 비교해 볼 때 매우 짧아 출생, 사망, 이주로 인한 총 인구수의 변화에는 큰 영향이 없는 것으로 가정한다. 즉, SIR의 합 N 은 시간에 상관없이 항상 동일하다. 또한 위의 미분방정식 모두를 동시에 만족시키는 해는 수치적 방법을 통해 구하며 비례상수인 β, γ 그리고 S, I, N 은 0 이상의 값을 갖는다.

⁹ 한국의 경우 현재 검역을 통해 해외유입을 통제하고 있으며 전염병 확산과정의 시초부터 회복단계까지 걸리는 시간이 사람 평균수명과 비교해 볼 때 매우 짧아 출생, 사망, 이주로 인한 총 인구수 N 의 변화에는 큰 영향이 없는 것으로 가정한다. 또한 SIR 모델 적용 시 COVID-19의 재감염 가능성은 없다고 가정하였다.

¹⁰ $S_0 > S_t, t > 0$.

¹¹ R_0 값은 시간에 따라 변하기 때문에 정기적으로 업데이트된 R_t 값을 구해 정확도를 높일 수 있다.

¹² 2020년 1월 기준 인구 사용, KOSIS

¹³ SIR 모델 예측 시 γ 값은 다수의 매체에서 보도된 코로나 회복기간인 평균 20일을 적용해 0.05로 정하였으며 (실제 자료로 추정된 γ 값으로 구한 회복기간은 20일보다 더 오래 걸림) 31번 확진자(신천지 대구교회 신도)가 발생한 시점인 2월 18일부터 '고강도 사회적 거리두기' 시행 직전인 3월 21일까지의 자료를 사용해 β 값을 구하였다. 이때 얻어진 β 값은 0.2184인데 '고강도 사회적 거리두기'가 시행된 3월 22일부터 5월 5일간의 실제 자료로 구한 β 값은 0.0165로 '고강도 사회적 거리두기'를 통해 약 1/13만큼 줄어든 것을 확인할 수 있다. R_0 값 역시 4.36에서 0.33으로 개선되었다.

¹⁴ https://imnews.imbc.com/replay/2020/nwdesk/article/5885206_32524.html.

¹⁵ 2020년 7월 기준 인구 사용, KOSIS

¹⁶ <https://www.yna.co.kr/view/AKR20200827072900530?input=1195m>.

¹⁷ <https://www.yna.co.kr/view/AKR20200827092651530?input=1195m>.