

# A Year of Work in Covid-19 Statistical Modeling

Jorge Luis Romeu, Ph.D.

Emeritus SUNY; Adjunct Prof., Dpt. MAE, 263 Link Hall, Syracuse University, NY 13224

## Abstract

We overview a year of statistical work developing tutorials, models and analysis applications to Covid-19 issues to illustrate our methods and persuade researchers from other professions to include statistics and statisticians in their work. Section 2 describes the web page created, and its background material. Section 3 overviews Quality Control and Design of Experiments papers that, using Covid-19 data, assess whether state counties or regions should be opened or locked-down. Section 4 overviews Reliability and Logistics papers that, using patient data, help take Triage decisions. Section 5 summarizes Principal Components and Discriminant Analysis papers that, using patient data, assess the results of Covid-19 vaccine and treatment clinical trials. Section 6 overviews Stochastic Process papers that implement Markov Chain models to evaluate decisions on opening colleges and assess Covid-19 hospital and ICU needs to avoid being overwhelmed. Section 7 presents Socio-economic analyses of problems derived from the Covid-19 pandemic. Section 8 presents Quality Engineering approaches to the Design and Improvement of Covid-19 vaccine distribution systems. Section 9 summarizes our work.

**Keywords:** public health, DOE, multivariate, discrimination, survival analysis, Markov chains.

## 1. Introduction

In this article we summarize ten months of pro-bono work on statistical modeling of Covid-19 topics. It all started on March 4th, when we wrote a communication to the Grand Lodge of New York, raising our concerns about continuing Lodge meetings, given the speed with which the Coronavirus was spreading from China, to Europe, to New York, California, Washington, etc.

We stated how, in Washington State, 10 deaths were attributed to the virus. Assuming that if the death rate was in fact 2%, as in China, there could be up to  $10/0.02 = 500+$  infections. Authorities had identified only 100+. Thence, there existed the possibility of having 400+ cases, still remaining undetected, potentially infecting others. In addition, we provided a numerical example describing how the infection process worked, that made its way to the media and became a video tutorial.

Shortly after, we proposed to the American Statistical Association and American Society for Quality<sup>1</sup> to organize a group of retired members to volunteer with Covid-19 statistical efforts. In April, we wrote the proposal <https://web.cortland.edu/matresearch/Covid-19Proposal2020.pdf> and started distributing and hanging our papers in ResearchGate, LinkedIn and other Internet forums.

We started developing statistical models and applications of methods that could illustrate statistics usage, persuade researchers and practitioners from other professions to include statistics in their work, as well as to invite more statisticians to collaborate with them. These tutorials would also help educate political

---

<sup>1</sup> GLNY, the ASA, and ASQ are three Civil Society organizations we belong to, and work with.

leaders (<https://web.cortland.edu/matresearch/CORONAVARUSListEngNov2020.pdf>) as well as the general public. The research material was well-received, according to its LinkedIn and ResearchGate hits: <https://web.cortland.edu/matresearch/SELECTEDREADINGSRESEARCHGATE.pdf>

In the rest of this paper we summarize the contents and uses of the statistics papers written. In Section 2 we describe the web page we created, and its introductory section. In Section 3 we deal with papers on Design of Experiments and Quality Control Applications. In Section 4, we deal with Reliability and Logistics Applications. In Section 5 we summarize the papers using Multivariate Analysis (Principal Components and Discriminant). In Section 6 we overview Stochastic Process papers that implement Markov Chain models. In Section 7, we present Socio-economic analyses of some problems derived from dealing with Covid-19. In Section 8, we conclude.

## **2.0 History and basic information:**

In this section we present some Basic Information regarding the Study of the Coronavirus Pandemic and its outbreak, from both historical and sociological, and numerical and medical perspectives.

First, this Pandemic has been ill-named the *Chinese or Kung-Fu virus*, just as the 1918 Flu Pandemic: <https://web.cortland.edu/matresearch/HISTORY%201918%20FLU%20PANDEMIC.pdf> was named the *Spanish Flu*, and perhaps with a similar objective.

The *Economist* article: <https://web.cortland.edu/matresearch/CoronavirusEconomist.pdf> appearing at the beginning of the Pandemic, analyzes its origin and implications.

We developed an illustrative Numerical Example of the transmission process of the virus:

[https://www.researchgate.net/publication/339936386\\_A\\_simple\\_numerical\\_example\\_that\\_illustrates\\_the\\_dangers\\_of\\_the\\_Coronavirus\\_epidemic](https://www.researchgate.net/publication/339936386_A_simple_numerical_example_that_illustrates_the_dangers_of_the_Coronavirus_epidemic)

Said numerical example made its way to the Syracuse newspapers, which adapted it into a U-Tube video for educational use: <https://www.syracuse.com/coronavirus/2020/03/how-fast-can-coronavirus-spread-statistics-professor-explains-why-we-need-to-act-now-video.html>

We wrote a formal Proposal, asking the retired community of statisticians, academics and researchers, to contribute their time and experience, pro-bono, to help fight the Covid-19 Pandemic. It is found in: [https://www.researchgate.net/publication/341282217\\_A\\_Proposal\\_for\\_Fighting\\_Covid-19\\_and\\_its\\_Economic\\_Fallout](https://www.researchgate.net/publication/341282217_A_Proposal_for_Fighting_Covid-19_and_its_Economic_Fallout)

The Johns Hopkins University global map, illustrating the world-wide expansion of the Coronavirus <https://gisanddata.maps.arcgis.com/apps/opsdashboard/index.html#/bda7594740fd40299423467b48e9ecf6> was included in the web page, for general information.

Similar Information ,about the spread of Coronavirus infection in the State of New York, is found in: <https://www.syracuse.com/coronavirus-ny/>

In Syracuse University's MFE634, a graduate Quality Engineering course, students applied Quality tools to analyze several Covid-19 issues: <https://web.cortland.edu/matresearch/Covid-19COPQGrp1.pdf>, and

[https://web.cortland.edu/matresearch/Coronavirus\\_Assess\\_Grp2.pdf](https://web.cortland.edu/matresearch/Coronavirus_Assess_Grp2.pdf),  
<https://web.cortland.edu/matresearch/Covid-19FTAgrp5.pdf>

Prior MFE634 Quality Engineering students also had Experiences with Epidemiology (Ebola & Zika):  
<https://web.cortland.edu/matresearch/EbolaGageR&R2016.pdf>  
<https://web.cortland.edu/matresearch/2017ZikaVirusFinPres.pdf>

The Section ends with a list of our Covid-19 papers, including *hits* from LinkedIn and ResearchGate:  
<https://web.cortland.edu/matresearch/SELECTEDREADINGSRESEARCHGATE.pdf> that provide some indication about how often and how much these papers were read, and about the readers' institutions.

### **3.0 Design of Experiments (DOE) and Quality Control (SPC) Applications:**

In this section we present tutorials on specialized statistical methods such as DOE and SPC that can be successfully applied to the Study of the Coronavirus Pandemic issues. We give the title of the papers, the link where it can be found, and a short description of its contents:

*Monitoring Community Infection Levels of Covid-19 Virus using Quality Control Techniques*

<https://web.cortland.edu/matresearch/AplicatSPCtoCovid19MFE2020.pdf>

This short paper illustrates some practical aspects in the implementation of *Control Charts* for proportions and mean/range, *to assess and monitor* community infection levels of Covid-19.

*Design of Experiments in Identification of Factors impacting Community Spread of Covid-19*

[https://www.researchgate.net/publication/341532612\\_Example\\_of\\_a\\_DOE\\_Application\\_to\\_Coronavirus\\_Data\\_Analysis](https://www.researchgate.net/publication/341532612_Example_of_a_DOE_Application_to_Coronavirus_Data_Analysis)

This paper walks the reader through the *detailed implementation of a DOE*, *to assess* the Effect on *Community Infection Spread*, of Factors Social Distancing, Face Masks and Day Schedule. The data set used was created by the author, using his experience and readings of Covid-19.

*Fractional Factorial DOEs in Identification of medical treatments that reduce Covid19 infection*

[https://www.researchgate.net/publication/344924536\\_Design\\_of\\_Experiments\\_DOE\\_in\\_Covid-19\\_Factor\\_Screening\\_and\\_Assessment](https://www.researchgate.net/publication/344924536_Design_of_Experiments_DOE_in_Covid-19_Factor_Screening_and_Assessment)

This paper is *a tutorial on Fractional Factorial DOEs* used *to screen and identify* experimental treatments and characteristics that may affect Covid-19 patients. Designs used include *Plackett-Burnam*, which helps *to reduce the number of experiments* done, saving valuable researcher time and effort. The Covid-19 data was also created by the author, using his experience and readings.

### **4.0 Reliability and Logistics Applications:**

In this section we present, as above, tutorials on some specialized statistical methods such as *Reliability and Logistics*, which can be successfully applied to *Triage*, *and to ICU* overcapacity problems.

*Design and performance evaluation of ICU units, using Reliability*

[https://www.researchgate.net/publication/342449617 Example of the Design and Operation of an ICU using Reliability Principles](https://www.researchgate.net/publication/342449617)

Health care facilities require implementing *efficient and effective Logistics* (e.g. ensuring PPEs, support equipment and workers are available). *Failure Modes and Effects Analysis* (FMEAs) and *Fault Tree Analysis* (FTAs) are two valuable *reliability techniques*. Finally, a reliability analysis example illustrates how to establish *efficient replacement schedules* for key ICU equipment.

*A summary of Fault Tree Analysis (FTA) and Failure Mode and Effects Analysis (FMEA)*

<https://web.cortland.edu/matresearch/FMEA&FTASummaryS2017.pdf>

This *Power Point presentation* provides a quick and simple *introduction* to the techniques of *Fault Tree Analysis* (FTA) and *Failure Mode and Effect Analyses* (FMEA). It includes real examples and the corresponding icons used to build FTAs.

*An Example of Survival Analysis of Covid-19 using ICU & Patients Ventilator data*

[https://www.researchgate.net/publication/342583500 An Example of Survival Analysis Data Applied to Covid-19](https://www.researchgate.net/publication/342583500)

First, we consider patient Ventilator *Time to Death or to Recovery*. We then include patient *covariate information* (e.g. age, number of *Comorbidities*, and if termination was due to death, or to recovery). We analyze the probability of *patient survival*, given Ventilator Time, age, Co-morbidities; and strength and direction of covariates effects in patient's Ventilator sojourn. We also implement *Discriminant and Regression Analyses* between recovering and dying Patients, *to assess which factors* impact them. The purpose of these procedures is *to help evaluate Triage* in the unfortunate case its application is required, *due to ICU overcapacity*.

*Statistical Methods to Accelerate Covid-19 Vaccine Clinical Trials using sequential analysis*

[https://www.researchgate.net/publication/344193195 Some Statistical Methods to Accelerate Covid-19 Vaccine Testing](https://www.researchgate.net/publication/344193195)

*Clinical Trials* are about assessing the success or failure *of a treatment or a vaccine*. In this paper, *industrial sampling plans* are used to help shorten Clinical Trials. We discuss *testing methods* and their *sample size requirements*, then, *sampling plans for attributes* (pass or fail) that follow the Binomial Distribution. We then explore *double and sequential sampling plans*, the *Expected Sample Number*, that assesses the efficiency of multi-stage sampling plans. We conclude with a *detailed tutorial on sequential probability ratio tests*. We illustrate these plan discussions via *numerical and practical examples*. All these methods can *help accelerate clinical trials* without affecting their quality or reliability.

*Survival Methods to Establish Covid-19 Vaccine Life (Length of its Effectiveness)*

[https://www.researchgate.net/publication/344495955 Survival Analysis Methods Applied to Establishing Covid-19 Vaccine Life](https://www.researchgate.net/publication/344495955)

This paper provides a *detailed three-step tutorial on Survival Analysis*, applied to the study and assessment of *Life Length of Covid-19 Vaccines*. First, survival analysis was implemented to assess *rate of decay of Vaccine Life Length*, in time, to establish its *Useful Life*. Secondly, we studied how Life may be modified by *population factors* such as *age, socioeconomic conditions, gender, number of co-morbidities*, etc. Then, *Discriminant Analysis* between the two vaccinated groups (those later infected with Covid-19, and those that were not) was implemented to *identify which factors impacted Vaccine Life and Efficacy*. Finally, a *multiple regression* analysis was implemented to *quantify* such impact. It was determined that, as patient Age and Number of Co-morbidities increases, the length of time (Life Length) and efficacy of the Vaccine diminishes. Such *results* are used to *determine the number of Vaccine doses required for immunization*.

*Covid-19 ICU Staff and Equipment Requirements using the Negative Binomial*

[https://www.researchgate.net/publication/345914205\\_Covid-19\\_ICU\\_Staff\\_and\\_Equipment\\_Requirements\\_using\\_the\\_Negative\\_Binomial](https://www.researchgate.net/publication/345914205_Covid-19_ICU_Staff_and_Equipment_Requirements_using_the_Negative_Binomial)

In this article we *estimate a health care system* (e.g. hospital, ward, ICU) *load and operating requirements* (number of beds, of doctors, nurses, ventilators, etc.) to successfully cope with a possible system overload, during a given Covid-19 wave. We implement *Poisson and Negative Binomial Distributions* to assess *ICU overflow*, and estimate *Staff and Equipment requirements*. We *simulate patient admissions* using these distribution parameters, obtaining *average, best or worst cases*, evaluating these results using *survival analysis*. We *compare and discuss ICU staff and equipment requirements*, estimated under patient overflows, obtaining *Confidence Intervals* with different probabilities of admission. *Probabilities of survival on an ICU*, given patient age and co-morbidities, are estimated, including *expected times to death*. All this *information* may be used to *determine patient allocation* to an ICU and ventilator facility, *in case of Triage*.

## **5.0 Multivariate Analysis (Principal Components; Discriminant Analysis):**

In this section we present, in the same format as done above, *Principal Components and Discriminant Analysis* statistical *tutorials*. These methods can be successfully *applied to Triage, and ICU overcapacity* problems.

*Principal Components and Discriminant Analysis of Covid-19 data: Part I*

[https://www.researchgate.net/publication/341385856\\_Multivariate\\_Stats\\_PC\\_Discrimination\\_in\\_the\\_Analysis\\_of\\_Covid-19](https://www.researchgate.net/publication/341385856_Multivariate_Stats_PC_Discrimination_in_the_Analysis_of_Covid-19)

The *analysis* below uses *real data taken from New York State Regions*, released on V/13/20 by NYS authorities. It illustrates the power of *multivariate analysis* in the Coronavirus Pandemic struggle. A *Principal Components* analysis was implemented to *assess variable impact* on each Region's infection rate. Principal Components *reduces the number of analysis variables* to a *smaller set of highly significant* ones. A *Discriminant Analysis* was then used to assess which of these variables does a better job of *separating the Regions into* those that can *safely open* their economy, and *those that should wait*. Using such variables within *Fisher Discriminant equation*, NYS regions can be *periodically re-evaluated*, using subsequently obtained NYS data. This is a preliminary study, subject to further data validation and verification. We encourage other researchers to substitute our data in the tutorial and input theirs, and get similar results for their regions.

*More on Principal Components and Discriminant Analysis of Covid-19 data: Part II*

[https://www.researchgate.net/publication/342154667\\_More\\_on\\_Applying\\_Principal\\_Components\\_Discrimination\\_Analysis\\_to\\_Covid-19](https://www.researchgate.net/publication/342154667_More_on_Applying_Principal_Components_Discrimination_Analysis_to_Covid-19)

This *paper is a continuation of the above Multivariate Statistics in the Analysis of Covid19*. We use the 62 New York state counties to (1) assess *which metrics* better *separate* them, and (2) obtain an *equation to classify* regions into *high and low risk* groups, according to said metrics. Variables used are: (1) *percent positives per 10K* (of county population) and (2) *deaths per 10K*. Since we don't have county population and thence, can't obtain county density, we use, (3) subjective *Urban v. Rural* status. We implement *Principal Components* to identify (1) which *variables* most significantly help us *differentiate* between *high and low infection counties*. Then, using identified variables, we develop (2) *Discrimination Functions* to *classify NYS counties* into high and low risk groups, according to their Covid-19 metrics. We use the *Principal Component scores* to plot the different counties and *assess how counties are similar (clustered), or different, according to the Axes: Positives and Deaths*.

*Principal Components and Discriminant Analysis of Covid-19 data: Power Point*

<https://web.cortland.edu/matresearch/ApplyPrincCompDSCC-Cov19.pdf>

This *Power Point* was used to *present the two above Multivariate Analysis papers* to the *Data Science Conference on COVID-19* August 28, 2020. It will help better understand their content.

*Logistic Regression/Discriminant Analysis Identifying key Covid Vaccine Clinical Trial Factors*

[https://www.researchgate.net/publication/346956247\\_Logistic\\_Regression\\_in\\_Factor\\_Identification\\_of\\_Covid-19\\_Vaccine\\_Clinical\\_Trials](https://www.researchgate.net/publication/346956247_Logistic_Regression_in_Factor_Identification_of_Covid-19_Vaccine_Clinical_Trials)

A *tutorial on the use of Logistic Regression to identify key factors in Covid-19 Clinical Trials*. The data analyzed was produced by this researcher using his experience and information. *We compare Logistics Regression with Discriminant Analysis methods* by analyzing same clinical trial data to identify Covid-19 factors that affect vaccine effects. By either method *we detect the same factors* that differentiate between both infected groups (vaccinated and placebo). This *approach can be reproduced by substituting, as responses, placebo v. vaccinated; by infected v. not infected; by deceased v. surviving; by Vaccine A v. Vaccine B, etc*. Researchers *may also use, as factors, other patient characteristics* such as weight, age, gender, occupation, number (or types) of co-morbidities and levels of interaction. We conclude that both *Logistics Regression and Discriminant Analysis results are equivalent. Logistics provides the probability of a data element belonging to one or the other group under consideration*. Statistically significant factors identify the *key model components* on which researchers should undertake further research.

## **6.0 Stochastic Processes Applications (Markov Chains):**

In this section we present, following the same manner as done in the sections above, several *tutorials on modeling* the trajectories of Covid-19 processes as *Markov Chains*. Through their movement between the different Markov states, we can obtain transition probabilities and sojourn times, useful to assess logistics, compare rates, triage, etc. By running different infection rates and facility logistics, we can compare the relative efficiency of different solutions, and select the best one.

*A Markov Chain model to study the spread of the Covid-19 virus*

[https://www.researchgate.net/publication/343021113 A Markov Chain Model for Covid-19 Survival Analysis](https://www.researchgate.net/publication/343021113)

*This Markov Chain illustrates the power that Markov modeling offers to Covid-19 studies. This article models the trajectory of Covid-19 infected patients into an ICU, and up to their death through a Markov Chain. We first consider a simple three-state, recurrent model. Its steady state probabilities are obtained for efficient and inefficient system comparisons. We then include additional absorbing states, to account for more complex situations. Using TPM we obtain the (1) probability of death of a Patient; and using their sojourns in the different states, (2) their expected time to death. The results are useful in establishing (1) logistic requirements of health care units, to provide excellent patient care, and (2) some objective Triage procedures, if ever such extremes are required. This Markov Chain tutorial has been, by the number of hits in its LinkedIn and ResearchGate web pages, our most read Covid-19 report.*

*A Two-Absorbing-States Markov Chain to study the problem of Covid-19 Herd Immunization*

[https://www.researchgate.net/publication/343345908 A Markov Model to Study Covid-19 Herd Immunization](https://www.researchgate.net/publication/343345908)

*This second Markov model assumes that the virus will infect a large part of the population, thus preventing further community spread and yielding Herd Immunization. Our previous Markov Chain assumed there was neither a vaccine nor a treatment for Covid-19. Also that, if current infection rates remained unchecked, everyone would eventually die. This paper assumes that Covid-19 survivors become immune, thence, cannot become re-infected. There is much debate about employing Herd Immunity as an alternative solution for combating Covid-19. Our Markov Chain quantitatively analyzes such situation. The model obtains (1) the probability of a Patient death or immunization. Also, the (2) expected times to death (or to immunization) when starting from different states in the Space (which can be used in Triage situations). Transition rates can help compare efficient and inefficient medical strategies, as well as help establish an acceptable infection rate. Times spent in a State (Sojourn) help estimate the required size of health care facilities that will treat patients. Statistics models help answer many health questions, as well as compare the performance of different public health strategies, in a more objective, non-partisan way. As George Box, a well-known statistician, once said: all models are wrong; some models are useful.*

*A Markov Chain to study the problem of Re-opening Colleges under Covid-19*

[https://www.researchgate.net/publication/343825461 A Markov Model to Study College Re-opening Under Covid-19](https://www.researchgate.net/publication/343825461)

*This Markov Model studies the dilemma of Re-opening Colleges under Covid-19. We analyze the situation using a Markov Chain defined over a nine element state space that moves through a set of Transient states, eventually leading to two Absorbing States: Expulsion or Coursework Completion. The model, due to its specific State Spaces and transition probabilities is very useful to compare reopening plans. Through the infection (transition) rates we study their impact on the probabilities of Expulsion and Course Completion. Differing infection rates depend on student compliance with community public health measures such as face covering, social distancing, etc. By assigning different values to these rates, their impact can be assessed and compared. Once updated and fine tuned (or rebuilt) Markov models can be used by college authorities to re-assess and improve their reopening plans, by faculty and students, to assess their risks in such openings, and by government authorities, to assess the validity and safety of such plans, thus allowing or proscribing them. This model can also be modified to assess re-opening the public school system.*

GSS/ASA Workshop (U-Tube) on Covid-19 ( our Markov Chain presentation is the last one)

[https://www.youtube.com/watch?v=ByGEUGbU\\_JU&feature=youtu.be](https://www.youtube.com/watch?v=ByGEUGbU_JU&feature=youtu.be)

U-Tube recording of the *Applications in Using Blended Data Workshop*, sponsored by the ASA Government and Social Statistics Sections. Our Markov Chain presentation is the last one.

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

[https://www.researchgate.net/publication/347441411\\_A\\_Markov\\_Model\\_to\\_Assess\\_Covid-19\\_Vaccine\\_Herd\\_Immunization\\_Patterns](https://www.researchgate.net/publication/347441411_A_Markov_Model_to_Assess_Covid-19_Vaccine_Herd_Immunization_Patterns)

This *Markov model* assesses different *patterns of vaccination*, which may *affect* achieving (or not) *Herd Immunity*. The urgency of this paper stems from *polls suggesting* a significant number of *people are not willing to become vaccinated*. Herd Immunity can be acquired by (a) letting the virus infect most of the population. Weaker ones (the elderly, those with co-morbidities etc.) will die) and those surviving will become immunized; alternatively, (b) by vaccinating a large part of the general population. In either case, *the majority of the population must be actively involved*. Vaccination entails two aspects: one is individual and the other social. First, the *vaccine protects the individual*. Secondly, if enough individuals in the general population are vaccinated, the *activity has an effect over the Pandemic*. With few new customers to infect, the virus starves and disappears. If insufficient persons become vaccinated, this latter benefit is lost. By *changing the vaccination parameters* (e.g., infection rates, participation, and immunization percentages), model results will differ, allowing the *comparison of different public health strategies*.

## **7.0 Socio-economic Analyses of Problems derived from Covid-19:**

In this section we present four papers that discuss *conceptual issues affecting* the results or the interpretation of the statistical analysis of *Covid-19 problems*. Format is similar to that of previous ones.

*Fallouts of Off-Shoring (outsourcing) and Tax payers' contributions, to the Coronavirus Pandemic*

[https://www.researchgate.net/publication/341685776\\_Off-Shoring\\_Taxpayers\\_and\\_the\\_Coronavirus\\_Pandemic](https://www.researchgate.net/publication/341685776_Off-Shoring_Taxpayers_and_the_Coronavirus_Pandemic)

This paper discusses how the *Off-shoring American jobs* during a quarter of a century, was one of the main causes that *prompted the election of President Donald Trump*. His *initial handling* of Coronavirus Pandemic was *not efficient -his handling of Covid-19 vaccine research and production, was*. In addition, extended *Off-shoring affected* many industries that could have provided *essential* resources to help in the fight against Covid-19. The *result* of these factors was that, the *number of Covid-19 infections and deaths* occurred was *possibly larger* than it should have been.

*A Digression on the Interaction between Race, Ethnicity, Class and Coronavirus.*

[https://www.researchgate.net/publication/343700072\\_A\\_Digression\\_About\\_Race\\_Ethnicity\\_Class\\_and\\_Covid-19](https://www.researchgate.net/publication/343700072_A_Digression_About_Race_Ethnicity_Class_and_Covid-19)

This paper discusses the *allegation that, if African-Americans and Hispanics* were impacted by Covid-19 *above their population representation*, it was because *genetic* factors were involved. We indicate how it



*may not be race or ethnicity, but poverty and other negative socio-economic conditions such as insufficient or no health care, poor living environments, working in high risk jobs, living in overcrowded quarters, perhaps over more than one generation, what causes these populations to become infected in larger percentages. Notice and compare how, middle and upper class African-American and Hispanic households (e.g., President Obama, Secretary Condoleezza Rice, Professor Romeu) are affected in equivalent numbers, as their middle- and upper-class peers, in mainstream and white America.*

*A Digression about Aspects of Clinical Trials for the new Vaccine against Covid-19.*

[https://www.researchgate.net/publication/346305686 A Digression on Covid-19 Vaccine Clinical Trials and its Consequences](https://www.researchgate.net/publication/346305686)

We digress about *Covid-19 vaccine accelerated Clinical Trials* and the early release, by CDC, of several candidate vaccines made possible by the application of new techniques. We wrote two previous papers:

[https://www.researchgate.net/publication/344495955 Survival Analysis Methods Applied to Establishing Covid-19 Vaccine Life](https://www.researchgate.net/publication/344495955) about the use of survival analysis for assessing vaccine life length, and [https://www.researchgate.net/publication/344193195 Some Statistical Methods to Accelerate Covid-19 Vaccine Testing](https://www.researchgate.net/publication/344193195) about efficiently assessing vaccine test results.

*Some drawbacks on vaccine early release stems from limited testing time: failure to identify some side-effects, especially long-term ones; failure to determine the time length of immunization provided, or the particular cohorts (e.g. elderly) that each of the vaccines will benefit (or hurt). The world needs a working vaccine, even with potentially minor problems. After a few years of world-wide vaccination, herd immunity will be achieved. Meanwhile, we have to make-do with what we now have.*

*A Digression on Covid-19 Vaccine Rollout.*

[https://www.researchgate.net/publication/348607971 A Digression on Covid-19 Vaccine Rollout](https://www.researchgate.net/publication/348607971)

*Research for a Covid-19 vaccine was a complete success! However, the US rollout of such vaccines has proven a failure. Such poor rollout suggests the need to re-examine its approach, and redefine some of its methods. For comparison, in June of 1944 the Allied army successfully landed in Normandy. Over 1,200 aircraft and more than 5,000 vessels carrying 160,000 troops, medical supplies, water, food, ammunition, etc. were moved in a few days. Operations Research, the technique that made possible the invasion's essential Logistics, was born. The political and military leaders had the brilliance of establishing the key goals of the military operation, while letting OR people take care of the technical details of its logistics implementation. This article provides examples of proven OR, logistics and quality engineering techniques that have been successfully applied to improve many situations. Some of the methodology presented in this paper, may also help improve the current Vaccine rollout performance.*

## **8.0 Additional Work**

After completing the initial series of papers, we continued expanding their scope. Several months later (Fall of 2021) we wrote a paper on the use of *Quality Engineering Methodology in the Study, Assessment and Development of Covid-19 Systems*, either already deployed or to be soon released.

[https://www.researchgate.net/publication/352998703 Quality Engineering Methodology in Covid-19 Systems Design and Improvement](https://www.researchgate.net/publication/352998703)

In the Spring 2022 we developed, with our MFE634 Quality Engineering graduate students, a project on using *Quality Engineering Methodology for Improving the Distribution of Covid Vaccines Worldwide*, especially in Third World Countries, through the World Health Organization:

<https://web.cortland.edu/matresearch/WorldwideCovidDistrFinalGrp3.pdf>

## 9.0 Discussion & Conclusions

Most work reported here was undertaken between March of 2020 and February of 2021. In those twelve months, this researcher wrote *over two dozen papers, reports and tutorials*, on Covid19. After Spring 2021, we continued working on Covid19 issues, individually, and with our Quality Engineering students.

We followed *Eight Steps in developing* our Covid-19 oriented, *statistical papers*:

1. Select a current Covid-19 topic of interest
2. Research the topic from a Public Health perspective
3. Select a statistical method to deal with this topic
4. Review the theory of the statistical method to be used
5. Find/develop an appropriate data set for the analysis
6. Analyze the data using the stated method
7. Obtain/interpret the Results, and Conclude
8. Distribute the papers (via web, institutions etc)

The reader may estimate the *number of hours invested in this effort*. Certainly, we have put in several hundred hours of pro-bono work to help defeat Covid-19. We hope it has contributed.

*Combining the results* from the Markov Chains, Survival, and Multivariate Analyses, as well as other methods used in this series, *may provide tools* for public health and other professionals to *help establish rules for Triage* procedures, if such situation ever becomes necessary.

In our work we want to *reach four audiences*: (1) *public health* professionals and researchers, (2) *medical doctors*, (3) *statisticians* and (4) *the public* in general.

We want to *encourage public health and medical professionals* to use more statistical procedures and *do more joint work with statisticians* -not only after the data have been collected, but also at the time that the experiments are being designed.

*We want to encourage statisticians*, especially those retired, who have the experience, financial support (their pension), and the free time to provide such assistance, *to contribute in helping with the planning, implementation and analysis* of statistical procedures –or with writing about them.

*We want to provide illustrative examples* to doctors, public health researchers, government, and general public, to help them better understand what statisticians do, *fostering greater collaboration*.

Finally, *this series of papers on statistical analysis of Covid-19* problems can also be used as part of a *biostatistics course* in a public health or medical curriculum, or in an *applied modeling and data analysis graduate course*, in a statistics department.

## References

### Textbooks

- Acheson, J. Quality Control and Industrial Statistics (5th Edition). Irwin, 1986
- Anderson, T.W. An Introduction to Multivariate Statistical Analysis. John Wiley & Sons. New York. Second Edition. 1971.
- Beyer, W., Editor. Handbook of Tables for Probability and Statistics. The Chemical Rubber Co. (CRC). Ohio. 1966.
- Box, G., Hunter, W. G., and J. S. Hunter. Statistics for Experimenters. Wiley. New York. 1978.
- Cinlar, E. Introduction to Stochastic Processes. Prentice Hall. NJ. 1975.
- Gryna, F.; Chua, R. and J. DeFeo, Juran's Quality Planning & Analysis for Enterprise Quality (5th Ed.). McGrawHill, NY. 2007.
- Heyman, D. and M. Sobel. Handbooks in Operations Research and Management Science. Vol. 2: Stochastic Models. North Holland. Amsterdam. 1990.
- Kalbfleisch, J. D. and R. L. Prentice. Statistical Analysis of Failure Time Data. John Wiley & Sons. New York. 1980.
- Mann, N. R.; R. E. Schafer and N. D. Singpurwalla. Methods for Statistical Analysis of Reliability and Life Data. Wiley, 1974.
- Montgomery, D. C. Design and Analysis of Experiments. 2nd Ed. Wiley, New York. 1984.
- Reliability Toolkit*. Reliability Analysis Center. RAC.
- Scheffe, H. The Analysis of Variance. Wiley. New York, 1959.
- Taylor, H. and S. Karlin. An Introduction to Stochastic Modeling. Academic Press. NY. 1993.
- Walpole, R. E. and R. H. Myers. Probability and Statistics for Engineers and Scientists. Prentice-Hall. <http://www.elcom-hu.com/Mshtrk/Statistics/9th%20txt%20book.pdf>

## Reports

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

Romeu, J. L. *Determining the Experimental Sample Size*. Journal of Systems Reliability Center. (SRC): 3rd Qtr. 2005 (pp. 11-21).

Romeu, J. L. *Design of Experiments for Reliability Improvement*. Quanterion Reliability Ques. <https://www.quanterion.com/design-of-experiments-for-reliability-improvement/>

Romeu, J. L. *Quality Control Charts*. RAC START Sheet. Volume 11, Number 4. <https://web.cortland.edu/matresearch/QCChartsSTART.pdf>

Romeu, J. L. *Statistical Modeling of Reliability Data*. Reliability Analysis Center/RAC Journal. <https://web.cortland.edu/matresearch/StatModelingArtRomeuRac4q2001.pdf>

Alternative access to these papers:

*Quality, Reliability and Continuous Improvement Institute* applied statistics web site: QR&CII <https://web.cortland.edu/romeu/QR&CII.htm> (In English)

*Instituto de Estadística Aplicada y Mejora Continua* applied statistics web site: IEA&MC <https://web.cortland.edu/matresearch/QR&CIInstPg.htm> (In Spanish)

Coronavirus Page: <https://web.cortland.edu/matresearch/CORONAVARUSListEngNov2020.pdf>