A Markov Model to Assess Covid-19 Vaccine Herd Immunization Patterns

Jorge Luis Romeu, Ph.D.
https://www.researchgate.net/profile/Jorge_Romeu
http://web.cortland.edu/romeu/; Email: romeu@cortland.edu

1.0 Introduction

This Markov model *assesses different patterns of vaccination*, which will affect achieving *Herd Immunity*. Such immunization approach assumes the virus will affect a large percentage of the general population, starving it for new *customers* and thus preventing community spread.

Herd Immunization can be acquired *by two* methods: (a) letting the *virus infect most of the population*. Weaker members (the elderly, those with co-morbidities, etc.) will die, while those surviving will become immunized. Alternatively, *Herd Immunization* can be acquired by (b) *vaccinating a large percent of the general population*. Such is achieved by inoculating a small, controlled virus infection. In either case, a large percent of the population must be involved.

The *urgency of the present situation* is that many people still refuse to take the vaccine. It is imperative to address the need for the population to agree. For this reason this paper will be very short and straight to the point. *For Markov Chain background material* and discussions, readers are sent to *our previous papers*: *A Markov Chain Model for Covid-19 Survival Analysis*: https://www.researchgate.net/publication/343021113_A_Markov_Chain_Model_for_Covid-19_Survival_Analysis and *A Markov Model to Study Covid-19 Herd Immunization,* found in: https://www.researchgate.net/publication/343345908_A_Markov_Model_to_Study_Covid-19_Herd_Immunization?channel=doi&linkId=5f244905458515b729f78487&showFulltext=true

We will *assume* that, either *by infection or by vaccination*, *individuals* will become *immunized* and will, therefore, *not become sick again nor pass down the infection* to anyone else. We will consider *two levels of vaccination*: *acceptable* (70% of the *available*[1] population) *and low* (40%). All *other model parameters* including vaccine effectiveness (70%) remain the *same*. We *develop Markov models* for both situations, *obtain and compare* their respective *performance measures*, and discuss the results. We conclude with some general remarks about vaccination policies.

2.0 Previous Covid-19 work

Following our proposal to the retired Academic and Research Communities, to fight Covid-19: https://www.researchgate.net/publication/341282217_A_Proposal_for_Fighting_Covid-19_and_its_Economic_Fallout we have written a number of articles on statistical models:

[1] Total population, minus those individuals already infected and thence naturally immunized, etc.

https://www.researchgate.net/publication/346956247_Logistic_Regression_in_Factor_Identification_of_Covid-19_Vaccine_Clinical_Trials on identifying important key factors observed in Clinical Trials, https://www.researchgate.net/publication/346305686_A_Digression_on_Covid-19_Vaccine_Clinical_Trials_and_its_Consequences and ICUs and hospital staffing using the Negative Binomial distribution: https://www.researchgate.net/publication/345914205_Covid-19_ICU_Staff_and_Equipment_Requirements_using_the_Negative_Binomial screening DOEs: https://www.researchgate.net/publication/344924536_Design_of_Experiments_DOE_in_Covid-19_Factor_Screening_and_Assessment using statistical methods to establish a new *Vaccine Life*: https://www.researchgate.net/publication/344495955_Survival_Analysis_Methods_Applied_to_Establishing_Covid-19_Vaccine_Life as well as to help *accelerate vaccine testing*: https://www.researchgate.net/publication/344193195_Some_Statistical_Methods_to_Accelerate_Covid-19_Vaccine_Testing and a *Markov model to study* problems of *reopening college*: https://www.researchgate.net/publication/343825461_A_Markov_Model_to_Study_College_Re-opening_Under_Covid-19 and *Markov Model to study* the *effects of Herd Immunization*: https://www.researchgate.net/publication/343345908_A_Markov_Model_to_Study_Covid-19_Herd_Immunization?channel=doi&linkId=5f244905458515b729f78487&showFulltext=true as well as of *general survival:* https://www.researchgate.net/publication/343021113_A_Markov_Chain_Model_for_Covid-19_Survival_Analysis about socio-economic and racial issues affected by Covid-19: https://www.researchgate.net/publication/343700072_A_Digression_About_Race_Ethnicity_Class_and_Covid-19 and developing *A Markov Chain Model for Covid-19 Survival Analysis:* https://www.researchgate.net/publication/343021113_A_Markov_Chain_Model_for_Covid-19_Survival_Analysis and *An Example of Survival Analysis Applied to analyzing Covid-19 Data*: https://www.researchgate.net/publication/342583500_An_Example_of_Survival_Analysis_Data_Applied_to_Covid-19, and *Multivariate Statistics in the Analysis of Covid-19 Data*, and *More on Applying Multivariate Statistics to Covid-19 Data*, both of which can also be found in: https://www.researchgate.net/publication/341385856_Multivariate_Stats_PC_Discrimination_in_the_Analysis_of_Covid-19, and the implementation of *multivariate analyses* methods such as: https://www.researchgate.net/publication/342154667_More_on_Applying_Principal_Components_Discrimination_Analysis_to_Covid-19 *Design of Experiments* to the Assessment of Covid-19: https://www.researchgate.net/publication/341532612_Example_of_a_DOE_Application_to_Coronavarius_Data_Analysis *Offshoring*: https://www.researchgate.net/publication/341685776_Off-Shoring_Taxpayers_and_the_Coronavarus_Pandemic and *reliability methods* in ICU assessment: https://www.researchgate.net/publication/342449617_Example_of_the_Design_and_Operation_of_an_ICU_using_Reliability_Principles and *Quality Control* methods *for monitoring* Covid-19: https://web.cortland.edu/matresearch/AplicatSPCtoCovid19MFE2020.pdf *Numerical Example* https://www.researchgate.net/publication/339936386_A_simple_numerical_example_that_illustrates_the_dangesrs_of_the_Coronavarus_epidemic

3.0 A Markov Chain Model with Two Absorbing States.

We analyze *two Covid-19 vaccination patterns* using a Markov Chain. The Markov Chain model is defined over a *nine-element state space*, which is flexible and allows implementing multiple options. Such Markov Chain models a *non-recurrent process* that, moving through a number of *Transient States*, eventually *ends in one of its two Absorbing States: Death or Immunization*, where they will remain forever.

Because the *state space* contains both transient and recurring (absorbing) states there is no steady state system solution. Instead, we pursue the *long run (asymptotic) probabilities of either dying or becoming immunized*, when starting in any of its Transient States.

Let $\{X_n; n >= 0\}$ be a *Markov Chain* over a *nine-element State Space* defined as:

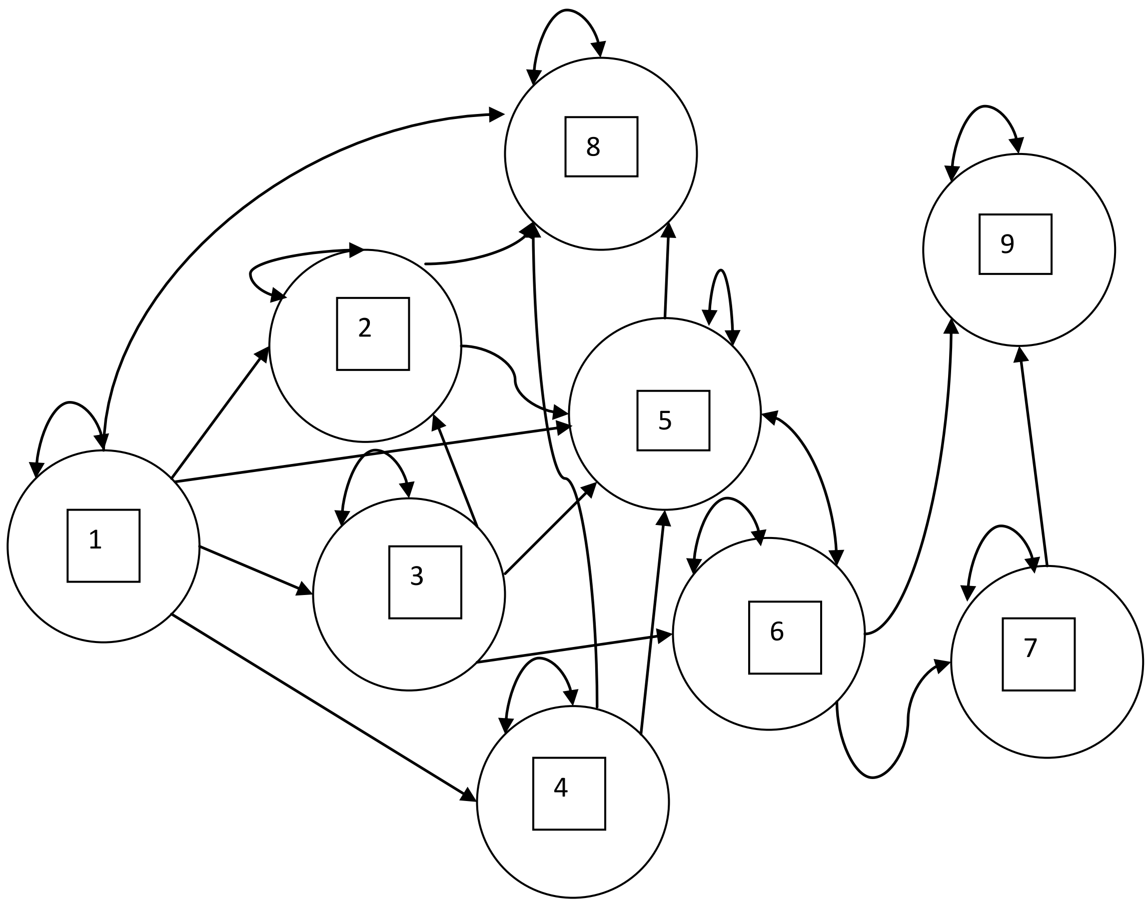
(1) *Non Infected* persons in the General Population;
(2) *Mildly Infected & Isolated* (discomfort; but no medical attention required);
(3) *Non-Vaccinated* (those that rejected vaccine)
(4) *Vaccinated persons* (those that agreed to take the vaccine)
(5) *Severely Infected, Detected & Quarantined* (treated at home)
(6) *Hospitalized patients* (after becoming severely ill with Colvid-19);
(7) *Patients in the ICU* (extremely sick);
(8) *Immunized* (an absorbing state);
(9) *Death* (an absorbing state)

We run the Markov Chain model twice. First, we develop *an optimistic framework*, assuming that *70% of the available population* will become vaccinated. Second, we develop *a pessimistic framework*, assuming that *only 40%* of the *available population* will become *vaccinated*. All other model parameters (i.e. all other state transition rates) remain the same. This will allow for a fair comparison of both vaccination *frameworks* under consideration, and their consequences.

The *Transition Probability Matrix P* for the First (70% vaccination) *Markov Chain* is:

Pop	**MildInf**	**NonVac**	**Vaccine**	**SevereCov**	**Hospital**	**ICU**	**Immune**	**Dead.**
0.05	**0.05**	**0.23**	**0.54**	**0.03**	**0.00**	**0.00**	**0.10**	**0.00**
0.00	0.70	0.00	0.00	0.05	0.00	0.00	0.25	0.00
0.00	0.10	0.70	0.00	0.10	0.10	0.00	0.00	0.00
0.00	0.20	0.00	0.07	0.03	0.00	0.00	0.70	0.00
0.00	0.00	0.00	0.00	0.30	0.20	0.00	0.50	0.00
0.00	0.00	0.00	0.00	0.20	0.57	0.20	0.00	0.03
0.00	0.00	0.00	0.00	0.00	0.20	0.65	0.00	0.15
0.00	0.00	0.00	0.00	0.00	0.00	0.00	1.00	0.00
0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	1.00

Figure 1: The *State Space Diagram* for the Markov Chain



RATIONALIZATION

The Markov Chain *unit time* is *a week*. *Transitions* refer to the State changes that occur from *one Monday to the following Monday*. State (1): the *General population*. Asymptomatically infected individuals, unaware of their condition, pass the virus on to others and then become immunized. State (2): *Mild Covid-19*; detected persons are quarantined at home. State (3): *Non Vaccinated* persons (refused vaccination). Some may spread the virus in the community. Some may develop symptoms, are tested, and are quarantined at home. If they become very ill, they are sent to the hospital. State (4): *Vaccinated*; some individuals may suffer mild Covid-19, stay at home, and eventually 70% become immunized. Some others may get sick and require medical attention. State (5): *infected* and requiring *medical attention at home*. State (6): persons with *serious illness* which require *hospitalization and treatment*. Some recover and are sent home; others get worse, and are placed in ICUs. State (7): *patients in ICU*. Some improve and are sent back to the ward, and others are placed on Ventilators or die. State (8): *herd immunization*; and *State (9): death.* These two latter states are *Absorbing: once you enter them, you cannot leave.*

We obtain, from the TPM, Sub-Matrix Q, corresponding to the seven Transient states:

Matrix Q of the Seven Transient States:

Pop	MildInf	NonVac	Vaccine	SevereCov	Hospital	ICU
0.05	0.05	0.23	0.54	0.03	0.00	0.00
0.00	0.70	0.00	0.00	0.05	0.00	0.00
0.00	0.10	0.70	0.00	0.10	0.10	0.00
0.00	0.20	0.00	0.07	0.03	0.00	0.00
0.00	0.00	0.00	0.00	0.30	0.20	0.00
0.00	0.00	0.00	0.00	0.20	0.57	0.20
0.00	0.00	0.00	0.00	0.00	0.20	0.65

We then subtract matrix Q from the Identity Matrix, yielding (I-Q),

0.95	-0.05	-0.23	-0.54	-0.03	0.00	0.00
0.00	0.30	0.00	0.00	-0.05	0.00	0.00
0.00	-0.10	0.30	0.00	-0.10	-0.10	0.00
0.00	-0.20	0.00	0.93	-0.03	0.00	0.00
0.00	0.00	0.00	0.00	0.70	-0.20	0.00
0.00	0.00	0.00	0.00	-0.20	0.43	-0.20
0.00	0.00	0.00	0.00	0.00	-0.20	0.35

We invert the above: **Matrix inverse S = $(I-Q)^{-1}$**

Populat	MildInf	NonVac	Vaccine	SevCov	Hospital	ICU.
1.05263	0.85191	0.80702	0.61121	0.39130	0.50350	0.28772
0.00000	3.33333	0.00000	0.00000	0.29071	0.18416	0.10524
0.00000	1.11111	3.33333	0.00000	1.04665	1.71885	0.98220
0.00000	0.71685	0.00000	1.07527	0.11879	0.07525	0.04300
0.00000	0.00000	0.00000	0.00000	1.74428	1.10497	0.63141
0.00000	0.00000	0.00000	0.00000	1.10497	3.86740	2.20994
0.00000	0.00000	0.00000	0.00000	0.63141	2.20994	4.11997

The Potential Matrix, of which S (above) is a sub-matrix, *contains the Long-Run Sojourns*, or the average number of visits *to each of the States* (matrix columns), when starting from any other of the States (represented by the matrix rows). Sojourns that occur outside matrix S are either Zero (if the states cannot be reached) or Infinite (if the transition is into an Absorbing state).

We then obtain *Matrix G = SB* where S is as above, and B is the 7x2 matrix formed with the last two columns of the initial TPM, corresponding to the two Absorbing states. Such matrix G then provides the probabilities of ever reaching said two absorbing states (immunization or death), when starting from any of the seven transient states (i=1, … 7) below, denoted by:

Populat(1); MildInf(2) ;NonVaccin(3) ;Vaccin(4) ;SevereCovid(5) ;Hospitalized(6) ;ICU(7)

St	Inmunizat	Deaths.
1.	0.941738	0.058262
2.	0.978690	0.021310
3.	0.801105	0.198895
4.	0.991293	0.008707
5.	0.872139	0.127861
6.	0.552486	0.447514
7.	0.315706	0.684294

We will *interpret this table* of results *through a numerical example*. Assume there is a population of *one million adults*, aged 18 and above (which is the age for taking the Covid vaccine). From Line 1 (i.e. *Population*) we have that *94.2% of these individuals* (or 942,000) will be *immunized*, while *5.8% of them* (or 58,200) *will die*. If we look at Line 2 (i.e. *Mildly Infected* and naturally *immunized*) we see that, from those individuals, *97.8% will survive* and *2.1% will die*. If we look at Line 3 (i.e. *Non Vaccinated*[2] sub-population, because they *rejected the vaccine*), 80.1% will be immunized and *19.8% will die*. Regarding Line 4 (i.e. *Vaccinated*[3] sub-population), *99% will survive* immunized, and 0.87% (less than 1%) will die of Covid-19. For Line 7 (i.e. those *in an ICU*), *31.5% will survive* and become immunized, while the remaining 68.4% will die.

Beware that, *except Line 1* (which is the only one that concerns the total population), each of the *percents of other Lines apply only to the sub-totals* of such sub-population. *For example*, for the Non-Vaccinated sub-population of 0.23*1,000,000 = 231,000, 19.8% will die (231000*0.198 = 45,738 persons) and 231,000-45,738 = 185,262 will survive and become naturally immunized.

Follow exactly the same calculations above, for the *pessimistic framework of 40% vaccinated*:

The *Transition Probability Matrix P* for the Second (40% vaccination) *Markov Chain* is:

Pop	MildInf	NonVac	Vaccine	SevereCov	Hospital	ICU	Immune	Dead.
0.05	**0.05**	**0.46**	**0.31**	**0.03**	**0.00**	**0.00**	**0.10**	**0.00**
0.00	0.70	0.00	0.00	0.05	0.00	0.00	0.25	0.00
0.00	0.10	0.70	0.00	0.10	0.10	0.00	0.00	0.00
0.00	0.20	0.00	0.07	0.03	0.00	0.00	0.70	0.00
0.00	0.00	0.00	0.00	0.30	0.20	0.00	0.50	0.00
0.00	0.00	0.00	0.00	0.20	0.57	0.20	0.00	0.03
0.00	0.00	0.00	0.00	0.00	0.20	0.65	0.00	0.15
0.00	0.00	0.00	0.00	0.00	0.00	0.00	1.00	0.00
0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	1.00

After implementing the same algorithmic process developed above, we obtain:

[2] Non-Vaccinated transition rate is: 0.3*(1-0.1-0.05-0.05-0.03) = 0,3*0.77 = 0.231
[3] Vaccinated transition rate is: 0.7*(1-0.1-0.05-0.05-0.03) = 0.3*0.77 = 0.539

Again the Matrix G = SB, where S = $(I - Q)^{-1}$, and B is the 7x2 matrix formed by the last two columns of the initial TPM corresponding to the two Absorbing states. Matrix G provides the probabilities of ever reaching said two absorbing states (immunization or death), when we start from any of the seven transient states (i=1, … 7) below, denoted by:

Populat(1); MildInf(2) ;NonVaccin(3) ;Vaccin(4) ;SevereCovid(5) ;Hospitalized(6) ;ICU(7)

St	Inmunizat	Deaths.
1.	0.900908	0.099092
2.	0.978690	0.021310
3.	0.801105	0.198895
4.	0.991293	0.008707
5.	0.872139	0.127861
6.	0.552486	0.447514
7.	0.315706	0.684294

Again we interpret this table of results through a *numerical example*. Assume a population of one million adults, aged 18 and above (which is the age for taking Covid19-19 vaccine). From Line 1 (i.e. *Population*) we have that *90.2%* of these individuals (or 902,000) will be *immunized,* while *9.9%* of them (or 99,000) *will die*. The results from the rest of the Lines are the same as in the previous example. How, then, do we explain a larger number of deaths in the general population?

The answer is straight forward. *Now*, *46% of the total* population of one million (i.e. 460,000 individuals) *are Non-Vaccinated*. Thence, *19.9% of them (i.e. 91,080) will die*, instead of only 45,738 persons, in the previous example, and *mostly from those who refused* to take the vaccine.

4.0 Discussion

We have *implemented a Markov Chain twice*, *changing only the parameter*s that represented the *vaccination pattern*. The model *results* are unquestionably *different*. *We have used this technique before*, applying a *Linear Programming (LP) model* to the US manufacturing off-shoring process during the last quarter of a century. We considered, in our LP model, activities such as providing health care, retraining, unemployment benefits, etc. to off-shored-displaced workers, which were not included (or paid for) in off-shoring companies' LPs. They were paid by the US government (tax payers). *Both LP model results* https://www.researchgate.net/publication/341685776_Off-Shoring_Taxpayers_and_the_Coronavarus_Pandemic) were *also* unquestionably *different*.

In either case, the *reader may challenge* our choice of *parameter values*. What is indisputable is how, in both cases, there is *an obvious difference in model results*, independently of the value of model parameters used. Such *difference stems from the different patterns of behavior* modeled.

Then, vaccination activity carries two aspects: one *individual and* the other*, social*. The vaccine *protects the individual*. In addition, if enough individuals in the general population take it, the vaccine *has an effect over the Pandemic*. With few new customers to infect, *the virus starves* and ends up disappearing! If not enough people become vaccinated this latter, *greater benefit* is lost.

Markov models, due to their State Spaces (E_j) and transition probabilities (X_{ij} for i,j=0, … ,k), are particularly *useful to model Immunization patterns*. For example, State (1) is *General Population* and State (8), *Immunized Population*. Then, $X_{18} = 0.1$ corresponds to the 10% of asymptomatic persons that become immunized; $X_{12} = 0.05$ corresponds to the 5% persons with mild infections that are detected, quarantined, do not require medical attention, and finally become immunized. Also, $X_{15} = 0.03$ represents the 3% of the infections that require treatment at home. If infections are positively resolved, patients become immunized; if not, patients are hospitalized.

There are 70% (or 40%) *persons vaccinated* (resulting in $X_{14} = 0.54$ or 0.31 once persons already naturally immunized, sick, younger than 17 years, etc. are subtracted). And there are *23%* (or 46%) persons *Non-Vaccinated* (resulting in $X_{13} = 0.46$ or 0.23, just as above). These *rates can be changed*, and model results can then be assessed and compared, *to find efficient strategies*.

One important parameter is $X_{48} = 0.7$ (*transition from Vaccinated to Immunized*). We assumed that the *vaccine was 70% efficient* (accounting for all different efficiency rates for diverse ages, co-morbidities, etc.). *Other values can be used* and model results can be assessed and compared. The same can be said about transition rates related to hospitalization, ICU etc. that depend on the load such hospitals and health care practitioners have, which in turn depend on infection rates.

The *time unit* utilized *for transitions* was *a week* (other units could have been used: hours, days, months, etc.). For example, the *weekly rate of asymptomatic infection* ($X_{18} = 0.1$), for members of the general population who contract Covid-19 and then become immunized, means it is 10%.

Probabilities of persons dying or immunized, combined with population estimates, helps assess the number of cases, deaths, immunized persons, etc. that result from the different scenarios.

Results from the present Markov model, complement results obtained in our previous papers, as different models address the different aspects of Covid-19 problems. For example, the article https://www.researchgate.net/publication/342583500_An_Example_of_Survival_Analysis_Data_Applied_to_Covid-19 can be *used in Triage situations, to evaluate the probability of a patient dying*, given its co-morbidities, and allocate scarce ICU beds or doctors to those patients with a higher chance of making the best possible use of these resources.

One advantage of *achieving a Vaccine Herd Immunization* is the possibility of *avoiding* the bad *triage situation* mentioned above. If the healthcare systems overheat, due to the large number of Covid-19 cases, Triage becomes a reality.

5.0 Conclusions

Again, *the urgency of this paper* stems from recent polls results that suggest that *a large number of people are not yet willing to take the vaccine*. And this is *a very serious problem*. Some argue that *there are still many unanswered questions* regarding the vaccine: How long does immunity last? Does it prevent passing on the virus? What are its secondary effects? *However, there is a key question*: which is the greater risk, taking the vaccine and becoming protected, with all the mentioned caveats, or contracting Covid-19, passing it on to those we love, and perhaps dying.

Vaccination activity carries two aspects: *one individual and the other social*. First, the *vaccine protects the individual*. *In addition*, if enough individuals in the general population become vaccinated, *it has an effect over the Pandemic*. With few new customers to infect, *the Covid-19 virus starves,* and ends by disappearing, which is the only real solution to the Covid-19 problem. *If not enough people become vaccinated, this latter, larger benefit is lost*.

Our *Markov model*, as well as its transition rates and its states, were all *based on our readings about Covid-19 and our 45+ years of statistical experience* in modeling and data analysis. This author is not a public health specialist, but a statistician. *Our intent is to place these models in the hands of health specialists*, so they can implement them with better information. *Models* may *provide* more *neutral rules* to establish *medical procedures*, if such is ever necessary.

Finally, *our statistical models* do not intend to compete with, but to *complement and enrich*, the *models developed by Epidemiologists* and other Public Health professionals. One example[4] says:

The total population is partitioned into eight stages of disease: S, susceptible (uninfected); I, infected (asymptomatic or pauci-symptomatic infected, undetected); D, diagnosed (asymptomatic infected, detected); A, ailing (symptomatic infected, undetected); R, recognized (symptomatic infected, detected); T, threatened (infected with life-threatening symptoms, detected); H, healed (recovered); E, extinct (dead). The interactions among these stages are shown in Fig. 1. We omit the probability rate of becoming susceptible again after having recovered from the infection. Although anecdotal cases are found in the literature, re-infection rate value appears negligible.

The reader will recognize how their eight stages of the disease are somewhat similar to our Nine State Space Elements. Also, how said article's *Figure 1* is also somewhat similar to our own *State Space Diagram*, presented in Section 3.0. Modeling tools used are different.

Both models are fine and pursue a similar objective: to help study and overcome Covid-19. Researchers are allies, not competitors.

As statistician George Box once said: *all models are wrong; some models are useful*.

[4] See: https://www.nature.com/articles/s41591-020-0883-7#Sec2

Bibliography

Taylor, H. and S. Karlin. An Introduction to Stochastic Modeling. Academic Press. NY. 1993.

Cinlar, E. Introduction to Stochastic Processes. Prentice Hall. NJ. 1975.

Heyman, D. and M. Sobel. Handbooks in Operations Research and Management Science. *Vol. 2: Stochastic Models*. North Holland. Amsterdam. 1990.

Box, G., Hunter, W. G., and J. S. Hunter. Statistics for Experimenters.Wiley. New York.1978.

Walpole, R. E. and R. H. Myers. Probability and Statistics for Engineers and Scientists. Prentice-Hall. http://www.elcom-hu.com/Mshtrk/Statstics/9th%20txt%20book.pdf

Romeu, J. L. *Operations Research and Statistics Techniques*. Proceedings of Federal Conference on Statistical Methodology. https://web.cortland.edu/matresearch/OR&StatsFCSMPaper.pdf

About the Author:

Jorge Luis Romeu retired Emeritus from the State University of New York (SUNY). He was, for sixteen years, a Research Professor at Syracuse University, where he is currently an Adjunct Professor of Statistics. Romeu worked for many years as a Senior Research Engineer with the Reliability Analysis Center (RAC), an Air Force Information and Analysis Center operated by IIT Research Institute (IITRI). Romeu received seven Fulbright assignments: in Mexico (3), the Dominican Republic (2), Ecuador, and Colombia. He holds a doctorate in Statistics/O.R., is a C. Stat. Fellow, of the Royal Statistical Society, a Senior Member of the American Society for Quality (ASQ), and Member of the American Statistical Association. He is a Past ASQ Regional Director (and currently a Deputy Regional Director), and holds Reliability and Quality ASQ Professional Certifications. Romeu created and directs the Juarez Lincoln Marti International Ed. Project (JLM, https://web.cortland.edu/matresearch/), which supports (i) higher education in Ibero-America and (ii) maintains the Quality, Reliability and Continuous Improvement Institute (QR&CII, https://web.cortland.edu/romeu/QR&CII.htm) applied statistics web site.