**THE AFTERMATH OF THE COVID19 PANDEMIC ON WOMEN’S HEALTH: THE PREVENTABLE CARE LAPSE IN CANCER SCREENINGS**

Anthony S. Fargnoli, Senior Research Scientist

New York University Langone Medical Center

**Background and Significance:**

COVID19 has undoubtedly imposed global life-changing suffering and the adoption of health seeking behaviors in the post curve era as the world slowly re-opens for normalcy. The uncertainty of the "second wave", and many other factors have also influenced health seeking behaviors and normal healthcare within the shutdown period that has continued as of this writing. Core arguments suggest that women have been disproportionately affected by the aftermath in terms of societal factors in caring for children and family members affected by the crisis. Their risk of death to coronavirus is also known to be much lower as men are nearly 4-5 times as likely to die.

The triad of preventable cancers specific to women's health includes breast, colon, and cervical cancers which represent a very large lethal percentage for the female population in the USA. The readers may also have at least one family member – mother, aunt, cousin, close friend affected by one of these serious diseases as at least 1 in 8 women will develop one condition over their lifetime. The collective diagnosis rate alone, with 68% testing compliance is well over 300,000. It is well known mammograms for breast cancer detection, colonoscopies for colon cancer, and pap smears for cervical are high volume procedures that reduce the risk of cancer death ranging from 50-80% with early detection. Therefore, it is logical to conclude a "silent women's wave" of morbidity will be upon us as a community in the next year as critical testing facilities supporting these were shut down during the COVID19 pandemic.

Despite learning much more about the true impact of the COVID19 pandemic and its associated risks for relatively healthy women to safely receive these life saving screenings, draconian measures and fear around daily posted COVID19 statistics discouraged normal rates of screenings. Recent publications in cancer journals have indicated this problem with greater than 80-90% reduction in women attending their regularly scheduled screening, and or their screenings delayed for 6 months to 1 year to catch up the back volume of previously canceled tests. The real issue, however, as reported by most private practice, is the fear of not receiving routine care due to the fear of catching re-surged COVID19.

This project was inspired by private physicians group calling for the need to generate clear medical media communication targeted for healthy, but at-risk women who have neglected their routine care. What are the actual risks of an 18-70 years old female with no co-conditions to catch and die from COVID 19 vs. these 3 high rate killing cancers? What is the societal impact of not spreading this urgent message to our mothers, aunts, and loved ones? This project captures the best estimated statistics on COVID19 via computing adjusted risks against the actual risks for cancer in a clear, meaningful manner to encourage women to immediately obtain their annual cancer screenings.

**Methods:**

COVID19 numbers have been a mainstay in virtually everyone’s mind since the tracking of the media. John’s Hopkin’s University posted the first live cumulative global and regional death tolls and nearly over 10,000 sites now monitor actual COVID19 death statistics based on hospital pathology reporting and national institutes of health detailing many other pandemic related statistics. Many of these are ripe for data science and have been posted on Kaggle for analysis for different studies.

This project sought to scrape an updated reliable source on World O meter.org, which is known for accurate up to date peer-reviewed data feed that updated by the hour. The web site linked for scraping: <https://www.worldometers.info/coronavirus/#countries>.

Data fields on the site are displayed clearly with sortable table features in HTML. The fields specific to COVID19 of interest to scrape for this project were: Country, the total number of confirmed cases via test, total cases per 1 million population, total population, serious cases, total tests, and total tests per 1 million population. These statistics and others update daily, however, it is the well known debate that the true risks are not captured by these raw numbers because of these basic pandemic facts:

1. COVID19 disproportionately affects those over 65, specifically in nursing/long term care facilities and men; thus women under 65 are much less at risk
2. COVID19 has a large, modeled but not very certain, population that is either asymptomatic OR have some form of speculated pre-existing immunity from the previous SARS1. These issues are known and have been published
3. These statistics do not stratify other certain risks for people to weigh their options on the best course of action such as obtaining cancer screenings.

Scrapy fields were created in the items file for many of the variables in the table and programmed into the spider file. Careful attention to the X tags in this particular table was noted and mapped accordingly in the row names xpaths. Please see attached files and R file on supplemental materials and GitHub for complete details.

Once data was successfully scraped from the World O meter site after validation of accuracy, the Coronovirus\_stats.csv file output was loaded into R package for data science analysis. Dplyr was used to clean data based on these criteria:

1. Countries with at least 5,000 published deaths as of this writing on June 17, 2020. This was to ensure very accurate data relevant to contemporary times in societies with full health systems and also accurate COVID19 data as many countries listed that were smaller had missing data in key elements. This narrowed the field to 16 countries.
2. The country data reflects what the news media projects daily in actual death/morbidity toll that the average person views, but are not adjusted.
3. True risks using the best available data and academic model sources computed the adjusted risks of dying from COVID19 based on the status of woman, age, and not being in a long term care/nursing home facility. These data were plotted by the country.
4. Following the country analysis which provided measured risk on actual and adjusted data, the goal was then to shift to identify cancer risks from national sources. These statistics commented in the R file were taken from posted sources. It is well known to the public and has been written, that cancer kills at a much higher rate than COVID19 based on what we know.
5. Specific screening, and pandemic non-screening cancer toll impact risk computed due to the decline in preventative care during shutdown and slower volume recently reported despite re-opening.
6. Summarizing, the organization of this data is to generate a strong push for women to continue and immediately schedule their delayed routine OBGYN care. Please pass this message on as it can save lives – 80-90% of many deaths are preventable with routine screening.

**Results**

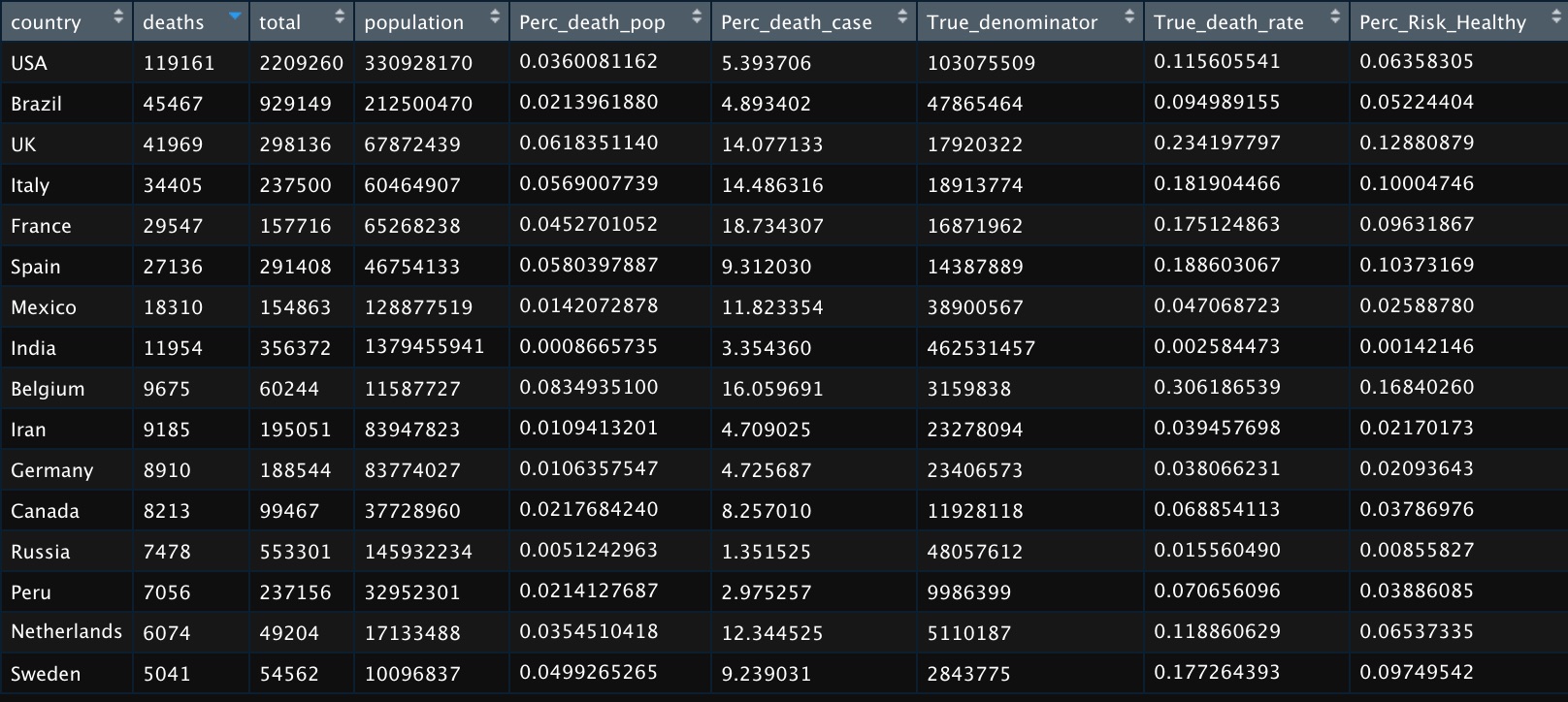


Figure 1. Main data frame after importing of scrapy derived coronavirus\_stats.csv file

The fields of country, deaths, total (i.e. total confirmed COVID19 cases), population per country were selected from the range of columns in the primary format. From here, assumptions and basic statistics were applied to encapture real numbers in terms of COVID19 death risk per country. The purpose of this portion of the analysis was to provide a frame of reference to women, particularly in the USA, to present other percentages to interpret global risk. Note, the original file had over 200+ countries, but the top 16 countries in terms of deaths >5000 were held for further analysis for two key reasons: 1. Each of these 16 countries had the infrastructure for both testing and reporting in hospital systems 2. The death toll of less than 5000 most likely indicated the country likely had more deaths or an inability to accurately track. This has been the basis of much reporting error in models and other media outlets whereby percentages could be skewed in either direction with small testing sample sizes. These 16 countries as shown by robust testing numbers had the proper capacity to produce further, less skewed analysis.

*Mutated Columns for Key Rate Statistics*

These columns were added via mutate function in R and used the information and several assumptions to provide additional insight:

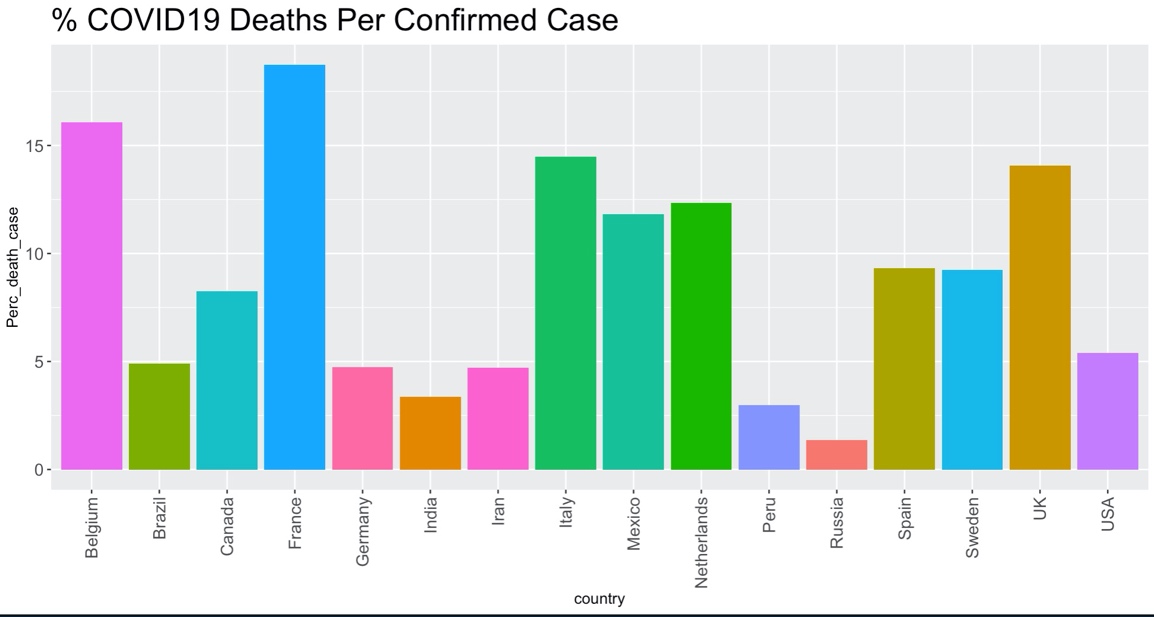
1.Perc\_death\_case – Simply divided the total confirmed deaths by the total column or confirmed number of cases times 100 to obtain a kill rate percentage in terms of the cases known. These statistics have been reported somewhat accurately in the media and is the consensus of what the public views as the true COVID19 kill rate in all general population. A quick walk shows the USA at 5% which is shown on media usually, followed by coming poor performing countries like Italy and France, who had the older population and fewer resources to deal with the pandemic with rates as high as 18%. These numbers are a function of many variables including testing, demographics, health system quality, pandemic preparedness, and others. Generally unreliable alone in many opinions in the aftermath.

2. Perc\_death\_population – More sobering analysts critical of the COVID19 response that has advocated for no quarantines have used this statistic dividing current deaths by the total population of the country which is expected to be low. For example in the USA where many argue the death toll has been exaggerated due to confounding tallies and co-morbidities, even 115,000 deaths as of this writing divided by 330M people results in a kill rate of 0.036% of the population. Granted this number may grow, but it will never reach the level of original models projecting at least 1, some as high as 5% of the population dead in one year.

3. True denominator derived “True death rate”– The skeptics with solid reason have identified the over projected bias and error in percentages based on several facts regarding what the true denominator should be to calculate the true kill rate of COVID19. We've learned recently in several large research reports that the spread of COVID19 was likely before the January 2020 timeline and that a large % of the eastern and western USA coasts were already exposed. The argument is that by the time we were able to "test" the presence of the virus it was well along its curve and that other factors were likely not explaining the rate of spread vs. what is now accepted as a relatively lower kill rate in healthier groups. Deferring to the literature without taking much of one side to this polarized debate, are two facts: 1. Asymptomatic carry the COVID19 and maybe transmit, but have no actual illness or morbidity 2. Recently publications in Nature and other top science journals hypothesize and explain that a subset of people who recovered from the previous SARS have to overlap pre-existing T cell immunity. In two large antibody studies, another controversial analysis showed that between 25-40% of LA and in some NYC populations who were healthy and never presented with COVID19 had antibodies, meaning they were infected and exposed to the SARS2 strain of COVID. To model this, we took a random variable generated in R and averaged up to 10 iterations for each country of a modified, more true denominator assuming on average 20% were unaffected due to different immune profile but could spread, not die, and another 20-40% random assumption of previously infected with antibodies that are confirmed immune and never had symptoms. Additionally, children or young adults under the age of 18, notwithstanding rare co-morbidities and illness, were nearly 100% immune to COVID19 issues due to not having the receptor in their lungs to incubate the virus. Thus from the general population, the model here discounts each population by 20% to remove the significantly young population and then discounts a range between 20-40% that either had pre-existing immunity or asymptomatic carrier. This results in a number much higher than the "confirmed cases" which are biased since most testing is only for people that are either hospitalized or have symptoms. The deaths were divided by this true estimate to provide a revised estimate of deaths per including modeled people that were infected but never tested or reported symptoms. This reduces the cases kill rate in the USA from 5% confirmed to less than 0.12%. The critics use these statistics to demand instantaneous re-opening since the seasonal flu kills at around this rate of 0.1-0.2% in bad seasons as found in the past.

Women’s Risk of Perc\_Risk\_Healthy – this field was created by further adjusting the True Death rate by risk adjustment for women’s age [18-65 years] and non-critical care [no long term care or nursing home] condition. The general COVID19 consensus against the severe draconian shutdown measures is based on three key facts for women in particular: 1. Nursing homes and long term care facilities have accounted for >50 to up to 70% of total deaths were at least age-related or 1 or more comorbidities was the true underlying cause 2. Women's risk of death to COVID19 ranges 33%-50% lower than men. The best available data demonstrates muscle mass wherein men iron stores increase ferritin levels, a biomarker of inflammation in the COVID19 cytokine storm, and other factors such as smoking/alcohol which have always been higher in men 3. Women less than 65, many in childbearing years, make up a very large % of demographics in most countries. In the USA our estimates show about 107 million. This statistic reduced the overall death totals in each country by a factor of 55% roughly matched to non-nursing home care, female, and age less than 65, then used the true denominator of the female population to estimate the risk of COVID19 death for this demographic.

COVID19 Chart Results



**Figure 2. Percentage of actual deaths per confirmed tested positive case.**

The chart in Figure 2 demonstrates the range of data in terms of actual cases where patients died divided by the denominator of COVID19 diagnosed or tested positive. The USA as of this writing sits about 5% and known countries that struggled such as Italy, France, and surprisingly Belgium, as high as 15%. Demographics and public health measures such as age, smoking history, and healthcare quality affect these figures. The most interesting to view is Sweden, which unlike other countries did not impose lockdowns. Sitting at 10%, they did no shutdown to the effect of Italy, Mexico, Netherlands, and France, yet had a lower rate. This chart as part of the project to communicate to patients that risk is diverse and not based on what specific measures any one country took to prevent deaths and that such prevention needs to be evaluated in terms of the individual patient.

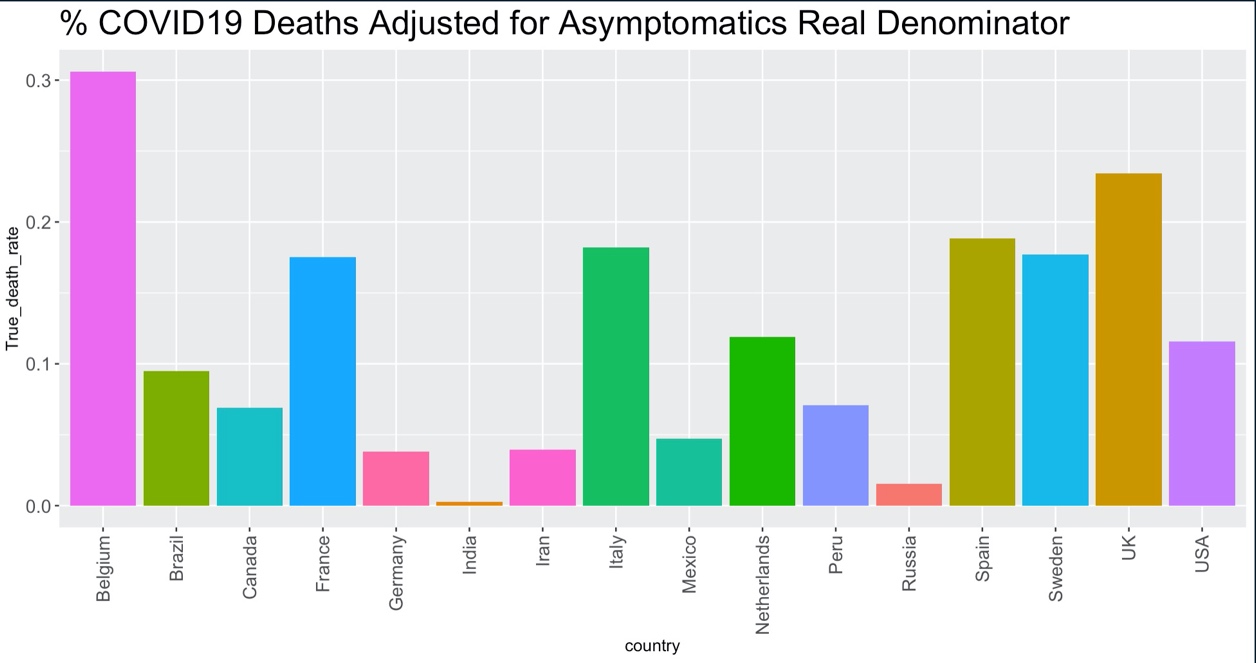


Figure 3. Corrected death rates based on asymptomatic and some potential pre-existing immunity to COVID19.

This realistic projection is what most epidemiologists now certify as the "true" kill rate of COVID19, expanding the denominator of asymptomatic 20-40%, excluding children, that contracted COVID19 or had antibodies but did not succumb to death. The USA rate sits at about 0.12% in our assessment others have it reaching 0.2% per model, however, this is significantly lower than 5%. This important figure combined with above communicates to patients their risk, if healthy, is extremely low as it stands >99.8% would survive an infection.

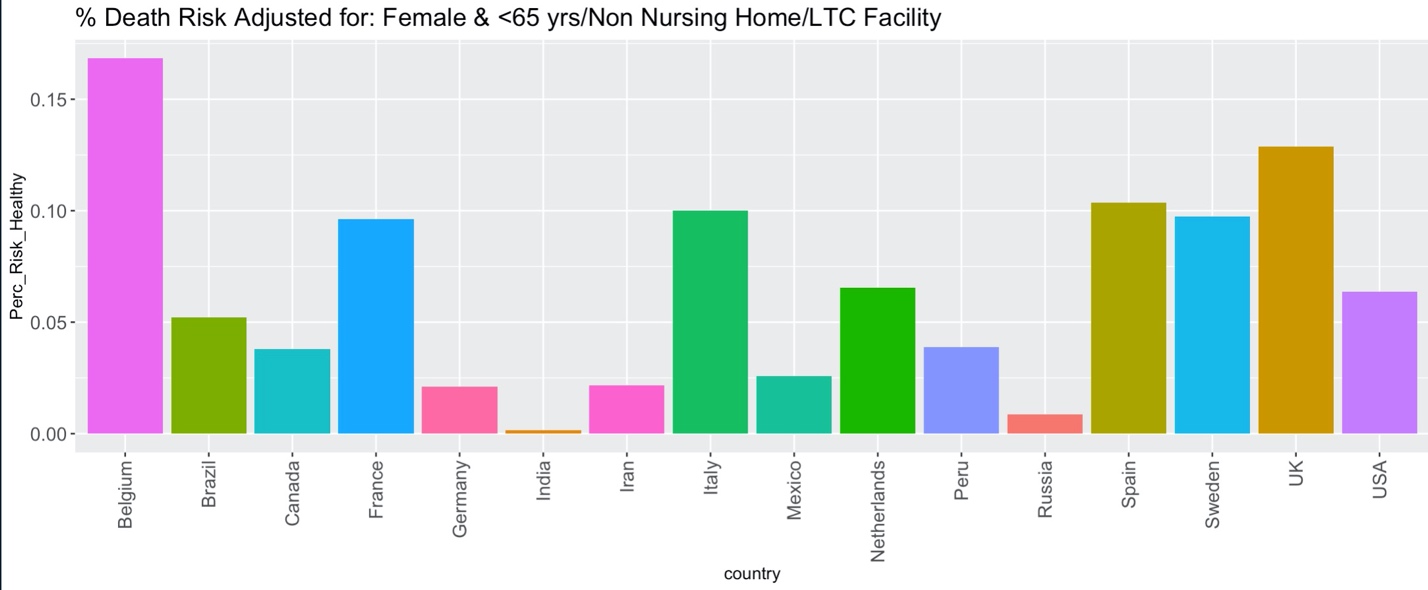


Figure 4. COVID19 Death Risk for Women in the low-risk category of age less than 65 and not in a long term care facility.

This project was determined by estimating the % of total women in the USA <65 years of age plus not having a long term care living situation. This was estimated by taking the total number of female coronavirus deaths and discounting further by 55%, which on average is the number of deaths due to long term care and or nursing home arrangements or less than 65 years of age. This number was also expected to be lower since females as a whole only account for 33-40% of all COVID19 deaths. The results are clear that in normal demographics the kill rate drops to a very low level of 0.057% in the USA. Notably, in Sweden where no social distancing or shutdown occurred, only 0.1% in this demographic died or 99.9% survival.

**The central question did the risk mitigation of COVID19 by governments do more harm than good for the majority of women? In terms of cancer, we examine this case, and to advocate the risk of not having routine care may kill many more in the long run through 2021.**

Diagnosis Statistics for Women's Health-Related Cancer as Per Routine Screenings

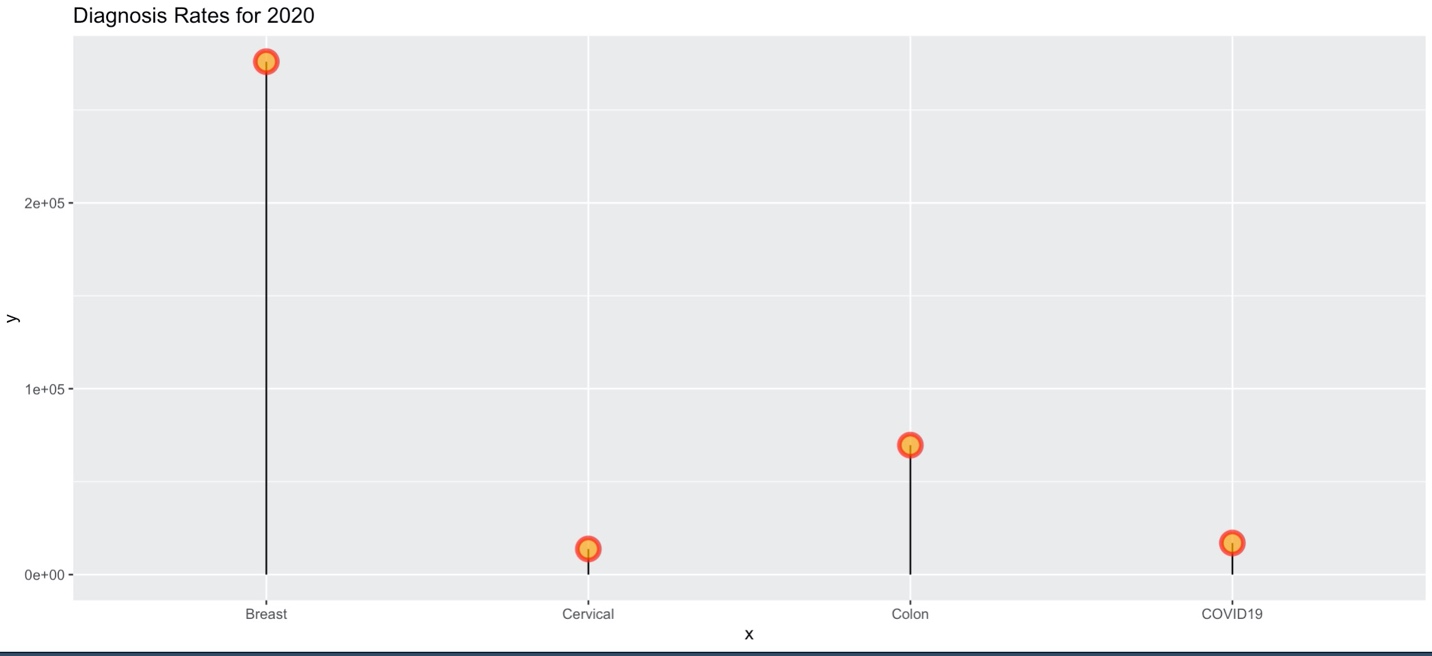


Figure 5. Diagnosis rates experienced yearly by women for the 3 major preventable death cancers via early detection of Breast, Cervical, and Colon compared with the projected number of COVID19 women who will be diagnosed remain of 2020.

This chart communicates several points, first and foremost breast cancer by a wide margin at a diagnosis rate per biopsy of 276,000 is at a high rate. Of this percentage diagnosed early, many will survive or will have a much less severe breast cancer course. To a lesser but still important degree, cervical about 13,000 annual cases diagnosed. Moreover, colon is still at a 57,000 case rate per year. The best available models demonstrate the downward trajectory of COVID19 cases, the most models project is only 20,000 additional women diagnosed in 8-10 months. Given that cancer is far deadlier, rates between 20-40% and as high as 90% in worst cases, any change in normal preventative care could have a major impact.

**Top academic reports in journals have recently shown with lockdowns, delays, and public fear – screenings have dropped on average 80-92% in all breast, colon, and cervical cancer screenings.**

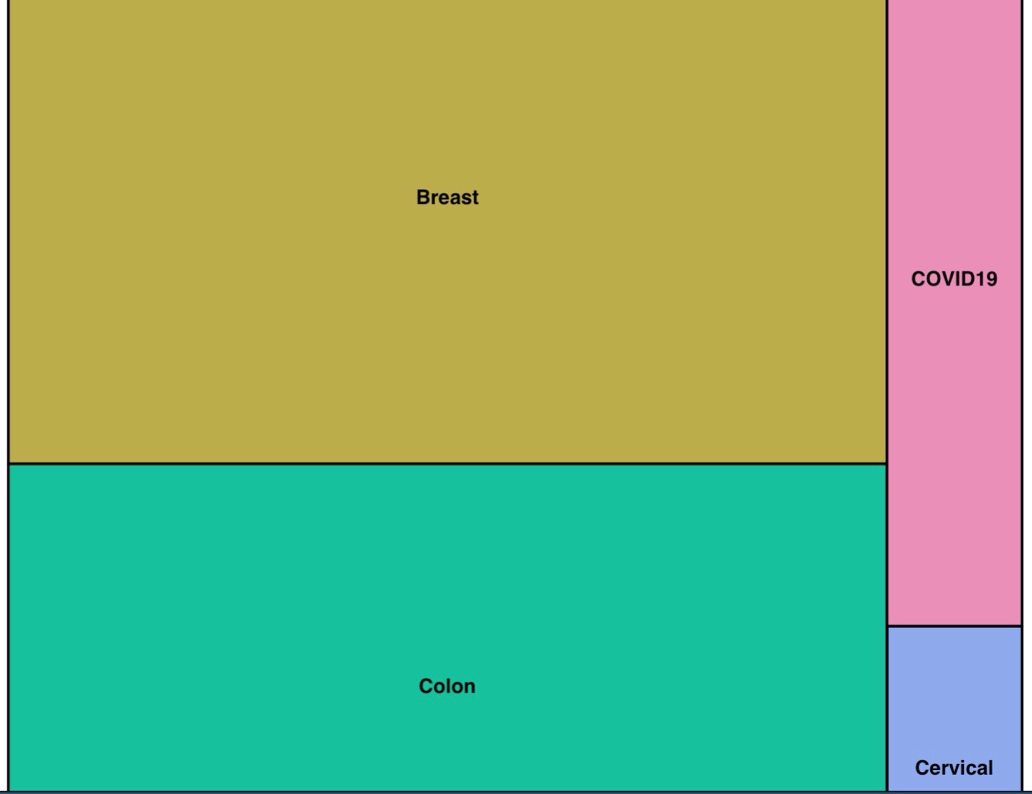
****

Figure 6. Tree diagram representing proportional total deaths for cancer-based on historical rates and contemporary COVID19 in women <65 years. Breast cancer dominates with on average 42,710 deaths, Colon cancer 24,570 deaths, and cervical 4,270 deaths.

Annually, it is observed that breast cancer alone claims the lives on average of 42,710 lives. Colon cancer claims 24,570 deaths. Projected COVID19 cases in the USA for women in the low-risk demographic, many of which even have co-morbidities, will reach only 21,234, of which only 1-10 would die. Cervical cancer is at a lower rate of incidence of 4290 but still sizable and increasing in younger urban populations.

With some basic assumptions of how cancer death rates will increases in 2020-2021 due to the lapses in critical screening, this projection below has been developed. A basic percentage model of lives saved as a function of screening was applied in reverse with projected breast, colon, and cervical cancer rates. Deaths can be projected given that exams have been reported 80-90% lower since pandemic through current stages, even with gradual volume increase there is still a large percentage who have missed exams or are afraid to schedule due to restrictions and fear of COVID19.

**In summary, total death rates to these cancers could increase by up to 60% due to this lapse. This would translate to summing average annual deaths of all three which is 115,000 roughly, and increase to 184,000. The best projection available for women aged 30-65 years acquiring and dying from COVID19 is less than 100 women from diagnosis of only 21,000.**

What are the actual risks in a clear numbers view?

The final purpose of this work was to generate a clear, simple chart demonstrating if women in this demographic either choose to do nothing or not have access to their screenings, what the risk is of acquiring cancer vs. dying from COVID19. Women need informed percentage with solid numbers to see this risk of staying in fear of not scheduling health exams to weight the risk fairly.

In this analysis we took the average diagnosis rates of all three cancers and divided by female population eligible for most screenings, ages 33-85:

Breast = 276,000 / 89e6 = 0.31%

Colon=69,650 / 89e6 = 0.07%

Cervical=13,800/89e6=0.015%

Total risk = 0.4%

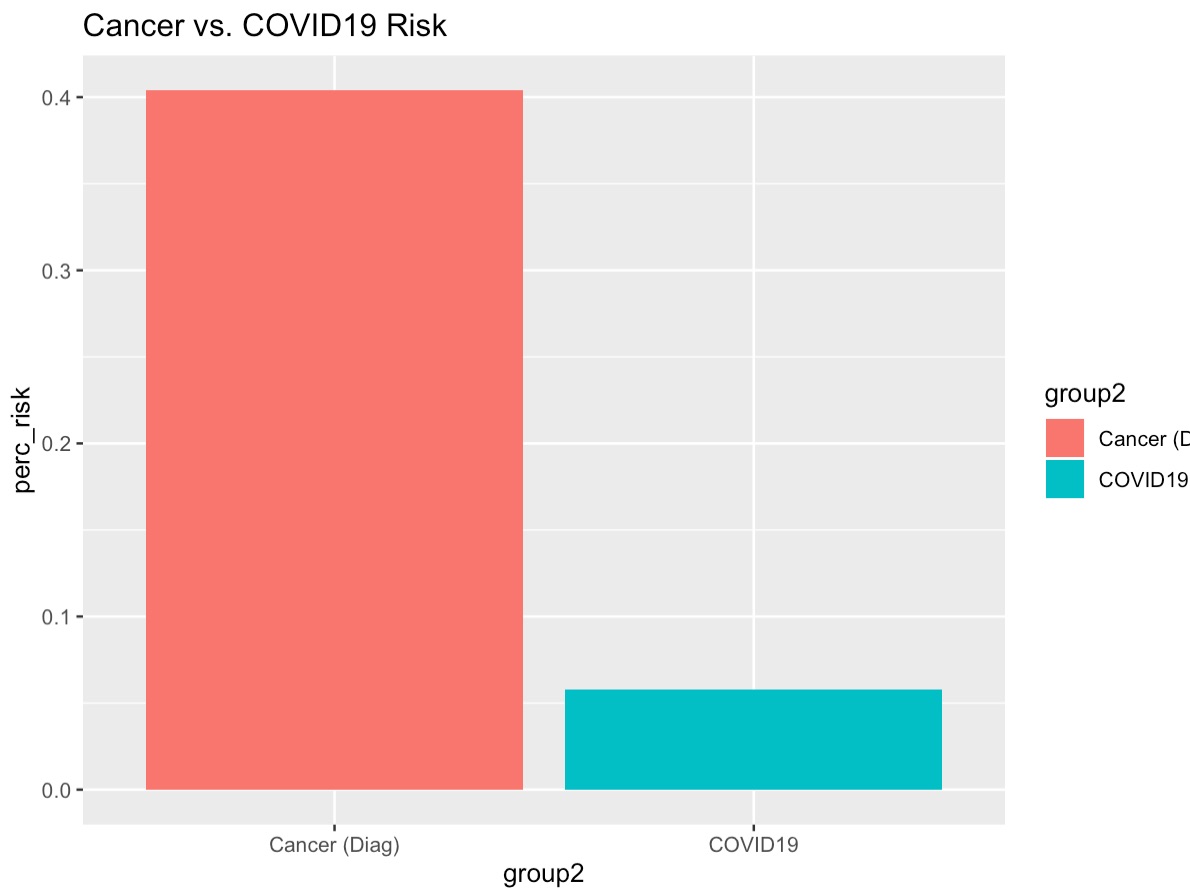


Figure 7. Risk of having a cancer diagnosis vs. acquiring and dying of COVID19. This summarized chart reveals a baseline risk of 0.4% of receiving a positive diagnosis with routine examinations against the risk of COVID19 death at 0.057%.

Limitations and Conclusions:

There are several limitations and possible improvements to this analysis. There are many other behavioral and societal factors affecting the women population from rescheduling their annual screenings. Many could be economical hardship, caretaker role, and or combination and not driven by fear of COVID19. However, the general perception of younger populations that have lower, but still risk for cancer in 30-50 years can be attributable to not having the correct risk information. Additionally, as in other models, more sophisticated projections could have been made however the goal was to use robust baseline numbers based on historical cancer data and actual USA deaths.

Another important factor is not in any way is this analysis reducing the serious situation of COVID19 pandemic. While it can be argued avoiding screenings may have limited the spread in healthy, which in turn limited the spread to the at-risk; this conclusion has to be weighed against the lives we may lose of women who would normally be saved via screening. These women are mothers, drivers of the economy, and integral to society. A greater percentage of losses in this demographic would most definitely have a long term impact, especially in minority and underprivileged communities where women have an even stronger presence in family structures.

**Conclusions**

COVID19 is a serious disease for a risk population in long term care situations, but the risk to healthy women is low. The risk of acquiring and dying from cancer is higher and should serve as a motivating factor to schedule screenings as soon as possible now that restrictions have been lifted. Government policymakers in the future of pandemic management, should not shut down essential women's health services as the aftermath could be deadlier than the pandemic itself for this demographic. Losing this demographic could have much worse societal implications than in any other since women serve as family caretakers, educators, leaders, and contributing members to society.

**References**

Pink Breast Cancer Foundation

<https://www.nationalbreastcancer.org/breast-cancer-awareness-month?gclid=EAIaIQobChMIuITEvdKN6gIVS7LICh2iywP4EAAYAyAAEgJ20vD_BwE>

Cancer.org

<https://www.cancer.org/about-us>

COVID-19 Antibody Seroprevalence in Santa Clara County, California

Eran Bendavid, Bianca Mulaney, Neeraj Sood, Soleil Shah, Emilia Ling, Rebecca Bromley-Dulfano, Cara Lai, Zoe Weissberg, Rodrigo Saavedra-Walker, James Tedrow, Dona Tversky, Andrew Bogan, Thomas Kupiec, Daniel Eichner, Ribhav Gupta, John Ioannidis, Jay Bhattacharya

doi: <https://doi.org/10.1101/2020.04.14.20062463>

Presence of SARS-CoV-2 reactive T cells in COVID-19 patients and healthy donors

Julian Braun, Lucie Loyal, Marco Frentsch, Daniel Wendisch, Philipp Georg, Florian Kurth, Stefan Hippenstiel, Manuela Dingeldey, Beate Kruse, Florent Fauchere, Emre Baysal, Maike Mangold, Larissa Henze, Roland Lauster, Marcus Mall, Kirsten Beyer, Jobst Roehmel, Juergen Schmitz, Stefan Miltenyi, Marcel A Mueller, Martin Witzenrath, Norbert Suttorp, Florian Kern, Ulf Reimer, Holger Wenschuh, Christian Drosten, Victor M Corman, Claudia Giesecke-Thiel, Leif-Erik Sander, Andreas Thiel

doi: https://doi.org/10.1101/2020.04.17.20061440