



## Article

# Evaluating COVID-19 Vaccine Efficacy Using Kaplan–Meier Survival Analysis

Waleed Hilal <sup>1</sup>, Michael G. Chislett <sup>2,\*</sup>, Yuandi Wu <sup>1</sup>, Brett Snider <sup>3</sup>, Edward A. McBean <sup>2</sup>, John Yawney <sup>4</sup> and Stephen Andrew Gadsden <sup>1</sup>

<sup>1</sup> Department of Mechanical Engineering, McMaster University, Hamilton, ON L8S 4L8, Canada; hilalw@mcmaster.ca (W.H.); wuy187@mcmaster.ca (Y.W.); gadsden@mcmaster.ca (S.A.G.)

<sup>2</sup> School of Engineering, University of Guelph, Guelph, ON N1G 2W1, Canada; emcbean@uoguelph.ca

<sup>3</sup> College of Engineering, Mathematics & Physical Sciences, University of Exeter, Exeter EX4 4QF, UK

<sup>4</sup> Adastra Corporation, Toronto, ON M5J 2J2, Canada; yawneyj@mcmaster.ca

\* Correspondence: chislett@uoguelph.ca

**Abstract:** Analyses of COVID-19 vaccines have become a forefront of pandemic-related research, as jurisdictions around the world encourage vaccinations as the most assured method to curtail the need for stringent public health measures. Kaplan–Meier models, a form of “survival analysis”, provide a statistical approach to improve the understanding of time-to-event probabilities of occurrence. In applications of epidemiology and the study of vaccines, survival analyses can be implemented to quantify the probability of testing positive for SARS-CoV-2, given a population’s vaccination status. In this study, a large proportion of Ontario COVID-19 testing data is used to derive Kaplan–Meier probability curves for individuals who received two doses of a vaccine during a period of peak Delta variant cases, and again for those receiving three doses during a peak time of the Omicron variant. Data consisting of 614,470 individuals with two doses of a COVID-19 vaccine, and 49,551 individuals with three-doses of vaccine, show that recipients of the Moderna vaccine are slightly less likely to test positive for the virus in a 38-day period following their last vaccination than recipients of the Pfizer vaccine, although the difference between the two is marginal in most age groups. This result is largely consistent for two doses of the vaccines during a Delta variant period, as well as an Omicron variant period. The evaluated probabilities of testing positive align with the publicly reported vaccine efficacies of the mRNA vaccines, supporting the resolution that Kaplan–Meier methods in determining vaccine benefits are a justifiable and useful approach in addressing vaccine-related concerns in the COVID-19 landscape.

**Keywords:** COVID-19; delta variant; Kaplan–Meier; omicron variant; survival analysis; vaccines; variants; variants of concern



**Citation:** Hilal, W.; Chislett, M.G.; Wu, Y.; Snider, B.; McBean, E.A.; Yawney, J.; Gadsden, S.A. Evaluating COVID-19 Vaccine Efficacy Using Kaplan–Meier Survival Analysis. *BioMed Informatics* **2024**, *4*, 2117–2132. <https://doi.org/10.3390/biomedinformatics4040113>

Academic Editor: Themis Exarchos

Received: 21 July 2024

Revised: 22 August 2024

Accepted: 30 September 2024

Published: 12 October 2024



**Copyright:** © 2024 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

## 1. Introduction

The rollout of vaccinations to help protect against contracting SARS-CoV-2 and potentially incurring serious illness has been a priority of many public health authorities and governments worldwide since their initial emergency approval in late 2020, including Ontario, Canada [1,2]. While clinical efficacies of the available vaccines indicate immunization-based protection against COVID-19, population-driven probabilities have important potential in describing the success of vaccines.

Survival analysis is a statistical concept wherein the probability of survival, from a defined event, can be developed using time-to-event data. In COVID-19 research, applications of survival analysis have been only scarcely published, despite having significant potential to inform on the trajectory of the epidemic, in a variety of facets. One such facet is the use of survival estimates to understand the probability of testing positive for the virus among populations of vaccinated individuals. The results of these analyses represent significant potential to indicate the merits of vaccines.

An Ontario-wide collection of positivity data for COVID-19 testing was assembled for individuals who had also received one or more doses of a Health Canada-approved vaccine. From these data, a Kaplan–Meier survival analysis approach determined the survival probabilities for testing negative for the virus. Since “survival” in this case is “not receiving a positive PCR test”, the inverse survival probability was then calculated to determine the probability of testing positive for the virus. Applied to separate age categories, these results show the age- and vaccination-specific probabilities of receiving a positive test for COVID-19.

The objective of this study is to quantify and describe the success and benefits of approved COVID-19 vaccines using a survival analysis approach. The novelty of applying Kaplan–Meier methods in the context of calculating probabilities of contracting COVID-19 despite having received multiple doses of a vaccine places this investigation at the forefront of pandemic research. This article introduces various contexts in which survival analyses were applied, beginning with an overview of the Kaplan–Meier method in Section 2. Section 3 details the methodology and procedures used to manipulate the Ontario dataset for analysis. Section 4 reports the results of the survival analysis for different age groups and various vaccines. Section 5 provides a discussion on the reasoning and significance of the survival probabilities within these age groups, highlighting the impact of COVID-19 variants of concern. Finally, Section 6 addresses the assumptions and limitations inherent in the modeling methods. Collectively, these data provide important insights into the success of different vaccination combinations in Canada’s COVID-19 landscape.

## 2. Survival Analyses in Epidemiological Contexts

Statistical characterization of possible events is a common approach to assessing the temporal nature of imposed conditions. Survival analysis calculates estimates of the probability of occurrence for a specific event, based on previously recorded timestamps and incidence counts for that event, also known as “time-to-event” data [3]. Due to its simplicity in the required calculations and processing of data, survival analysis procedures offer versatility for an array of problems, such as predicting pipe failures or the longevity of business plans, as examples [4,5]. In medicine, survival analysis is often used to describe the probability of becoming ill or facing mortality over time, derived from collected data in which patients had previously become seriously or fatally ill [6].

Survival analysis has a potentially significant role in opportunities to monitor epidemics, where the individual impacts of an infectious disease are typically binary in nature; that is, infected (a confirmed “infected” datapoint) or not (either a confirmed “negative” case, or an unresponsive censored case) [7]. Since parameters such as contact interval or initial infection may be right-censored, survival analysis has been increasingly used by those studying the impacts and longevity of epidemics [8].

Generalized insights into the progression of infectious disease transmission are considered useful for decision-making by public health authorities [9]. Survival analysis provides the ability to draw generalized conclusions, as the estimated long-term parameters drawn from previously collected data on the virus of concern can be statistically inferred, thereby ignoring the impact of external factors such as the social behaviors of a population, co-morbidities, and age-based vulnerabilities, among others [9,10].

### 2.1. The Kaplan–Meier Method

Survival analysis models can produce errors when an outcome is reached by means that have ignored the initial conditions (e.g., a patient dies due to factors unrelated to the monitored illness) [11]. In such cases where observations are considered as incomplete, the Kaplan–Meier method of survival analysis may be implemented to generate an estimate that takes the form of a step-function containing discontinuities at the time of observed non-survival, without any assumptions regarding the distribution, thus maintaining a non-parametric approach [11].

In practice, non-survival is considered as failing to meet the desirable outcome in a defined context. Traditionally, this has typically referred to the occurrence of patient death [11]. In the context of COVID-19 testing, this could simply be a “positive” result from a tested individual. Therefore, survival would dictate that the patient was not carrying the virus: a “negative” result. This nominal categorization (a clinical “yes” or “no”) is described by Equation (1) (where  $x$  is the survival status of an individual,  $k$ ) [12].

$$x_k = \begin{cases} 1 & (\text{non-survival}) \\ 0 & (\text{survival}) \end{cases} \quad (1)$$

Hence, for an analysis surrounding COVID-19 positivity,  $x = 1$  would imply a “positive” test result, and  $x = 0$  indicates a “negative” result.

The survival probability describes the probability that a member of the population will not face the undesirable outcome at a given time. The probability at a defined time will influence the probability at a later time, since the total non-survival counts ( $q_i$ ) will increase temporally, and, therefore, the probability of survival is assumed to increase more substantially over time. This is more concisely described in Equation (2) [13], although more detailed mathematics are provided in Section 3.

$$\hat{S}(t_i) = \prod \left( 1 - \frac{q_i}{N_i} \right) \quad (2)$$

Several estimators of the variance in Kaplan–Meier survival probabilities exist, but a common and often preferred calculation method is the Greenwood formula [14]. Following this paper’s syntax, the Greenwood formula is defined in Equation (3).

$$\widehat{\text{Var}}(\hat{S}(t_i)) = \hat{S}(t_i)^2 \sum \left( \frac{q_i}{N_i(N_i - q_i)} \right) \quad (3)$$

As introduced, Kaplan–Meier curves are a collection of step-wise estimates. For this reason, the graphical presentation of the results from a Kaplan–Meier analysis implies significant context on the type of data used, and the reliability of any conclusions. If a plot includes fewer steps with larger “jumps”, it can be assumed that there were relatively few data points for that particular analysis (or a high number of censored cases), which would suggest less precision within the results [15]. Likewise, a collection of very small steps that are too numerous to count would indicate better precision and more defensible conclusions from the analysis [15].

The censoring of data in this estimator method introduces the concept of a nonparametric maximum probability, or a “missing data problem” [16,17]. This method in survival analysis allowing right-censored data is very beneficial in scenarios with truncated or interval-based censoring within datasets [16,18,19].

## 2.2. Application of the Kaplan–Meier Method

Applications of the Kaplan–Meier method in epidemiological studies have been few, despite being one of the most common approaches to modeling survival estimates [20]. Primarily, applications of the method are most common in the prediction of the probability of death with respect to time, typically in medical fields such as oncology, nephrology, and rheumatology [21–24].

Similarly, the success of vaccines and their responses within a population have not been extensively studied using an analysis with Kaplan–Meier estimates. Studies on “breakthrough infections” of COVID-19 have become a later focus of pandemic research, using Kaplan–Meier analyses to determine how likely it is that one will become infected with COVID-19 despite having already received one or more doses of a vaccine [25–27]. In three studies focusing on breakthrough infections by the end of 2021, two concluded that significant protection against infection can be declared, considering relatively low probabilities of testing positive for two to three months under Kaplan–Meier survival probabilities [25,27].

One of the three studies, however, found that *Moderna* vaccine recipients were less likely to have a breakthrough infection compared to those who received *Pfizer* [26]. The significance of this discrepancy is open to interpretation since the difference between the cumulative incidents of breakthroughs for each vaccine was reported as 12% for the third month after having received both doses [26].

Limited vaccine-related applications outside of COVID-19 include characterizing the uptake of the immunization program for vaccines preventing *Haemophilus influenzae* type b infection (the Hib vaccine) in young children and predicting the success of phase II trials for a vaccine targeting maternal cytomegalovirus (CMV) [28,29]. In the study analyzing Hib vaccine uptake, a population segment in Germany was determined to be insufficiently following the WHO-recommended immunization schedule, using a Kaplan–Meier back-calculation from positive infections to define the theoretical age at which a child would have been vaccinated [28]. A study of immediate-postpartum mothers in Tuscaloosa, Alabama, concluded from a Kaplan–Meier analysis that the vaccine under trial for CMV was at least 10% more likely to maintain non-infection for a period of 42 months [29].

In COVID-19 contexts, there have been very few studies using survival analysis and modified applications of Kaplan–Meier curves in the interpretation of various facets of the pandemic. Within the first year of global transmission, an early study used Kaplan–Meier estimators to determine recovery times from an infected population in India; the survival probability was calculated at daily intervals (based on time-to-recovery data) to allow for a determination that, at the time, there was only a 4% probability of recovering from the virus within 10 days of infection [30]. The trial process for using remdesivir to treat those hospitalized with COVID-19 also used this method of survival analysis, estimating the probability of recovery at a given time post-injection [31]. Kaplan–Meier curves have also been used to declare the significance of potential co-morbidities or illness factors in conjunction with COVID-19, such as an analysis proving that patients with diabetes were more likely to experience more severe illness if they also contracted COVID-19 [32]. Results from the study showed a survival probability over 30% less in the diabetes group than in the non-diabetes control group, following 20 days of hospitalization [32].

Applications of Kaplan–Meier survival analysis to study the success of COVID-19 vaccines are novel in pandemic-related research and are shown below to provide useful insights into the time-based benefit of immunization against the virus.

### 3. Methods: Modeling Vaccinated Ontarians

Characterizing the trends of positive testing over time is an essential tool in understanding the benefits and long-term protection implications of the various vaccines over time. A Kaplan–Meier method of survival analysis was implemented to determine the probability of testing positive for the virus over time as characterized by time since the receipt of vaccines.

Data for the model were sourced from a data package consisting of the Ontario population tested for COVID-19 from a period of March 2020 to late December 2021. These data were sourced from ICES under an agreement to conduct COVID-19 research in Canada. These datasets were linked using unique encoded identifiers and were analyzed at ICES. The legal parameters associated with privacy legislation that govern the research of ICES permit the collection and analyses of Ontario's health care data without consent, in applications aimed at safeguarding public health.

Prior to implementing the survival estimation model, the dataset underwent a preparation process to remove erroneous data. The entire population is grouped based on "dose status" indicating how many doses of any vaccine were received. Each dose status receives its independent dataframe; population with one dose dataframe (df\_dose\_1), population with two doses dataframe (df\_dose\_2), and population with three doses dataframe (df\_dose\_3). All individuals who were not vaccinated were removed from the dataset.

The data were then separated by specified vaccine type. The possible conditions considered are detailed in Table 1.

**Table 1.** Segregated vaccine types for the dataset, where a dosage regime may comprise various vaccine combinations.

Doses		Vaccines under Consideration				
1	Pfizer (1)	Moderna (1)	AstraZeneca (1)	-	-	-
2	Pfizer (2)	Moderna (2)	AstraZeneca (2)	AZ (1) + Pfizer (1)	AZ (1) + Moderna (1)	-
3	Pfizer (3)	Moderna (3)	-	-	-	-

Individuals who may have had multiple tests had their testing results merged to constitute a single data point (thereby ensuring that an individual person is not considered more than once in the analysis). Only positive and negative test results were analyzed, and statuses of any tests that were categorized as indeterminant, canceled, pending, or rejected were all removed from consideration in the dataset. The testing data consisted of 614,470 tested individuals with two doses of a COVID-19 vaccine and 49,551 tested three-dose recipients.

Two periods of analysis were chosen on the basis of the dominant strain of the virus at that time, and the population sizes were modified as follows:

- Period 1 (“Delta period”): 77,220 possible persons, 5 July 2021 to 12 August 2021 [inclusive],
- Period 2 (“Omicron period”): 21,658 possible persons, 16 November 2021 to 24 December 2021 [inclusive].

Both periods span 38 days. For the “Omicron period”, the dates were chosen based on the mean date for when the population would have received a third dose of the vaccine (30 November), backdated 14 days prior to the mean date to account for individuals who would have become “fully immunized” by 30 November. This was then extended until the last available date of data collection (24 December), which encompasses a total of 38 days. The reliability of data was compromised following 24 December 2021, when Ontario announced major modifications to the criteria and volume of COVID-19 testing [33]. The “Delta period” was adjusted to reflect the same duration as the “Omicron period” (in this case, starting 14 days prior to the mean date of the second dose).

The 14-day lag period described above was selected in alignment with findings that the majority of vaccines are effective and provide maximum protection (against COVID-19 and the Omicron variant) two weeks after receiving a second dose of the vaccine [34].

Moreover, a third dose (or “booster” dose) is recognized to provide an increase in protection following the same period of time [35]. In practice, this would mean that for a two-dose analysis, time “0” represented 14 days following receipt of the second vaccine dose, and a three-dose analysis would represent 14 days following receipt of the third dose. Mathematically, this can be considered as an injection time occurring at  $t = -14$  days, with respect to the most recent dose. This is outlined more concisely below.

Although it has been reported that re-infection for COVID-19 is possible, a degree of “immunity” is notable in previously infected cases [36]. For this reason, individuals who had previously tested positive for the virus (prior to either of the isolated timeframes) were excluded from this analysis to ensure that the individuals within the population were more similar in their ability to contract the virus.

For the “Delta period”, all individuals who tested positive before 1 July 2021 were removed from consideration in the dataset. For the “Omicron period”, all individuals who tested positive before 1 November 2021 were removed from consideration in the dataset. The only test results that were accepted for these periods are those that occurred 14 days after a person’s most recent vaccination since they otherwise would have become infected while not yet “fully immunized” from their vaccine dosage. Tests that occurred prior to 14 days post-vaccine receipt were removed from the dataset.

It should be noted that the date ranges coincide with distinct variants of concern dominating as the primary form of the virus in monitored cases. The onset of the virus’



*Omicron* variant of concern developed into the dominant strain of the virus within a very short time over *Delta* in late 2021 [37]. For this reason, the basis of conclusions for the modeling results indicates the probability of testing positive for the virus, with a high probability that the case is associated with the defined period's dominant variant of concern. While it is not assured that every case associated with these results would belong to the dominant variant, the high abundance of variant-based cases implies that this assumption is an appropriate generalization.

The duration of vaccination is then calculated for each individual:

- $t = 0$  is defined as the start of a given period (either 5 July or 16 November 2021),
- $\text{date}_{\text{vaccine}}$  is defined as the date at which a person received their last vaccine dose,
- An observed test date occurs at any given number of days following  $t = 0$  (defined as  $t_{\text{test}}$ ),
- The duration of vaccination for any given person is calculated as  $t_{\text{test}} - (\text{date}_{\text{vaccine}} + 14)$ .

Using this common timeline, a Kaplan–Meier approach was then applied to the data. “Survival” in this application of a survival analysis indicates the status of an individual as “not testing positive for the virus”. At a given time,  $t$ , an individual may report a positive test (deemed a value of 1), or negative/no longer monitored for a positive result (a value of 0). If the initial population of the entire group of analysis is described by the variable  $N_i$ , and the number of individuals testing positive at  $t$  is provided as  $q_i$ , the probability of not testing positive, or the probability of survival, at  $t$  is determined by Equation (4) [38].

$$S(t_i) = \frac{N_i - q_i}{N_i} \quad (4)$$

The probability of not testing positive until the time defined at  $t$  is evaluated by taking the product of all probabilities of “survival” from the times precursive to  $t$  [38]. This is justified by the understanding that the survival of the population at a time,  $t_i$ , is influenced by the survival from the initial time ( $i = 1$ ) until the time immediately before  $i$  (simply,  $i - 1$ ) [13]. This is presented more concisely in Equation (5) (where  $\hat{S}(t)$  represents the cumulative survival probability).

$$\hat{S}(t_i) = S(t_i) \times \hat{S}(t_{i-1}) \quad (5)$$

Alternatively, the probability of testing positive (non-survival) at a given time can be determined by Equations (6) and (7).

$$D(t_i) = 1 - S(t_i) \quad (6)$$

$$\hat{D}(t_i) = 1 - \hat{S}(t_i) \quad (7)$$

In the survival analysis for the probability of becoming a confirmed case of COVID-19, the results of interest surround the value of  $D(t)$  at various timestamps following vaccination to assess and compare the temporal benefit of individual vaccines for specific age demographics. A 95% statistical confidence interval was then applied to all estimates (i.e., the inverse Kaplan–Meier curves).

## 4. Results

The data used for integration with the Kaplan–Meier model were a collection of positive test results from individuals living in Ontario with one, two, or three doses of a COVID-19 vaccine. The vaccines summarized in Table 2 were recognized in the dataset.

In addition to the three vaccines listed, data were also scrutinized for those who had received a combination of vaccines, specifically *AstraZeneca* as a first dose with *Pfizer* or *Moderna* as a second dose. The combination of mRNA vaccines (one *Pfizer* dose with one *Moderna* dose) was not studied. Overall, the available dataset encompasses a significant number of Ontarians, with 614,470 people receiving any form of two vaccine doses.

**Table 2.** Overview of COVID-19 vaccines used within the entire population of recorded data.

Vaccine	Branded Name	Name Used in this Study	Vaccine Type	Ages Eligible for Receipt (January 2022)
Pfizer-BioNTech's COVID-19 Vaccine (BNT162b2)	Comirnaty	Pfizer	mRNA	5 and older [37]
Moderna's COVID-19 Vaccine (mRNA-1273)	Spikevax	Moderna	mRNA	12 and older [38]

The vaccination-age groups describe the persons associated with a specific receipt of vaccination, segregated by age range. The largest vaccination-age group in the dataset was the *Pfizer* (two doses) for those aged 20 to 39, inclusive. The summarized *N* values for each vaccination-age group are provided in Table 3.

**Table 3.** N values for each vaccination-age group as determined from the Ontario dataset.

Vaccine	Two Doses (Delta Period)					Three Doses (Omicron Period)				
	All Ages	Under 20	20 to 39	40 to 64	65 and Older	All Ages	Under 20	20 to 39	40 to 64	65 and Older
Pfizer	59,061	14,583	18,660	16,537	9281	19,213	-	6500	7643	4843
Moderna	18,159	1188	7936	6079	2956	2445	-	405	855	1176

Due to the limited population size for those receiving *AstraZeneca* (either in full or in combination with an mRNA vaccine), the results were restricted to only *Pfizer* and *Moderna* recipients, which constitute a significant and considerable population of Ontarians for justifiable modeling.

The application of the Kaplan–Meier method for the vaccination-age groups is presented graphically. The probability of testing positive is depicted as a function of the elapsed time subsequent to the administration of different vaccine regimens, comprising either two or three doses, accordingly.

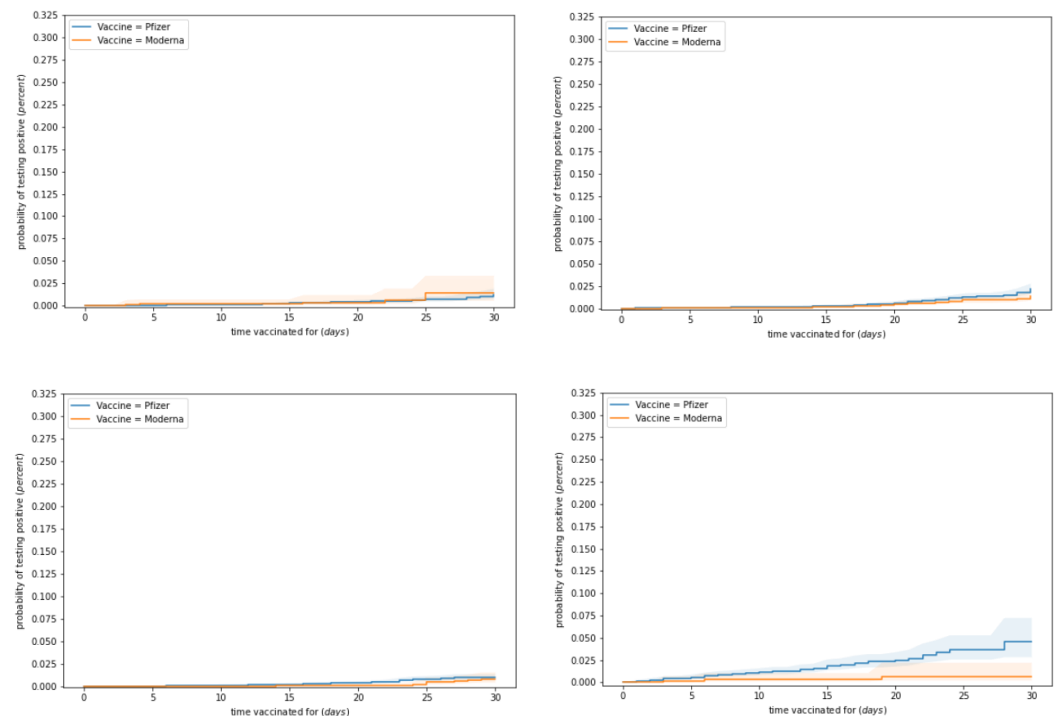
Figure 1a,b illustrate the probability of testing positive for the virus for all ages during the *Delta* and *Omicron* periods, respectively, for each of the vaccines and vaccine combinations. Note that the scale for Figure 1a is an exaggerated scale compared to that in Figure 1b, such that the visual presentation displays the marginal difference between either vaccine, which is less prominent in the *Delta* figure when compared to the *Omicron* period (Figure 1b). Although the difference in the curves is exaggerated in Figure 1a, this margin remains less than 1% for the duration of the plot.

When separated into individual age groups, the trends for the two-dose probability of testing positive during the *Delta* Period were similar across all curves and aligned with the results presented in the aggregated curve (Figure 1a). However, the “under 20” age group showed the only difference, with *Moderna* recipients showing a slightly higher probability of testing positive compared to those who received *Moderna*, following 25 days post-full effectiveness of two-dose vaccine protection. Despite this variance, all results indicate less than a 3% probability of testing positive at day 30 of either period, regardless of vaccine type. The two-dose plots for each vaccine-age group during the *Delta* Period are presented in Figure 2. Again, note that this scale is different from the plot for *Omicron* cases, for ease of interpretability.

With the exception of a span of 5 days in the “under 20” curve, *Pfizer* recipients show a slightly higher probability of testing positive when compared to those who received *Moderna* for both doses during this period although it is worth noting that this margin of difference is acutely small (less than 1% difference in probability), and it has been reported that a 1% to 2% difference in effectiveness is not recognized as sufficiently significant to discriminate between vaccines [39,40].



**Figure 1.** Comparison of COVID-19 infection rates for different vaccination regimes: (a) Two doses during the *Delta* period, and (b) Three doses during the *Omicron* period.

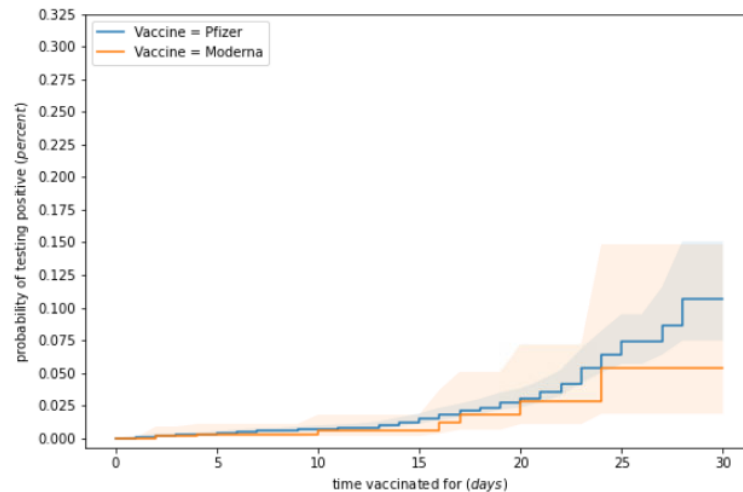


**Figure 2.** Individual vaccination-age groups with two doses of *Pfizer* and *Moderna* (*Delta* Period), exaggerated vertical axis. (**Top left**): ages under 20. (**Top right**): ages 20 to 39, inclusive. (**Bottom left**): ages 40 to 64, inclusive. (**Bottom right**): ages 65 and older.

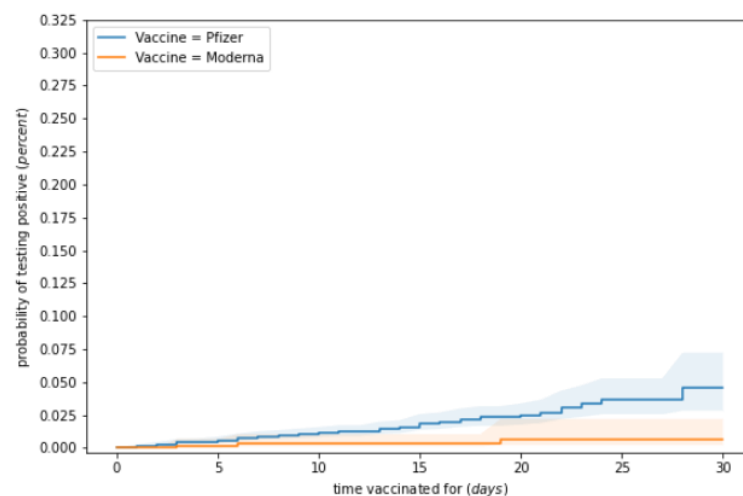
Contrasting with the marginally different curves during the *Delta* Period, the curves associated with *Omicron* display more discrepancy between the two vaccines, and overall higher probabilities of testing positive (presented in Figure 1b). In all cases, the Kaplan–Meier estimates produce multi-day periods with no change in the  $y$ -values of the curve, which indicates a zero-positivity rate during that period of time. This is accompanied by wider confidence intervals, exacerbated when compared to the marginal confidence intervals seen for the plots associated with two vaccine doses. The results associated with these charts indicate that the reduced size of the population in this period (due to the timing of the rollout for third doses) influenced the estimates computed using the Kaplan–Meier method.



Figures 3 and 4 represent the *Omicron* Period, three-dose probability curves for the age categories from 40 to 64 (inclusive) and 65 and older age groups, respectively. In both cases, *Pfizer* recipients maintained a slightly higher probability of testing positive for COVID-19 when compared to those who received *Moderna* for all three doses. Despite not showing the same low probabilities seen in the two-dose figures, both age groups produced significantly high survival probabilities (low probabilities of testing positive), with the 65 and older group remaining less than 5% in positivity probability for the 30-day period shown.



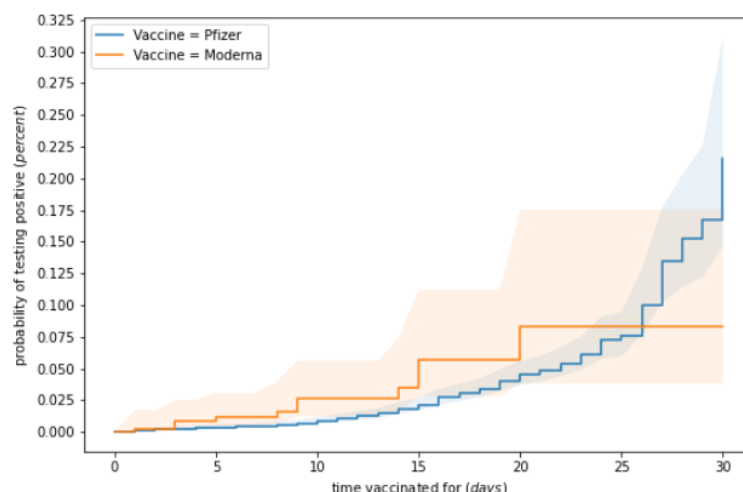
**Figure 3.** Ages 40 to 64 (inclusive) with three doses of *Pfizer* and *Moderna* (*Omicron* Period).



**Figure 4.** Ages 65 and older with three doses of *Pfizer* and *Moderna* (*Omicron* Period).

For those aged 20 to 39 (inclusive), Figure 5 differs slightly from these two figures, showing *Pfizer* and *Moderna* curves that differ over time; *Moderna* recipients show a higher probability of testing positive for approximately 25 days, and *Pfizer* recipients overtake the higher probabilities after this timestamp. It is important to note that this curve presents wider confidence intervals, and this contrast to other trends is likely attributable to the smaller population of young adults receiving a third dose of any vaccine by the end of 2021.

Due to the negligible population size of those under the age of 20 receiving three doses of any vaccine in 2021, this age group is not assessed for the *Omicron* Period analysis.



**Figure 5.** Ages 20 to 39 (inclusive) with three doses of *Pfizer* and *Moderna* (Omicron Period).

## 5. Discussion

In most of the figures described above, the results show that *Moderna* provides the most successful probability against testing positive for the virus, for all age groups, for the longest period of time. This result was consistent among the groups that had received two vaccine doses, as well as those who had received a third dose. Although these results favor the protection provided by the *Moderna* vaccine over the other mRNA-type vaccine, *Pfizer*, it is important to note that this difference is essentially negligible. In all age groups, a three-dose regime for either mRNA class of immunization, during a wave of *Omicron* cases, results in comparably low probabilities of testing positive for the virus (and, therefore, a relatively high survival probability) for extended periods of time: in many cases, less than 10% positivity probability past 30 days since the full effectiveness of a recipient's third dose.

It has been reported that two doses of *Pfizer* alone are associated with a vaccine effectiveness ('VE') of 95% [41]. Although the results of the survival analysis undertaken in this study do not imply or prove the efficacy of a vaccine, they do support the general findings associated with the benefits of the vaccines. That is, in all studied populations receiving two doses of *Pfizer* during a wave of *Delta* cases, a calculated probability of testing positive within a 30-day period never exceeded 5% (or, reversely, survival/not testing positive never fell below 95%).

The survival curves for both vaccines following two doses aggregated for all ages proved that, despite hierarchical differences in probability for testing positive, the difference between the "most successful" and "least successful" vaccines remains negligible (<1%) for at least a period of a full month during a wave of cases (following full effectiveness of a recipient's second dose). For this same period, both vaccines also remain significantly below the 5% probability of testing positive, regardless of age. These results indicate that the vaccines are highly beneficial in avoiding testing positive for the virus, aligning with reports that each of the available vaccines is highly successful at protecting the population and mitigating severe illness [42].

Despite the common trends in the three-dose plots, where *Pfizer* is commonly noted to result in a higher probability for testing positive than *Moderna*, the curves are accompanied by wider confidence intervals than those associated with two-doses, and atypical behaviors (steeper curves, and large jumps in difference between each curve) are evident in each plot. An important context provides insight into this inconsistency, primarily relating to vaccine rollout in these age groups. In late 2021, when *Omicron* variant cases were spiking in Canada, Canadians were only just beginning to receive three doses of the vaccine. For this reason (and as seen in Table 3), the population sizes for these vaccination-age groups are smaller than the matching groups from the summer period, and results are, therefore, more likely to be consistent with a broader range of conclusions. Vaccine rollout in younger

populations was modified further in 2021 to default (where possible) to the injection of *Pfizer* rather than *Moderna*, due to reports of myocarditis as a side effect in those under the age of 30 occurring more commonly in those who receive *Moderna* [43]. Hence, it is perhaps attributable to this preferential distribution of the *Pfizer* vaccine that results in a larger population size for those under the age of 30 and, therefore, a wider confidence interval on the *Moderna* curve for these plots (Figures 3–5).

### 5.1. The Role of Variants of Concern

Early publications on the onset of the *Omicron* variant have reported that those with three doses of a COVID-19 vaccine are less likely to constitute symptomatic infection groups, despite evolving mutations of the original strain [44]. Additionally, as previously noted, the data used for this survival analysis were narrowed to two timeframes, the second being late 2021, which would have been subject to an increasingly significant number of *Omicron*-specific cases in combination with *Delta* cases. Due to the higher transmissibility of *Omicron* when compared to the previous dominant strain, *Delta* [45], the results of the survival analyses are subject to variability.

In survival analysis, a strictly *Omicron* infection population is predicted to show higher probabilities of testing positive when compared to the *Delta* population, since more of the population will contract the virus, exhibit symptoms, and pursue testing. Although infection may cause equal or more severe illness in those with the *Delta* variant, the less transmissible behavior of the *Delta* variant leads to lower probabilities of testing positive in an exclusively *Delta* dataset.

The mixing of the two variants during the study period would, therefore, favor higher positivity rates in any group that happens to have more cases of the newer variant. Since data were not explicitly segregated by variant type in the dataset during the analyses, it is assumed that disproportionate totals of each variant in vaccination-age groups are possible, and a likely cause for minor discrepancies. To minimize this issue, the dates selected for analyses were chosen based on the dominant strain of the virus at that time, when caseloads were high.

### 5.2. Validation

As a simple validation of the use of Kaplan–Meier for this study, a log-rank test for each age group under the two variant time periods was conducted, and 5% and 1% significance levels were considered. For the tests, smaller significance levels suggest a more stringent analysis. A *p*-value that is less than a given significance level would indicate that there is a statistically significant difference in the survival analysis curves under consideration. This is summarized in Table 4.

**Table 4.** Log-rank *p*-values for different age groups and periods.

Age Group	Delta Period	Omicron Period
Under 20	0.47	–
20 to 39	0.07	0.04
40 to 64	0.06	0.30
65 and Older	0.07	0.005

For the analysis of “all ages”, the *Delta* log-rank test returned a *p*-value of 0.02, and 0.005 for *Omicron*. This suggests a statistically significant difference at the 1% level for *Omicron*, and the 5% level for both *Delta* and *Omicron*.

For the under-20 age group, *Delta* results had an associated *p*-value of 0.47 (with no applicable analysis for *Omicron*). Since this value is greater than both 5% and 1% significance levels, there were no statistically significant differences found for this aspect of the analyses.

In adults aged 20 to 39,  $p$ -values from the log-rank testing were 0.07 and 0.04 for *Delta* and *Omicron*, respectively. Again, these values suggest that the analyses had no statistically significant differences for this age group, with the exception of the 5% level for the *Omicron* period.

Log-rank  $p$ -values of 0.06 and 0.30 for *Delta* and *Omicron* were evaluated for the analyses associated with the 40 to 64 age category. For both considered significance levels, in either variant time period, it can be concluded that there were no statistically significant differences found.

Finally, the 65 and older age group only had a statistically significant difference at the 1% and 5% levels for *Omicron*, having associated  $p$ -values of 0.07 and 0.005 for *Delta* and *Omicron*, respectively.

The statistically significant differences between the populations and their estimated probabilities constitute a minority of the log-rank results, which broadly suggests a relatively stringent analysis and appropriate use of Kaplan–Meier with the available dataset.

## 6. Limitations of the Study

Several assumptions and simplifications within the dataset and the use of the Kaplan–Meier model for analyses in this study are directly linked to any discrepancy in data presentation and cases of low accuracy for select age group results. As introduced in the description of the three-dose plots, inconsistent curve shapes as compared to the similarly shaped curves in the two-dose plots are noted. Primarily, the three curves contain several extended time periods with no change in the value on the  $y$ -axis, therefore implying that there would be no greater probability of testing positive on one day following full vaccine protection when compared to several days later. These “flatline” sections are inconsistent with expected analysis results and are presented with significantly wide confidence intervals (*Moderna* showing wider confidence intervals and longer periods of flat behavior than *Pfizer*, although the low accuracy is noted for both curves).

Lower accuracy in three doses for individuals aged 20 to 39 is largely attributable to limited population size. Whereas the total three-dose *Moderna* population for those 65 and older was 11,713, the 20 to 39 group only had 527 reported recipients of three *Moderna* doses. This significant difference in total populations would suggest that Kaplan–Meier calculations for the smaller group could lead to longer periods of time between reported positive cases. This consequently leads to prolonged intervals where the function  $S(t)$  value does not change, since the number of individuals testing positive  $q(t)$  may persist at zero for several days in succession. Hence, a difference in sample size is a limitation of the model, which may impact the confidence in curve behavior. However, these differences do not inhibit the drawing of generalized conclusions from the model results.

Although the plots are accompanied by wide confidence interval bands, the generalized presentation of the results is still considered valid due to consistent alignment with the findings in other, more attuned Kaplan–Meier curves (i.e., the differences between *Pfizer* and *Moderna* are consistent among all age groups).

A key assumption in this application of survival analysis is that the only factor impacting positivity rates in the pandemic is vaccination—this approach neglects those factors which may cause increases in positivity. This is to say, populations are assumed to be equally interacting and behaving with no differences among age groups—all individuals, irrespective of age or social status, are anticipated to experience comparable levels of exposure to the virus. This assumption is categorically false, although necessary to employ.

It has been reported that in many applications of Kaplan–Meier models, independence among cohorts, constant conditions, and equivalent non-survival risks for all cohorts constitute required modeling assumptions to reflect the statistical nature of the particular model [46]. While the assumptions made may not be considered fully realistic in all scenarios, the disclosure of these assumptions as indicated herein provides an avenue of interpretation that is conscientious of the limitations of the model and the variability among members of a study population [46]. The unknown aspect of human behavior,

which impacts results such as those presented in survival analysis, can be considered a “social factor”.

In this study, the social factor is considered as the reason for the oldest age category having the lowest probabilities of testing positive, despite scientific knowledge that this age group is less immunologically responsive to a dose of vaccine [47]. Since biologically, the vaccine could work less effectively in those in the 65 and older demographic, the Kaplan–Meier curves showing low positive testing probabilities are likely a result of the social behaviors of this group, due to the cautious nature of those who are more aged and vulnerable to severe disease, and the distancing of families from older generations during the pandemic to protect the health of loved ones. It can be assumed that the social factor protects this age group from excessive exposure, leading to fewer positive test data points for computation in the survival analysis. Furthermore, the transient nature of younger groups, including the working population, would cause increased exposure to the virus, and, therefore, a higher number of positive cases, despite a potentially greater immunity gain from the vaccines. Although social factors may differ significantly between age groups, this trend is limited when comparing results within age groups.

Restrictions in the available data limited this study in comparing breakthrough infections with the un-vaccinated Ontario population and single-dose recipients. The two periods studied (*Delta* and *Omicron* peak periods) were matched intentionally with study populations having received two and three vaccine doses, respectively, for two primary reasons: to accommodate gaps in patient and testing data, and to accurately depict the impacts of the vaccines as inoculation actively occurs in tandem with a wave of the pandemic. Since most of the population would have received a second dose several months before the *Omicron* study period, only the three-dose regimen was considered in that period to better reflect the benefits of the vaccine as it is being introduced during the ongoing wave of infection.

While the application of the Kaplan–Meier model in this study does not expressly re-design or innovate the model itself, a key success of this application has been continuing to build consensus in statistical modeling while using the proven model in novel contexts. Kaplan–Meier and survival analysis in the areas of COVID-19 breakthrough infections and vaccine-related virus positivity for the province of Ontario is an untouched realm of pandemic research. Although a standard and proven model is used in the study, the population that was modeled and the relative imminence of further COVID-19 threats allow for the introduction of a new Canadian perspective on the pandemic. Hence, this report on Ontario’s temporal response and impacts of COVID-19 vaccines and waves of infection delivers notable new insights.

## 7. Conclusions

A survival analysis approach to modeling and predicting vaccine success in the COVID-19 pandemic is a virtually unexplored field, utilizing statistical estimation techniques with widespread discussions of the monumental COVID-19 pandemic. Kaplan–Meier estimates of the probability for testing positive show that marginal differences put multiple doses of the *Moderna* vaccine across all ages at a slight advantage to those who received *Pfizer*—maintaining this slight difference over time (in most cases, at least for upwards of 30 days, and predicted to continue for longer periods). Those individuals 65 and older returned a small confidence interval result, showing an extended period of time with negligible difference between the two vaccines, with less than 5% probability of testing positive for the virus within a month post-third dose. At the same timestamp, the 20 to 39 (inclusive) age group presented the highest probability of testing positive; however, the limited population size of this group caused wider confidence intervals. Although generally, the findings for this age category are indicative of higher positivity trends, the wider confidence interval suggests slightly less precision in the computed results.

Finally, for all ages following two doses of either vaccine, there is a significantly low probability of testing positive for as many as 30 days post-vaccination. The highest



probabilities of positivity resided with those who received *Pfizer* compared to the lowest probabilities associated with *Moderna* recipients.

The results of this survival analysis approach to vaccine characterization support the general findings and directions of previous immunological studies, as well as broader public health messaging that vaccines are successful, and although differences in each immunization exist, their overall variances are negligible. As the Food and Drug Administration in the United States set forth plans for the fourth doses of vaccines in fall 2022 [48], the modeling-based insights into the vaccines support that additional doses during times of peaking infection can substantially minimize risks of becoming infected (and infectious). A probability of testing positive reaching values in excess of 20% after 38 days of full effectiveness from a vaccine (Figure 3) indicates that the fall months function as a strategic time for the rollout of additional inoculation to curb further spread of the virus. Survival analysis is a practical approach to deriving generalized plots of post-vaccine probability of infection. It is evident that higher uptake in the vaccination program, and multiple doses, can help reduce the probability that individuals will experience COVID-19 symptoms, seek testing, and receive a positive result. On this trajectory, fewer seriously ill cases can keep the threat of the virus at bay and safeguard public health and safety.

**Author Contributions:** E.A.M., W.H., B.S., J.Y. and S.A.G.: conceptualization. W.H.: development/execution of methodology and software. M.G.C.: first draft of manuscript, development of subsequent drafts, initial results interpretation. All authors: validation. M.G.C., E.A.M. and W.H.: formal analysis. All authors: writing, review, and editing. E.A.M.: supervision, project administration, and funding acquisition. All authors have read and agreed to the published version of the manuscript.

**Funding:** This research was funded by the Natural Sciences and Engineering Research Council of Canada (NSERC), funding reference number 400677. Additionally, funding from the NSERC Alliance COVID-19 fund, reference numbers 401636 and 401641, is also recognized.

**Institutional Review Board Statement:** Not applicable

**Informed Consent Statement:** Patient consent was waived. The data utilized were sourced from ICES under an agreement to conduct COVID-19 research in Canada. The legal framework governing ICES research activities permits the collection and analysis of Ontario's health care data without individual consent.

**Data Availability Statement:** The data analyzed in this study are subject to the following licenses/restrictions: this study was conducted using data sourced from ICES, which is funded by an annual grant from the Ontario Ministry of Health (MOH) and the Ministry of Long-Term Care (MLTC). Data access during the COVID-19 pandemic is overseen through the Ontario Health Data Platform (OHDP), a Province of Ontario initiative to address Ontario's ongoing response to the pandemic and its related impacts. Requests to access these datasets should be directed to <https://www.ices.on.ca/DAS/Public-Sector/Access-to-ICESData-Process> (accessed on 20 July 2024).

**Acknowledgments:** The University of Guelph Leadership Chair program is also gratefully acknowledged. ICES, OHDP, and the staff at these organizations are recognized and appreciated for their data and guidance in the development of the research findings.

**Conflicts of Interest:** J.Y. was employed by Adastra Corporation. The remaining authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

## References

1. Mahase, E. Vaccinating the UK: How the covid vaccine was approved, and other questions answered. *BMJ* **2020**, *371*, m4759. [CrossRef] [PubMed]
2. Vilches, T.N.; Zhang, K.; Van Exan, R.; Langley, J.M.; Moghadas, S.M. Projecting the impact of a two-dose COVID-19 vaccination campaign in Ontario, Canada. *Vaccine* **2021**, *39*, 2360–2365. [CrossRef] [PubMed]
3. Chung, C.F.; Schmidt, P.; Witte, A.D. Survival analysis: A survey. *J. Quant. Criminol.* **1991**, *7*, 59–98. [CrossRef]
4. Snider, B.; McBean, E.A. Improving Urban Water Security through Pipe-Break Prediction Models: Machine Learning or Survival Analysis. *J. Environ. Eng.* **2020**, *146*, 04019129. [CrossRef]

5. Parsa, H.G.; Self, J.; Sydnor-Busso, S.; Yoon, H.J. Why Restaurants Fail? Part II—The Impact of Affiliation, Location, and Size on Restaurant Failures: Results from a Survival Analysis. *J. Foodserv. Bus. Res.* **2011**, *14*, 360–379. [\[CrossRef\]](#)
6. Rabbani, M.A.; Habib, H.B.; Islam, M.; Ahmad, B.; Majid, S.; Saeed, W.; Shah, S.M.A.; Ahmad, A. Survival analysis and prognostic indicators of systemic lupus erythematosus in Pakistani patients. *Lupus* **2009**, *18*, 848–855. [\[CrossRef\]](#)
7. Cole, S.R.; Hudgens, M.G. Survival analysis in infectious disease research: Describing events in time. *AIDS* **2010**, *24*, 2423–2431. [\[CrossRef\]](#)
8. Kenah, E. Contact intervals, survival analysis of epidemic data, and estimation of  $R_0$ . *Biostatistics* **2011**, *12*, 548–566. [\[CrossRef\]](#)
9. Kenah, E.; Britton, T.; Halloran, M.E.; Longini, I.M. Molecular Infectious Disease Epidemiology: Survival Analysis and Algorithms Linking Phylogenies to Transmission Trees. *PLoS Comput. Biol.* **2016**, *12*, e1004869. [\[CrossRef\]](#)
10. Henderson, R.; Jones, M.; Stare, J. Accuracy of point predictions in survival analysis. *Stat. Med.* **2001**, *20*, 3083–3096. [\[CrossRef\]](#)
11. Kaplan, E.L.; Meier, P. Nonparametric Estimation from Incomplete Observations. *J. Am. Stat. Assoc.* **1958**, *53*, 457–481. [\[CrossRef\]](#)
12. Motakis, E.; Ivshina, A.; Kuznetsov, V. Data-driven approach to predict survival of cancer patients. *IEEE Eng. Med. Biol. Mag.* **2009**, *28*, 58–66. [\[CrossRef\]](#) [\[PubMed\]](#)
13. Scanniello, G. Source code survival with the Kaplan Meier. In Proceedings of the 2011 27th IEEE International Conference on Software Maintenance (ICSM), Williamsburg, VA, USA, 25–30 September 2011; pp. 524–527. [\[CrossRef\]](#)
14. Tang, Y. Some new confidence intervals for Kaplan-Meier based estimators from one and two sample survival data. *Stat. Med.* **2021**, *40*, 4961–4976. [\[CrossRef\]](#) [\[PubMed\]](#)
15. Rich, J.T.; Neely, J.G.; Paniello, R.C.; Voelker, C.C.J.; Nussenbaum, B.; Wang, E.W. A practical guide to understanding Kaplan-Meier curves. *Otolaryngol. Neck Surgery Off. J. Am. Acad. Otolaryngol.-Head Neck Surg.* **2010**, *143*, 331–336. [\[CrossRef\]](#) [\[PubMed\]](#) [\[PubMed Central\]](#)
16. Satten, G.A.; Datta, S. The Kaplan–Meier Estimator as an Inverse-Probability-of-Censoring Weighted Average. *Am. Stat.* **2001**, *55*, 207–210. [\[CrossRef\]](#)
17. Kalbfleisch, J.D.; Prentice, R.L. *The Statistical Analysis of Failure Time Data: Kalbfleisch/The Statistical*; Wiley Series in Probability and Statistics; John Wiley & Sons, Inc.: Hoboken, NJ, USA, 2002. [\[CrossRef\]](#)
18. Woodroffe, M. Estimating a Distribution Function with Truncated Data. *Ann. Stat.* **1985**, *13*, 163–177. [\[CrossRef\]](#)
19. Turnbull, B.W. The Empirical Distribution Function with Arbitrarily Grouped, Censored and Truncated Data. *J. R. Stat. Soc. Ser. B (Methodol.)* **1976**, *38*, 290–295. [\[CrossRef\]](#)
20. Jager, K.J.; van Dijk, P.C.; Zoccali, C.; Dekker, F.W. The analysis of survival data: The Kaplan–Meier method. *Kidney Int.* **2008**, *74*, 560–565. [\[CrossRef\]](#)
21. Campigotto, F.; Weller, E. Impact of informative censoring on the Kaplan-Meier estimate of progression-free survival in phase II clinical trials. *J. Clin. Oncol. Off. J. Am. Soc. Clin. Oncol.* **2014**, *32*, 27. [\[CrossRef\]](#)
22. Shen, Y.; Cai, J. Maximum of the Weighted Kaplan-Meier Tests with Application to Cancer Prevention and Screening Trials. *Biometrics. J. Int. Biom. Soc.* **2001**, *57*, 837–843. [\[CrossRef\]](#)
23. Stel, V.S.; Dekker, F.W.; Tripepi, G.; Zoccali, C.; Jager, K.J. Survival Analysis I: The Kaplan-Meier Method. *Nephron Clin. Pract.* **2011**, *119*, c83–c88. [\[CrossRef\]](#) [\[PubMed\]](#)
24. Ruiz-Irastorza, G.; Egurbide, M.V.; Pijoan, J.I.; Garmendia, M.; Villar, I.; Martinez-Berriotxo, A.; Erdozain, J.G.; Aguirre, C. Effect of antimalarials on thrombosis and survival in patients with systemic lupus erythematosus. *Lupus* **2006**, *15*, 577–583. [\[CrossRef\]](#) [\[PubMed\]](#)
25. Mizrahi, B.; Lotan, R.; Kalkstein, N.; Peretz, A.; Perez, G.; Ben-Tov, A.; Chodick, G.; Gazit, S.; Patalon, T. Correlation of SARS-CoV-2-breakthrough infections to time-from-vaccine. *Nat. Commun.* **2021**, *12*, 6379. [\[CrossRef\]](#) [\[PubMed\]](#)
26. Abu-Raddad, L.J.; Chemaitelly, H.; Ayoub, H.H.; Tang, P.; Hasan, M.R.; Coyle, P.; Yassine, H.M.; Benslimane, F.M.; Al-Khatib, H.A.; Al-Kanaani, Z.; et al. Protection offered by mRNA-1273 versus BNT162b2 vaccines against SARS-CoV-2 infection and severe COVID-19 in Qatar. *medRxiv* **2021**. [\[CrossRef\]](#)
27. Zheutlin, A.; Ott, M.; Sun, R.; Zemlianskaia, N.; Rubel, M.; Hayden, J.; Neri, B.; Kamath, T.; Khan, N.; Schneeweiss, S.; et al. Durability of Protection against COVID-19 Breakthrough Infections and Severe Disease by Vaccines in the United States. *medRxiv* **2022**. [\[CrossRef\]](#)
28. Laubereau, B.; Hermann, M.; Schmitt, H.J.; Weil, J.; Von Kries, R. Detection of delayed vaccinations: A new approach to visualize vaccine uptake. *Epidemiol. Infect.* **2002**, *128*, 185–192. [\[CrossRef\]](#)
29. Pass, R.F.; Zhang, C.; Evans, A.; Simpson, T.; Andrews, W.; Huang, M.L.; Corey, L.; Hill, J.; Davis, E.; Flanagan, C.; et al. Vaccine Prevention of Maternal Cytomegalovirus Infection. *N. Engl. J. Med.* **2009**, *360*, 1191–1199. [\[CrossRef\]](#)
30. Barman, M.P.; Rahman, T.; Bora, K.; Borgohain, C. COVID-19 pandemic and its recovery time of patients in India: A pilot study. *Diabetes Metab. Syndr. Clin. Res. Rev.* **2020**, *14*, 1205–1211. [\[CrossRef\]](#)
31. Beigel, J.H.; Tomashek, K.M.; Dodd, L.E.; Mehta, A.K.; Zingman, B.S.; Kalil, A.C.; Hohmann, E.; Chu, H.Y.; Luetkemeyer, A.; Kline, S.; et al. Remdesivir for the Treatment of COVID-19—Final Report. *N. Engl. J. Med.* **2020**, *383*, 1813–1826. [\[CrossRef\]](#)
32. Yan, Y.; Yang, Y.; Wang, F.; Ren, H.; Zhang, S.; Shi, X.; Yu, X.; Dong, K. Clinical characteristics and outcomes of patients with severe covid-19 with diabetes. *BMJ Open Diabetes Res. Care* **2020**, *8*, e001343. [\[CrossRef\]](#)
33. Vasquez-Peddie, A. Omicron: Some Provinces Face COVID-19 Test Backlogs. 2021. Available online: <https://www.ctvnews.ca/health/coronavirus/some-provinces-face-covid-19-pcr-testing-backlogs-amid-omicron-surge-1.5721812> (accessed on 20 July 2024).

34. Tavailani, A.; Abbasi, E.; Kian Ara, F.; Darini, A.; Asefy, Z. COVID-19 vaccines: Current evidence and considerations. *Metab. Open* **2021**, *12*, 100124. [[CrossRef](#)] [[PubMed](#)]
35. Andrews, N.; Stowe, J.; Kirsebom, F.; Toffa, S.; Rickeard, T.; Gallagher, E.; Gower, C.; Kall, M.; Groves, N.; O'Connell, A.M.; et al. Effectiveness of COVID-19 vaccines against the Omicron (B.1.1.529) variant of concern. *medRxiv* **2021**. [[CrossRef](#)]
36. Fontanet, A.; Cauchemez, S. COVID-19 herd immunity: Where are we? *Nat. Rev. Immunol.* **2020**, *20*, 583–584. [[CrossRef](#)] [[PubMed](#)]
37. Vogel, L. An early look at Omicron. *Can. Med Assoc. J.* **2022**, *194*, E58. [[CrossRef](#)] [[PubMed](#)]
38. Goel, M.K.; Khanna, P.; Kishore, J. Understanding survival analysis: Kaplan-Meier estimate. *Int. J. Ayurveda Res.* **2010**, *1*. [[CrossRef](#)] [[PubMed](#)]
39. Self, W.H.; Tenforde, M.W.; Rhoads, J.P.; Gaglani, M.; Ginde, A.A.; Douin, D.J.; Olson, S.M.; Talbot, H.K.; Casey, J.D.; Mohr, N.M.; et al. Comparative Effectiveness of Moderna, Pfizer-BioNTech, and Janssen (Johnson & Johnson) Vaccines in Preventing COVID-19 Hospitalizations among Adults without Immunocompromising Conditions—United States, March–August 2021. *MMWR. Morb. Mortal. Wkly. Rep.* **2021**, *70*, 1337–1343. [[CrossRef](#)]
40. Puranik, A.; Lenehan, P.J.; Silvert, E.; Niesen, M.J.; Corchado-Garcia, J.; O'Horo, J.C.; Virk, A.; Swift, M.D.; Halamka, J.; Badley, A.D.; et al. Comparison of two highly-effective mRNA vaccines for COVID-19 during periods of Alpha and Delta variant prevalence. *medRxiv* **2021**. [[CrossRef](#)]
41. Mahase, E. COVID-19: Pfizer vaccine efficacy was 52% after first dose and 95% after second dose, paper shows. *BMJ* **2020**, *371*, m4826. [[CrossRef](#)]
42. Rashedi, R.; Samieefar, N.; Masoumi, N.; Mohseni, S.; Rezaei, N. COVID-19 vaccines Mix-and-match: The concept, the efficacy and the doubts. *J. Med. Virol.* **2021**, *94*, 1294–1299. [[CrossRef](#)]
43. Goldman, R.D. Myocarditis and pericarditis after COVID-19 messenger RNA vaccines. *Can. Fam. Physician Med. Fam. Can.* **2022**, *68*, 17–18. [[CrossRef](#)] [[PubMed](#)]
44. Accorsi, E.K.; Britton, A.; Fleming-Dutra, K.E.; Smith, Z.R.; Shang, N.; Derado, G.; Miller, J.; Schrag, S.J.; Verani, J.R. Association Between 3 Doses of mRNA COVID-19 Vaccine and Symptomatic Infection Caused by the SARS-CoV-2 Omicron and Delta Variants. *JAMA* **2022**, *327*, 639–651. [[CrossRef](#)] [[PubMed](#)]
45. Kumar, S.; Thambiraja, T.S.; Karuppanan, K.; Subramaniam, G. Omicron and Delta variant of SARS-CoV-2: A comparative computational study of spike protein. *J. Med. Virol.* **2021**, *94*, 1641–1649. [[CrossRef](#)] [[PubMed](#)]
46. Sjölander, A. A Cautionary Note on Extended Kaplan–Meier Curves for Time-varying Covariates. *Epidemiology* **2020**, *31*, 517–522. [[CrossRef](#)] [[PubMed](#)]
47. Brockman, M.A.; Mwimanzi, F.; Lapointe, H.R.; Sang, Y.; Agafitei, O.; Cheung, P.K.; Ennis, S.; Ng, K.; Basra, S.; Lim, L.Y.; et al. Reduced Magnitude and Durability of Humoral Immune Responses to COVID-19 mRNA Vaccines Among Older Adults. *J. Infect. Dis.* **2021**, *225*, 1129–1140. [[CrossRef](#)]
48. Howard, J. *A Fourth COVID-19 Shot Might be Recommended This Fall, as Officials 'Continually' Look at Emerging Data*; CNN Health: Atlanta, GA, USA, 2022.

**Disclaimer/Publisher's Note:** The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.