Journal of Neonatal Surgery
ISSN(Online): 2226-0439
Vol. 14, Issue 32s (2025)
https://www.jneonatalsurg.com

OPEN ACCESS

Weibull-Weighted Lindley Distribution: Modelling of Heterogeneous Survival Patterns in COVID-19 Data

Arvind Pandey[*1], Annu Chauhan[2], Ravindra Pratap Singh[3]

[*1,2,3]Department of Statistics, Central University of Rajasthan

Email ID: chauhanannu2428@gmail.com, Email ID: stats.rpsingh@gmail.com

[*]Corresponding author:

Arvind Pandey

Email ID: 9arvindpandey5@gmail.com

Cite this paper as: Arvind Pandey, Annu Chauhan, Ravindra Pratap Singh, (2025) Weibull-Weighted Lindley Distribution: Modelling of Heterogeneous Survival Patterns in COVID-19 Data. *Journal of Neonatal Surgery*, 14 (32s), 4478-4490.

ABSTRACT

This paper introduces the Weibull-Weighted Lindley (WWL) distribution, a new three-parameter frailty-based model for time-to-event data characterized by unobserved heterogeneity. The WWL distribution arises by compounding the Weibull distribution (as the baseline survival model) with a Weighted Lindley distribution (as the frailty component), resulting in a highly flexible distribution that 6can accommodate increasing, decreasing, and hazard rates. We derive key properties of the model, including its probability density function, cumulative distribution function, survival function, and hazard function. Multiple estimation methods are explored. A comprehensive simulation study is conducted to evaluate the performance of these estimators. The model's utility is demonstrated through its application to real-world COVID-19 survival datasets. In all cases, the WWL model provides an excellent fit compared to competing lifetime models, as measured by standard information criteria and goodness-of-fit tests. These findings highlight the WWL distribution as a powerful and adaptable tool for modeling complex survival data with latent heterogeneity.

Keywords: *COVID-19 data, Hazard function, Frailty models, Lifetime distributions, Parameter estimation, Time-to-event data, Survival analysis, Weibull distribution, Weighted Lindley distribution.*

1. INTRODUCTION

Modeling time-to-event (survival) data plays a fundamental role in diverse disciplines such as medicine, public health, engineering reliability, demography, economics, and the social sciences. In the medical context especially, survival models help estimate how long a patient might live after treatment, how long it takes for a disease to relapse, or how different risk factors affect mortality. The recent COVID-19 pandemic has underscored the value of such modeling: understanding variations in patient survival can inform resource allocation, treatment strategies, and public health decisions. Traditional parametric models like the exponential, Weibull, and gamma distributions have long served as the backbone of survival analysis due to their mathematical simplicity and interpretability Kleinbaum and Klein [2005], Lawless [2003]. These models assume that all differences in survival times can be explained by observed co-variates, such as age, gender, or treatment group. However, real-world data, especially that arise during health crises such as pandemics, often exhibit a more complex structure. Two individuals with similar medical histories may respond very differently to the same infection due to factors we cannot directly observe, such as immune system variability, genetic predispositions, environmental exposures, or prior asymptomatic infections. This unobserved variability is known as frailty or unobserved heterogeneity Duchateau and Janssen [2007], Hougaard [2000].

To capture the influence of unmeasured or latent factors on survival outcomes, frailty models introduce a random effect into the hazard function typically as a multiplicative term that adjusts an individual's risk of failure based on hidden characteristics. These models are particularly valuable in epidemiological research, where patient outcomes often reflect a complex interplay of both observed covariates (such as age or comorbidities) and unobserved heterogeneity (such as immune response or genetic predisposition). In recent years, frailty modeling has gained prominence in COVID-19 survival studies as a powerful tool for explaining differential patient outcomes. Several studies have applied frailty models to better understand survival dynamics during the pandemic. For instance, gamma frailty models used to analyze recovery times in hospitalized patients, revealing significant latent effects beyond clinical predictors. Similarly, Pandey et al. [2024] proposed a new continuous lifetime distribution and demonstrated its utility in modeling COVID-19 and reliability data using various

estimation methods. Other works, including those by Osmani and Ziaee [2023], Yazdani et al. [2024], and Jiang et al. [2025], have employed frailty based approaches to identify hidden risk factors influencing mortality and hospital readmission. Collectively, these studies highlight the importance of incorporating unobserved heterogeneity to improve the accuracy and interpretability of survival analysis models.

The most commonly used frailty distributions include the gamma and inverse Gaussian, both of which offer mathematical convenience and have well established properties. However, these distributions can be restrictive, especially when modeling data with pronounced skewness, over dispersion, or heavy tails features that are often seen in pandemic data or clustered survival outcomes. As a more flexible alternative, the Weighted Lindley (WL) distribution has recently gained attention. Introduced as an extension of the Lindley distribution, the WL distribution offers additional shape control and can more effectively accommodate skewed or long-tailed frailty effects Mazucheli et al. [2016]. Several studies have incorporated the WL distribution into frailty-based survival models, such as the WL-Gamma-Weibull (WL-GW) and WL-Generalized Log-Logistic (WL-GLL2) models Tyagi et al. [2021]. These models have demonstrated improved performance over traditional approaches in capturing variability and clustering in time-to-event data. However, many of these models rely on baseline distributions (like gamma or log-logistic) that may not fully capture the dynamics of hazard rates over time. For instance, in situations where the risk of failure steadily increases or decreases common in chronic diseases or during epidemics more flexible baseline hazard structures are needed. The Weibull distribution is widely known for its ability to model various hazard shapes, including increasing, decreasing, and constant rates. It is especially suited to survival settings where the risk of an event changes over time. Surprisingly, despite its popularity, the Weibull distribution has not yet been systematically paired with the Weighted Lindley frailty distribution in a comprehensive parametric model. This represents a notable gap in the current literature, particularly for contexts like COVID-19 survival data, where hazard dynamics and hidden heterogeneity are both prominent.

In this paper, we introduce the Weibull-Weighted Lindley (WWL) distribution, a new frailty-based survival model that combines the flexibility of the Weibull distribution with the adaptability of the Weighted Lindley frailty. The WWL distribution forms a three-parameter model that is both tractable and adaptable, capable of accommodating a wide range of hazard shapes (increasing, decreasing, and non-monotonic) while addressing unobserved heterogeneity in survival data.

The main contributions of this paper are:

- We define and derive the mathematical properties of the WWL distribution, including its PDF, CDF, survival, hazard, and moment functions.
- We explore estimation of its parameters using both classical methods (MLE, method of moments, maximum product of spacings) and robust alternatives (least squares, percentile estimation).
- We assess the model's estimation accuracy and parameter behavior through extensive simulation studies.
- We apply the model to multiple COVID-19 datasets and compare its performance against a wide array of classical and modern lifetime models.
- We analyze stochastic ordering and order statistics to further establish the model's theoretical strength and inter pretability in applied contexts.

By addressing limitations in existing frailty models and offering a flexible, interpretable alternative, the WWL distribution aims to advance the modeling of complex survival data, particularly in high-stakes public health scenarios.

The rest of the structure of this article is as follows. Section 2 introduces the proposed WWL distribution and illustrates the behavior of its probability density and hazard functions under various parameter settings. In Section 3, we derive important statistical and reliability properties of the WWL distribution, supported by numerical illustrations. Section 4 is devoted to the maximum likelihood estimation (MLE) method for estimating the model parameters, along with asymptotic confidence intervals based on the observed Fisher information matrix. A detailed simulation study is presented in Section 5 to evaluate the performance of the MLE method under varying sample sizes and parameter configurations. Section 6 demonstrates the practical applicability of the WWL distribution by fitting it to several real-world COVID-19 datasets and comparing its performance to existing lifetime models. Finally, concluding remarks and potential directions for future work are provided in Section 7.

2. MODEL FORMULATION

Let T denote a continuous lifetime random variable, and let W represent an unobserved frailty term. The conditional hazard function of T , given $W = w$, at time $t > 0$, is defined as

$$h(t \mid w) = wh_0(t)e^{\underline{X}\underline{\beta}} \tag{1}$$

Where $h_0(t)$ is the baseline hazard function, $\underline{X}$ is a row vector of observed covariates, and $\underline{\beta}$ is a corresponding column vector of regression coefficients. The frailty term w adjusts the baseline risk to account for latent factors not captured by the covariates.

The corresponding conditional survival function is given by

$$S(t \mid w) = \exp\left(-\int_0^t h(x \mid w)\, dx\right) = \exp\left(-wH_0(t)e^{\underline{X\beta}}\right) \quad (2)$$

Where, $H_0(t) = \int_0^t h_0(x)\, dx$ is the cumulative baseline hazard function.

The marginal survival function is obtained by integrating over the distribution of the frailty variable W, assuming it has density $f(w)$:

$$\begin{aligned} S(t) &= \int_0^\infty S(t \mid w) f(w)\, dw \\ &= \int_0^\infty \exp\left(-wH_0(t)e^{\underline{X\beta}}\right) f(w)\, dw \\ &= \mathcal{L}_W\left(H_0(t)e^{\underline{X\beta}}\right) \end{aligned} \quad (3)$$

Where $L_W(\cdot)$ is the Laplace transform of the frailty distribution.

In this study, we assume that the conditional distribution of $X \mid W$ is Weibull with scale $\beta > 0$ and shape $\gamma > 0$, and that the frailty term $W \sim$ Weighted Lindley (γ, τ). The Weighted Lindley (WL) distribution is a flexible two-parameter generalization of the Lindley distribution, defined by the probability density function

$$f_W(w; \gamma, \tau) = \frac{\gamma(\tau + 1)}{(\gamma + \tau)\Gamma(\tau)}\, w^{\tau-1}(1 + w)e^{-\gamma w}, \quad w > 0, \quad (4)$$

Where $\gamma, \tau > 0$, and $\Gamma(\cdot)$ denotes the gamma function. This distribution, proposed by Mazucheli et al. Mazucheli et al. [2016], effectively captures skewed and heavy-tailed frailty. To ensure identifiability, the constraint $\mathbb{E}[W] = 1$ is imposed, which leads to a reparameterization $\gamma = \sqrt{\tau(\tau + 1)}$. Letting $\theta = \tau$, the frailty distribution depends on a single heterogeneity parameter $\theta > 0$.

Given this structure, the marginal survival function of X is derived using the Laplace transform of the WL distribution:

$$S(x) = \left(\frac{\sqrt{\theta(\theta+1)} + (\beta x)^\gamma + \theta}{\sqrt{\theta(\theta+1)} + \theta}\right)\left(\frac{\sqrt{\theta(\theta+1)}}{(\beta x)^\gamma + \sqrt{\theta(\theta+1)}}\right)^{\theta+1} \quad (5)$$

The corresponding cumulative distribution function (CDF) is

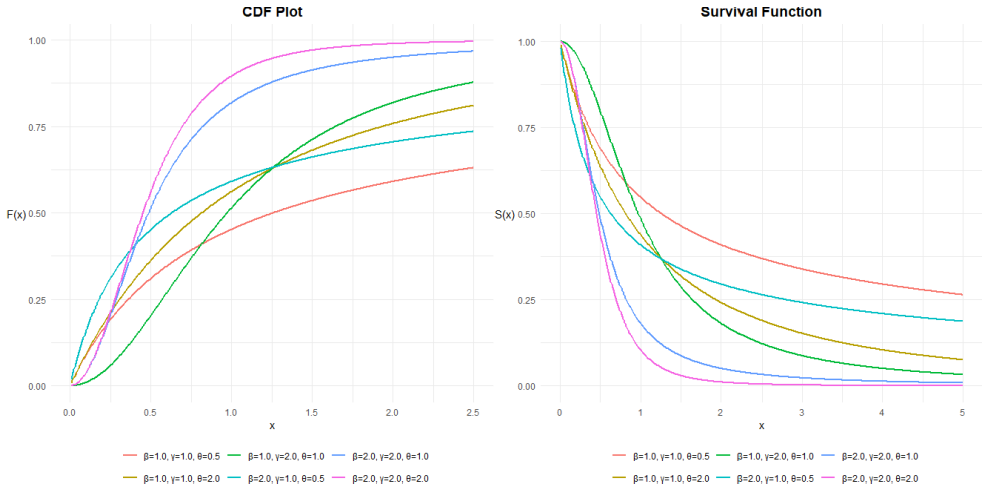
$$F(x) = 1 - S(x) = 1 - \left(\frac{\sqrt{\theta(\theta+1)} + (\beta x)^\gamma + \theta}{\sqrt{\theta(\theta+1)} + \theta}\right)\left(\frac{\sqrt{\theta(\theta+1)}}{(\beta x)^\gamma + \sqrt{\theta(\theta+1)}}\right)^{\theta+1} \quad (6)$$

Differentiating the survival function yields the probability density function (PDF):

$$\begin{aligned} f(x) = \gamma\beta^\gamma x^{\gamma-1} \Bigg[&-\left(\frac{1}{\sqrt{\theta(\theta+1)} + \theta}\left(\frac{\sqrt{\theta(\theta+1)}}{(\beta x)\gamma + \sqrt{\theta(\theta+1)}}\right)^{\theta+1}\right) + \left(\frac{\sqrt{\theta(\theta+1)} + (\beta x)^\gamma + \theta}{\sqrt{\theta(\theta+1)} + \theta}\right)(\theta+1) \\ &\cdot \left(\frac{\sqrt{\theta(\theta+1)}}{(\beta x)\gamma + \sqrt{\theta(\theta+1)}}\right)^{\theta+2} \cdot \left(\frac{1}{(\beta x)\gamma + \sqrt{\theta(\theta+1)}}\right)\Bigg] \end{aligned} \quad (7)$$

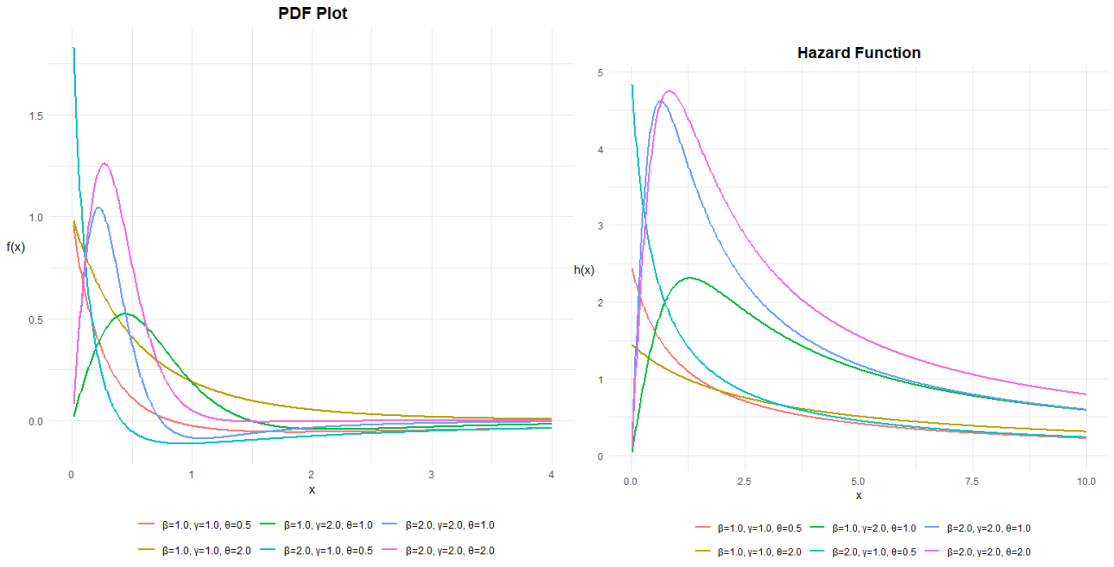
The hazard function, defined as $h(x) = f(x)/S(x)$, simplifies to:

$$h(x) = \gamma\beta^\gamma x^{\gamma-1}\left[\frac{1}{\sqrt{\theta(\theta+1)} + (\beta x)\gamma + \theta} + \frac{\theta+1}{(\beta x)\gamma + \sqrt{\theta(\theta+1)}}\right] \quad (8)$$



CDF Plot **Survival Function**

Figure 1: CDF and Survival Function Plots of the WWL Distribution



PDF Plot **Hazard Function**

Figure 2: PDF and Hazard Function Plots of the WWL Distribution

The WWL distribution has several useful properties for modeling lifetime data. Its support lies on the positive real line, $x \in [0, \infty)$, which is consistent with the nature of survival times. Figures 1(a) and 1(b) shows the behavior of the CDF, survival function, PDF, and hazard function of the WWL distribution for different values of the shape parameter γ and frailty parameter θ, with a fixed scale parameter β. In Figure 1(a), the CDF increases smoothly from 0 to 1 as expected, while the survival function shows a steady decline. Figure 2(a) illustrates the flexibility of the WWL PDF, which may take unimodal or decreasing forms. The corresponding hazard functions reveal a variety of patterns such as increasing, decreasing, or bathtub-shaped demonstrating the model's ability to capture a wide range of failure behaviors in survival analysis.

3. STATISTICAL PROPERTIES

Let $X \sim$ WWL (β, γ, θ). This section presents key statistical characteristics of the WWL distribution, including moments, shape measures, and location parameters. Owing to the complexity of the probability density function, many properties are evaluated numerically using standard numerical integration techniques by using R.

3.1 Moments

The *r*-th raw moment of the WWL distribution is defined as

$$\mu'_r = \int_0^\infty x^r f(x)\, dx,$$

Where $f(x)$ denotes the PDF of the WWL distribution, as given in Equation (2).

In particular, the first raw moment μ'_1 represents the mean of the distribution, while the second raw moment μ' is used to compute the variance. The expressions are given by:

$$\mu = \mu'_1, \quad \mathrm{Var}(X) = \mu'_2 - \mu^2$$

3.2 Quantile Function and Median

The quantile function plays a vital role in characterizing a probability distribution. It is especially useful for generating random samples and for computing positional statistics such as the median. The quantile $Q(p)$ of the WWL distribution can be obtained by numerically solving the following equation:

$$S(Q) = 1 - p = \left(\frac{\sqrt{\theta(\theta+1)} + (\beta Q)^\gamma + \theta}{\sqrt{\theta(\theta+1)} + \theta}\right)\left(\frac{\sqrt{\theta(\theta+1)}}{(\beta Q)^\gamma + \sqrt{\theta(\theta+1)}}\right)^{\theta+1}, \qquad (9)$$

where $p \in (0, 1)$ and $S(Q)$ is the survival function of the WWL distribution.

To compute the median, we set $p = 0.5$ in Equation (9). The median values for different parameter settings can be obtained numerically and analyzed to understand the effect of distribution parameters. Typically, the median decreases as the shape parameter γ or the frailty parameter θ increases, indicating a shift toward shorter survival times.

3.3 Mode

A value of the random variable that maximizes the PDF is referred to as the mode. In the case of the WWL distribution, the mode can be determined by solving the following equation:

$$\frac{\partial \log f(x)}{\partial x} = 0,$$

Subject to the condition:

$$\frac{\partial^2 \log f(x)}{\partial x^2} < 0.$$

Due to the nonlinear and complex structure of the WWL PDF, this equation does not yield a closed-form solution. Therefore, we have obtained the values of the mode numerically for different combinations of the parameters β, γ, and θ. These computed values are listed in Table 1. From these numerical results, it is evident that the WWL distribution exhibits unimodality. This conclusion is further supported by the shape of the PDF plots, which show a single peak for all tested parameter settings.

3.4 Skewness and Kurtosis

Skewness and kurtosis are important shape measures of a probability distribution and are derived from its higher-order moments. Skewness quantifies the asymmetry of the distribution, while kurtosis measures the heaviness of the tails and the peakedness of the distribution around the mean. These measures for the WWL distribution are calculated using the following standard formulas:

$$\text{Skewness} = \frac{\mu'_3 - 3\mu\mu'_2 + 2\mu^3}{(\mu'_2 - \mu^2)^{3/2}}, \text{Kurtosis} = \frac{\mu'_4 - 4\mu\mu'_3 + 6\mu^2\mu'_2 - 3\mu^4}{(\mu'_2 - \mu^2)^2},$$

Where, μ is the mean, and μ' denotes the r-th raw moment of the distribution.

The existence of the r-th moment requires that $\theta > \frac{r}{\gamma}$, ensuring the integrability of the moment expression. The parameter γ primarily governs the shape of the hazard function producing increasing hazard rates when $\gamma > 1$ and decreasing rates when $\gamma < 1$. The frailty parameter θ influences the tail behavior and dispersion of the distribution.

Table 1: Summary of Statistical Properties of the WWL Distribution for Various Parameter Combinations

β	γ	θ	Mean	Variance	Median	Mode	Skewness	Kurtosis
1.5	1.5	3	0.7357	0.4525	0.5618	0.28	3.4766	54.8635
2.0	2.0	3	0.5058	0.1102	0.4397	0.33	1.8569	12.0071
1.5	2.0	3	0.6744	0.1959	0.5863	0.44	1.8569	12.0071
2.0	1.5	3	0.5517	0.2545	0.4213	0.21	3.4770	55.5307
2.5	2.0	3	0.4046	0.0705	0.3518	0.26	1.8569	12.0072
1.5	2.5	3	0.6520	0.1148	0.6016	0.52	1.2379	6.7621
2.0	2.5	4	0.4767	0.0551	0.4464	0.39	0.9443	4.8815
2.5	1.5	4	0.4181	0.1218	0.3311	0.17	2.3567	16.3892
2.0	2.0	5	0.4786	0.0806	0.4305	0.34	1.1975	5.7279
3.0	2.0	5	0.3191	0.0358	0.2870	0.23	1.1975	5.7279

Table 1 reports the numerically computed mean, variance, median, mode, skewness, and kurtosis for various combinations of the WWL parameters. The results illustrate several key behaviors. First, the mean tends to decrease as either the shape parameter γ or scale parameter β increases, indicating tighter concentration of mass near the origin. Similarly, the variance decreases with larger parameter values, suggesting more peaked distributions. Skewness and kurtosis values reveal the tail characteristics of the distribution. High skewness and kurtosis values, especially for lower γ indicate long right tails and high peaked- ness. As γ or θ increases, the distribution becomes more symmetric and less dispersed. In all reported cases, the median consistently exceeds the mode, confirming the right-skewed nature of the WWL distribution.

3.5 Stochastic Ordering

Stochastic ordering provides a rigorous framework to compare random variables beyond simple summaries like mean or variance, focusing on the entire distributional behavior. Let X_1 and X_2 be nonnegative random variables with cumulative distribution functions (CDFs) F_1, F_2; probability density functions (PDFs) f_1, f_2; survival functions $S_1 = 1 - F_1$, $S_2 = 1 - F_2$; and hazard functions $h_1 = f_1/S_1$, $h_2 = f_2/S_2$.
Common stochastic orders include:

- **Stochastic Order** ($X_1 \leq_{st} X_2$): $F_1(x) \geq F_2(x)$ for all $x \geq 0$.
- **Hazard Rate Order** ($X_1 \leq_{hr} X_2$): $h_1(x) \geq h_2(x)$ for all $x \geq 0$.
- **Mean Residual Life Order** ($X_1 \leq_{mrl} X_2$): $\mathrm{MRL}_{X1}(x) \leq \mathrm{MRL}_{X2}(x)$ for all x, where $\mathrm{MRL}_X(x) = \mathrm{E}[X - x \mid X > x]$.
- **Likelihood Ratio Order** ($X_1 \leq_{lr} X_2$): The ratio $f_1(x)/f_2(x)$ is decreasing in x.

These orders satisfy the chain of implications (Shaked and Shanthikumar, 1994):

$$X_1 \leq_{lr} X_2 \implies X_1 \leq_{hr} X_2 \implies X_1 \leq_{mrl} X_2 \implies X_1 \leq_{st} X_2.$$

Theorem: Likelihood Ratio Ordering for WWL Distribution

Let $X_1 \sim \mathrm{WWL}(\beta_1, \gamma, \theta)$, $X_2 \sim \mathrm{WWL}(\beta_2, \gamma, \theta)$,
Where $\beta_1 > \beta_2 > 0$ and $\gamma, \theta > 0$ are fixed. Then:

$$X_1 \leq_{lr} X_2 \implies X_1 \leq_{hr} X_2 \implies X_1 \leq_{mrl} X_2 \implies X_1 \leq_{st} X_2.$$

Proof. Recall the WWL PDF:

$$f(x;\beta,\gamma,\theta)=\gamma\beta^{\gamma}x^{\gamma-1}\left[-\left(\frac{1}{\sqrt{\theta(\theta+1)}+\theta}\left(\frac{\sqrt{\theta(\theta+1)}}{(\beta x)\gamma+\sqrt{\theta(\theta+1)}}\right)^{\theta+1}\right)+\left(\frac{\sqrt{\theta(\theta+1)}+(\beta x)\gamma+\theta}{\sqrt{\theta(\theta+1)}+\theta}\right)(\theta+1)\times\left(\frac{\sqrt{\theta(\theta+1)}}{(\beta x)\gamma+\sqrt{\theta(\theta+1)}}\right)^{\theta+2}\cdot\left(\frac{1}{(\beta x)\gamma+\sqrt{\theta(\theta+1)}}\right)\right] \quad (10)$$

Define the likelihood ratio

$$R(x)=\frac{f(x;\beta_1,\gamma,\theta)}{f(x;\beta_2,\gamma,\theta)}.$$

Because β appears inside $(\beta x)^{\gamma}$, increasing β scales this term upward, affecting both numerator and denominator expressions in a way that makes $R(x)$ a decreasing function in x. Intuitively, the higher β corresponds to a distribution with more weight near smaller x, which causes the ratio to decrease with increasing x.

Formally, differentiating $R(x)$ with respect to x shows

$$\frac{d}{dx}R(x)<0,$$

Demonstrating that $R(x)$ is decreasing in x, i.e., $X_1 \leq_{lr} X_2$.

Therefore, by the chain of implications, the WWL distribution satisfies all the common stochastic orders when comparing two distributions with the same γ and θ, but different scale parameters $\beta_1 > \beta_2$. This implies that increasing the scale parameter β yields stochastically smaller lifetimes, which is a useful property in reliability and survival analysis, where earlier failures are associated with larger values of β.

4. PARAMETER ESTIMATION USING MAXIMUM LIKELIHOOD

In this section, we estimate the parameters $\beta > 0$, $\gamma > 0$, and $\theta > 0$ of the Weibull-Weighted Lindley (WWL) distribution using the method of maximum likelihood. Since the probability density function of the WWL distribution is mathematically complex, we use numerical techniques to find the maximum likelihood estimates (MLEs) of the parameters.

Let x = {x1, x2. . . xn} be a random sample of size n from the WWL distribution. The likelihood function for this sample is given by

$$L(\beta,\gamma,\theta)=\prod_{i=1}^{n}\gamma\beta^{\gamma}x_i^{\gamma-1}\left[\frac{(\theta+1)\left(\sqrt{\theta(\theta+1)}+(\beta x_i)^{\gamma}+\theta\right)}{\left((\beta x_i)^{\gamma}+\sqrt{\theta(\theta+1)}\right)^{2}}-\frac{1}{\sqrt{\theta(\theta+1)}+\theta}\right]\left(\frac{\sqrt{\theta(\theta+1)}}{(\beta x_i)^{\gamma}+\sqrt{\theta(\theta+1)}}\right)^{\theta+1}$$

Taking the natural logarithm of the likelihood gives the log-likelihood function:

$$l(\beta,\gamma,\theta)=\sum_{i=1}^{n}\log f\,(x_i;\beta,\gamma,\theta),$$

Where f (xi; β, γ, θ) is the probability density function evaluated at the observed data points.

To find the MLEs of the parameters, we solve the following system of equations:

$$\frac{\partial l}{\partial\beta}=0,\frac{\partial l}{\partial\gamma}=0,\frac{\partial l}{\partial\theta}=0.$$

These likelihood equations do not have closed-form solutions because the log-likelihood function is highly nonlinear. Traditional optimization methods like Newton–Raphson and BFGS may fail to converge or get stuck in local optima. To handle this, we use the Simulated Annealing (SANN) algorithm, a global optimization method that does not require derivatives. SANN is especially helpful when the likelihood surface is complicated and has many peaks. In all datasets studied, SANN gave accurate and stable parameter estimates, even in cases where traditional methods did not work well.

Asymptotic Inference and Confidence Intervals

Under regular mathematical conditions, the MLEs are known to be consistent and asymptot- ically normal. This means that, as the sample size increases, the distribution of the MLEs approaches a multivariate normal distribution:

$$\hat{\theta} \xrightarrow{d} \mathcal{N}\left(\theta, \mathcal{I}^{-1}(\theta)\right),$$

Where $\boldsymbol{\mathcal{X}}(\boldsymbol{\theta})$ is the Fisher information matrix. In practical applications, we use the observed information matrix, calculated as the negative of the second derivatives (Hessian) of the log-likelihood function, to estimate the variance and covariance of the estimators:

$$\widehat{Cov}\left(\hat{\beta}, \hat{\gamma}, \hat{\theta}\right) = \left[-\frac{\partial^2 l}{\partial\theta\ \partial\theta^{\top}}\right]^{-1}.$$

Using the standard errors obtained from this matrix, we construct approximate $100(1-\alpha)$ % confidence intervals for each parameter θ_j as:

$$\widehat{\theta_j} \pm z_{\alpha/2} \cdot \mathrm{SE}\left(\widehat{\theta_j}\right),$$

Where $z_{\alpha/2}$ is the critical value of the standard normal distribution and SE ($\hat{\theta}_j$) is the standard error of the estimate.

5. SIMULATION STUDY

To evaluate the performance of different estimation techniques for the parameters of the proposed distribution, a detailed Monte Carlo simulation study was conducted. The simulation was designed to compare the bias and mean squared error (MSE) of several estimators of MLE, LSE, MPSE, WLSE, and CoV. To generate synthetic data from the proposed distribution, Inverse transformation method was used. The simulation study was conducted under the different combinations of parameter. But due to lacke of space, only a set of parameter with two different sample size is reported. The simulation, presented in Table 2, summarize the average estimates and their corresponding MSEs for all methods and sample sizes.

Table 2: Simulation result

Method	Methods	$\beta = 2$	$\gamma = 2$	$\theta = 2$
n=30	MLE	2.690184(1.59841)	2.022948(0.1215246)	2.184353(0.3307766)
	COV	2.860183(1.31492)	2.062386(0.1415293)	2.226314(0.4039120)
	LSE	2.757783(1.82703)	2.066167(0.1483239)	2.136594(0.3654879)
	MPSE	2.767748(1.36487)	2.058083(0.1219794)	2.048412(0.2815523)
	WLSE	2.692130(1.74796)	2.056905(0.1160827)	2.131563(0.3221727)
n=100	MLE	2.088073(1.02090)	1.998004(0.0159074)	2.018383(0.0318133)
	COV	2.096131(1.05571)	2.018440(0.0214826)	2.047216(0.0527912)
	LSE	2.208290(1.18371)	2.018226(0.0209831)	2.028495(0.0491120)
	MPSE	2.096390(0.83066)	2.017562(0.0168246)	2.009999(0.0329734)
	WLSE	2.143115(0.75158)	2.011609(0.0161525)	2.023315(0.0335803)

It is observed that as sample size increases, the bias and variance of the parameter is decreeing. MLE is working best among the different estimation methods.

6. MODEL FITTING PERFORMANCE ON REAL-WORLD DATASETS

This section evaluates the fitting performance of the proposed Weibull Weighted Lindley (WWL) distribution using four real-world COVID-19 datasets. For comparison, we consider several established lifetime distributions that are widely used in survival and reliability analysis. These include the Gamma, Lognormal (LN), Log-logistic (LL), Inverse Gaussian (IG), Inverse Weibull (IW), Inverted Gamma (Inv-Gamma), and Inverse Lomax (ILO) distributions. All of these models are defined on the positive real line and are capable of capturing different levels of skewness and tail behavior, which are essential for modeling lifetime and failure-time data. The selected models represent a variety of hazard rate shapes, such as increasing, decreasing, and nonmonotonic patterns, providing a comprehensive basis for comparison. This enables a fair assessment of the WWL model's flexibility and goodness-of-fit relative to existing alternatives. The comparative results are presented in Tables 3–6, using six statistical criteria: Log- Likelihood (LogL), Akaike Information Criterion (AIC), Bayesian Information Criterion (BIC), Corrected AIC (CAIC), Hannan-Quinn Information Criterion (HQIC), and Kol- mogorov–Smirnov (KS) statistic along with associated p-values.

Dataset 1: Daily COVID-19 Case Rates in Italy

This dataset consists of n = 111 normalized daily case rates reported in Italy between April 1 and July 20, 2020, sourced from Daud et al. [2024]. These values capture infection rates normalized by population and reflect day-to-day transmission patterns during the early pandemic phase. The observations are: 0.2070, 0.1520, 0.1628, 0.1666, 0.1417, 0.1221, 0.1767, 0.1987, 0.1408, 0.1456, 0.1443, 0.1319, 0.1053, 0.1789, 0.2032, 0.2167, 0.1387, 0.1646, 0.1375, 0.1421, 0.2012, 0.1957, 0.1297, 0.1754, 0.1390, 0.1761, 0.1119, 0.1915, 0.1827, 0.1548, 0.1522, 0.1369, 0.2495, 0.1253, 0.1597, 0.2195, 0.2555, 0.1956, 0.1831, 0.1791, 0.2057, 0.2406, 0.1227, 0.2196, 0.2641, 0.3067, 0.1749, 0.2148, 0.2195, 0.1993, 0.2421, 0.2430, 0.1994, 0.1779, 0.0942, 0.3067, 0.1965, 0.2003, 0.1180, 0.1686, 0.2668, 0.2113, 0.3371, 0.1730, 0.2212, 0.4972, 0.1641, 0.2667, 0.2690, 0.2321, 0.2792, 0.3515, 0.1398, 0.3436, 0.2254, 0.1302, 0.0864, 0.1619, 0.1311, 0.1994, 0.3176, 0.1856, 0.1071, 0.1041, 0.1593, 0.0537, 0.1149, 0.1176, 0.0457, 0.1264, 0.0476, 0.1620, 0.1154, 0.1493, 0.0673, 0.0894, 0.0365, 0.0385, 0.2190, 0.0777, 0.0561, 0.0435, 0.0372, 0.0385, 0.0769, 0.1491, 0.0802, 0.0870, 0.0476, 0.0562, 0.0138

Table 3: Model comparison for Italian COVID-19 case data

Distribution	LogL	AIC	BIC	CAIC	HQIC	KS Stat	*p*-value
WWL	**128.5565**	**-251.1131**	**-242.9845**	**-250.8888**	**-247.8156**	**0.0668**	**0.9353**
Gamma	126.3586	-248.7172	-243.2981	-248.6061	-245.7776	0.0643	0.7490
ILo	125.9880	-247.9759	-242.5568	-247.8648	-245.7776	0.9819	0.0000
LL	122.4670	-240.9340	-235.5150	-240.8229	-238.7357	0.0751	0.5578
IG	112.3282	-220.6564	-215.2374	-220.5453	-218.4581	0.1688	0.0036
Inv-Gamma	99.7801	-195.5602	-190.1411	-195.4491	-193.3618	0.1865	0.0009
IW	88.8366	-173.6731	-168.2541	-173.5620	-171.4748	0.1906	0.0006

Dataset 2: Daily COVID-19 Mortality Rates in Saudi Arabia

This dataset includes n = 61 daily COVID-19 mortality rates reported in Saudi Arabia from July 6 to September 4, 2021 Suleiman et al. [2024]. The data represent the proportion of deaths per day due to COVID-19 and are used to evaluate model performance on mortality dynamics. The observations are: 0.3086, 0.3283, 0.2865, 0.2450, 0.2852, 0.3251, 0.2636, 0.3236, 0.2824, 0.2817, 0.3012, 0.2603, 0.2997, 0.2393, 0.2785, 0.2778, 0.2375, 0.2963, 0.2167, 0.2752, 0.2353, 0.2347, 0.1951, 0.2140, 0.2329, 0.2711, 0.2126, 0.2314, 0.1924, 0.2113, 0.2683, 0.2487, 0.2674, 0.1716, 0.2666, 0.2091, 0.2278, 0.1706, 0.2271, 0.1890, 0.2077, 0.2452, 0.1319, 0.2259, 0.1504, 0.1879, 0.1689, 0.2063, 0.2249, 0.1686, 0.1310, 0.1497, 0.1309, 0.1495, 0.1121, 0.1121, 0.1307, 0.1120, 0.1306, 0.1492, 0.0932

Table 4: Model comparison for Saudi Arabian COVID-19 mortality data

Distribution	LogL	AIC	BIC	CAIC	HQIC	KS Stat	*p*-value
WWL	**85.0767**	**-164.1534**	**-157.8208**	**-163.7324**	**-161.6716**	**0.0868**	**0.9327**
Gamma	82.0595	-160.1190	-155.8972	-159.9121	-158.4644	0.1109	0.4406
LN	80.0984	-156.1968	-151.9751	-155.9899	-154.5423	0.1311	0.2455
IG	80.0738	-156.1475	-151.9258	-155.9406	-154.4930	0.1367	0.2041
LL	79.1505	-154.3009	-150.0792	-154.0940	-152.6464	0.0995	0.5816
ILo	79.0729	-154.1458	-149.9241	-153.9389	-152.4913	0.9681	0.0000
Inv-Gamma	77.5243	-151.0486	-146.8268	-150.8417	-149.3940	0.1535	0.1131
IW	70.7702	-137.5404	-133.3186	-137.3335	-135.8858	0.1780	0.0420

Dataset 3: Daily COVID-19 Mortality Rates in Canada

This dataset comprises n = 54 daily mortality rates collected in Canada from November 1 to December 26, 2020 Ahmad et al. [2024]. These values illustrate pandemic severity during a winter surge and are normalized to facilitate cross-model comparisons. The values of the dataset are: 0.1622, 0.1159, 0.1897, 0.1260, 0.3025, 0.2190, 0.2075, 0.2241, 0.2163, 0.1262, 0.1627, 0.2591, 0.1989, 0.3053, 0.2170, 0.2241, 0.2174, 0.2541, 0.1997, 0.3333, 0.2594, 0.2230, 0.2290, 0.1536, 0.2024, 0.2931, 0.2739, 0.2607, 0.2736, 0.2323, 0.1563, 0.2677, 0.2181, 0.3019, 0.2136, 0.2281, 0.2346, 0.1888, 0.2729, 0.2162, 0.2746, 0.2936, 0.3259, 0.2242, 0.1810, 0.2679, 0.2296, 0.2992, 0.2464, 0.2576, 0.2338, 0.1499, 0.2075, 0.1834, 0.3347, 0.2362

Table 5: Model comparison for regional COVID-19 case data

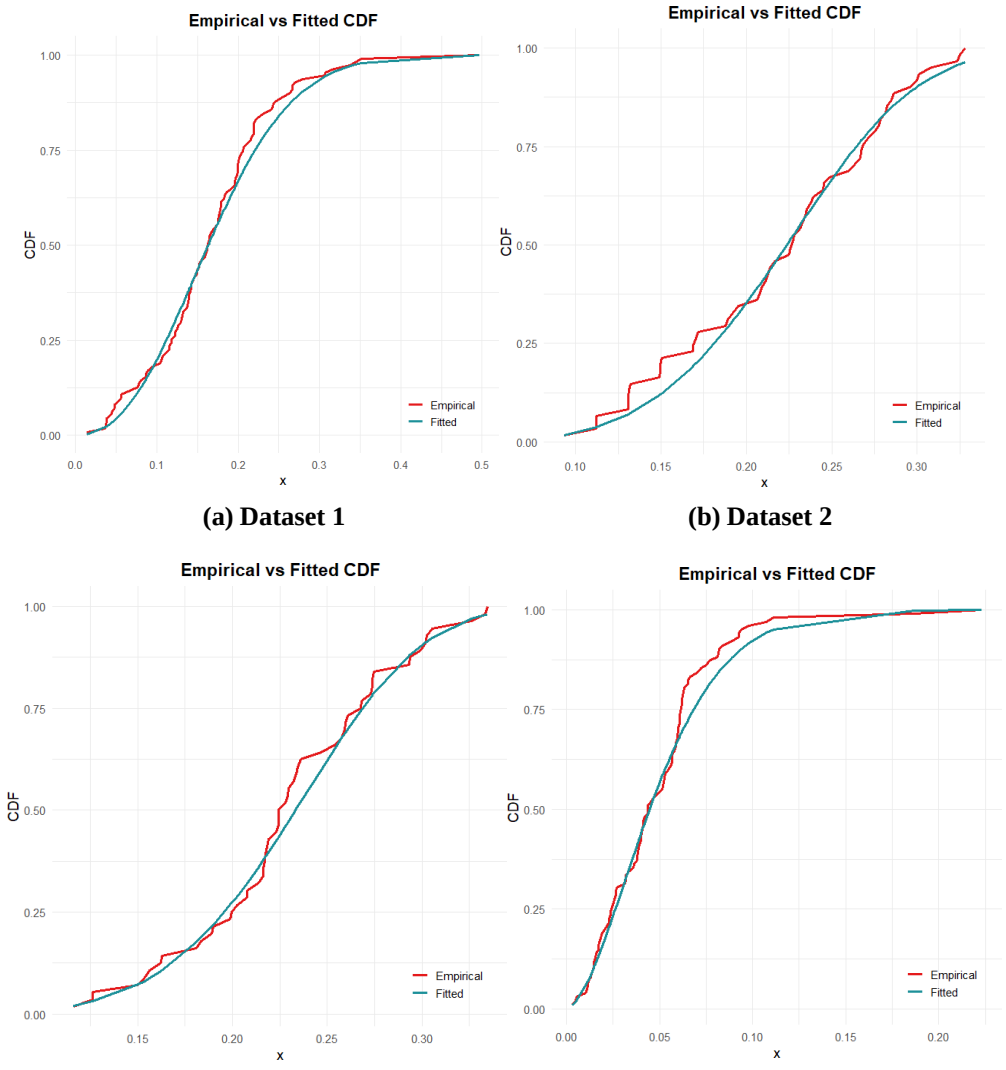
Distribution	LogL	AIC	BIC	CAIC	HQIC	KS Stat	*p*-value
WWL	**86.6007**	**-167.2014**	**-161.1253**	**-166.7399**	**-164.8457**	**0.1028**	**0.9077**
Gamma	85.3663	-166.7327	-162.6820	-166.5063	-165.1622	0.1020	0.6049
ILo	85.0348	-166.0696	-162.0189	-165.8431	-164.4991	0.9746	0.0000
LL	84.3775	-164.7550	-160.7042	-164.5285	-163.1845	0.0794	0.8721
LN	84.0358	-164.0716	-160.0209	-163.8451	-162.5011	0.1183	0.4137
IG	83.9449	-163.8898	-159.8390	-163.6633	-162.3193	0.1218	0.3768
Inv-Gamma	82.1736	-160.3471	-156.2964	-160.1207	-158.7767	0.1364	0.2486
IW	74.3971	-144.7941	-140.7434	-144.5677	-143.2237	0.1704	0.0773

Dataset 4: Daily COVID-19 Death-to-Case Ratios in the United Kingdom

This dataset contains n = 107 daily death-to-case ratios observed in the UK between March 12 and June 28, 2020 Alsuhabi et al. [2022]. The metric, calculated as daily deaths divided by daily cases, is a proxy for infection fatality rate and serves as an important epidemiological indicator. For instant access to the data, see the following: 0.0149, 0.0235, 0.0230, 0.0159, 0.0200, 0.0413, 0.0360, 0.0378, 0.0363, 0.0399, 0.0453, 0.0436, 0.0598, 0.0624, 0.0546, 0.0607, 0.0609, 0.0521, 0.0615, 0.0928, 0.2232, 0.0620, 0.0812,0.0629, 0.0651, 0.0840, 0.1072, 0.0821, 0.0567, 0.0559, 0.0606, 0.0380, 0.0586, 0.0980, 0.0925, 0.0631, 0.1869, 0.0049, 0.0176, 0.0495, 0.1112, 0.0890, 0.0940, 0.0600, 0.0652, 0.0413, 0.0588, 0.0665, 0.0816, 0.0753, 0.0579, 0.0436, 0.0527, 0.0382, 0.0568, 0.0613, 0.0531, 0.0767, 0.0400, 0.0406, 0.0237, 0.0471, 0.0722, 0.0595, 0.0597, 0.0389, 0.0265, 0.0518, 0.0419, 0.0566, 0.0516, 0.0390, 0.0245, 0.0266, 0.0314, 0.0701, 0.0410, 0.0436, 0.0320, 0.0255, 0.0171, 0.0268, 0.0259, 0.0333, 0.0318, 0.0188, 0.0172, 0.0112, 0.0155, 0.0229, 0.0184, 0.0621, 0.0146, 0.0114, 0.0216, 0.0103, 0.0129, 0.0134, 0.0117, 0.0143, 0.0032, 0.0054

Table 6: Model comparison for regional COVID-19 mortality data

Distribution	LogL	AIC	BIC	CAIC	HQIC	KS Stat	*p*-value
WWL	**221.2130**	**-436.4261**	**-428.5512**	**-436.1812**	**-433.2373**	**0.0894**	**0.711**
Gamma	222.8574	-441.7147	-436.4648	-434.4648	-439.5888	0.0803	0.5009
LL	219.9845	-435.9689	-430.7190	-435.8477	-433.8431	0.0952	0.3142
LN	218.5571	-433.1143	-427.8643	-432.9931	-430.9884	0.1221	0.0953
IG	213.5050	-423.0100	-417.7600	-422.8888	-420.8841	0.1624	0.0092
ILo	211.0788	-418.1576	-412.9076	-418.0364	-416.0317	0.9957	0.0000
Inv-Gamma	203.1266	-402.2532	-397.0032	-402.1320	-400.1273	0.1784	0.0030
IW	198.3012	-392.6024	-387.3525	-392.4812	-390.4766	0.1565	0.0135



(a) Dataset 1 **(b) Dataset 2**

(c) Dataset 3 **(d) Dataset 4**

Figure 3: Empirical vs Fitted CDF for all datasets using the WWL distribution.

The WWL distribution demonstrates strong performance across all datasets, and the proposed WWL model consistently ranks among the top performers across all evaluation metrics. These results support WWL's effectiveness for modeling diverse COVID-19 data characteristics, including case rates and mortality metrics.

To visually assess the goodness-of-fit of the WWL distribution to each dataset, we compare the empirical cumulative distribution function (ECDF) with the fitted WWL CDF derived using MLE. The plots below compare the fitted WWL CDF (in blue) with the ECDF (in red) for each dataset. A close match suggests that the WWL model adequately captures the empirical distribution.

Table 7 presents the estimated values of the WWL parameters β, γ, and θ for all four datasets. Each estimate is accompanied by its standard error (SE), and a 95% confidence interval (CI). These results help assess how well the WWL model fits each dataset. All computations were performed in R using the maxLik package. The SANN algorithm was used for log-likelihood maximization, and SEs were obtained from the inverse Hessian matrix. Confidence intervals were based on normal approximation.

7. CONCLUSION

This study introduced the WWL distribution, a new frailty-based model for lifetime data that combines the flexibility of the Weibull distribution with the heterogeneity handling capacity of the Weighted Lindley distribution. The WWL model offers a wide range of hazard rate shapes and captures unobserved variability in survival analysis. Distribution function is having close form, moments and stochastic ordering of WWL distribution exists. Hazard function is decreasing, increasing and then decreasing. We derived its key properties and applied various estimation methods, including MLE, MPSE, MoM, LSE, WLSE and CvME through simulations. The results are on expected line that bise and MSE is decreasing as sample size increases. Four covid data sets are used for the data analysis. First data is daily COVID-19 case rates in Italy. This data fits well to the WWL distribution with log likelihood =128.5565, KS= 0.0668 and p value=0.9353. Which is best. Second data is daily COVID-19 mortality rates in Saudi Arabia. This data fits well to the WWL distribution with log likelihood = 85.0767, KS= 0.0868 and p value=0.9327. Which is best. Third data set daily COVID-19 mortality rates in Canada. This data fits well to the WWL distribution with log likelihood = 86.6007, KS= 0.1028 and p value=0.9077. The fourth data set is daily COVID-19 death-to-case ratios in the United Kingdom. This data fits well to the WWL distribution with log likelihood = 221.2130, KS= 0.0894 and p value= 0.711. Data analysis of different COVID-19 data sets shows the superiority of fit compared to several existing models. Overall, the WWL model provides a valuable and flexible tool for modelling survival data, especially when dealing with complex hazard behaviour and latent heterogeneity. Future work may focus on extending the model to censored data and multivariate frameworks.

Table 7: MLEs, Standard Errors, and 95% Confidence Intervals for WWL Parameters

Dataset	Parameter	Estimate	Std. Error	95% CI
1	β	3.4379	0.3443	[2.7631, 4.1127]
	γ	1.9922	0.1698	[1.6594, 2.3250]
	θ	5.1608	1.5317	[2.1587, 8.1629]
2	β	3.1057	0.3325	[2.4540, 3.7574]
	γ	1.6453	0.1532	[1.3450, 1.9456]
	θ	8.0431	2.0783	[3.9696, 12.1166]
3	β	4.0033	0.4018	[3.2158,4.7908]
	γ	5.0735	0.5115	[4.0710, 6.0759]
	θ	17.1519	3.2131	[10.8543,23.4495]
4	β	18.1558	1.1528	[15.8963, 20.4153]
	γ	1.6779	0.1256	[1.4318, 1.9240]
	θ	21.3156	4.2243	[13.0362, 29.5950]

REFERENCES

[1] Aijaz Ahmad, Najwan Alsadat, Aafaq A Rather, MA Meraou, and Marwa M Mohie El-Din. A novel statistical approach to covid-19 variability using the weibull-inverse nadarajah haghighi distribution. *Alexandria Engineering Journal*, 107:950–962, 2024.

[2] Hassan Alsuhabi, Ibrahim Alkhairy, Ehab M Almetwally, Hisham M Almongy, Ahmed M Gemeay, EH Hafez, RA Aldallal, and Mohamed Sabry. A superior extension for the lomax distribution with application to covid-19 infections real data. *Alexandria Engineering Journal*, 61(12):11077–11090, 2022.

[3] Hanita Daud, Ahmad Abubakar Suleiman, Aliyu Ismail Ishaq, Najwan Alsadat, Mohammed Elgarhy, Abubakar Usman, Pitchaya Wiratchotisatian, Usman Abdullahi Ubale, and Yu Liping. A new extension of the gumbel distribution with biomedical data analysis. *Journal of Radiation Research and Applied Sciences*, 17(4):101055, 2024.

[4] Luc Duchateau and Paul Janssen. *The Frailty Model*. Springer, 2007.

[5] Philip Hougaard. *Analysis of multivariate survival data*, volume 564. Springer, 2000. Sihang Jiang, Johanna Loomba, Andrea Zhou, Suchetha Sharma, Saurav Sengupta, Jiebei

[6] Liu, Donald Brown, and N3C Consortium. A bayesian survival analysis on long covid and non-long covid patients: A cohort study using national covid cohort collaborative (n3c) data. *Bioengineering*, 12(5):496, 2025.

[7] David G. Kleinbaum and Mitchel Klein. *Survival Analysis: A Self-Learning Text*. Springer Science & Business Media, 2nd edition, 2005.

[8] Jerald F. Lawless. *Statistical Models and Methods for Lifetime Data*. John Wiley & Sons, 2003.

[9] Josmar Mazucheli, Alexandre F. B. Menezes, and Layla B. Fernandes. A new one-parameter lifetime distribution: the weighted lindley distribution. *Journal of Statistical Computation and Simulation*, 86(5):891–905, 2016.

[10] Freshteh Osmani and Masood Ziaee. Survival evaluation of hospitalized covid-19 patients with cox frailty approach. *Disaster Medicine and Public Health Preparedness*, 17:e233, 2023.

[11] Arvind Pandey, Ravindra Singh, Shikhar Tyagi, and Abhishek Tyagi. Modelling climate, covid-19, and reliability data: A new continuous lifetime model under different methods of estimation. *Statistics and Applications*, 22:1–27, 01 2024.

[12] Ahmad Abubakar Suleiman, Hanita Daud, Aliyu Ismail Ishaq, Mahmod Othman, Huda M Alshanbari, and Sundus Naji Alaziz. A novel extended k umaraswamy distribution and its application to covid-19 data. *Engineering Reports*, 6(12):e12967, 2024.

[13] Shikhar Tyagi, Arvind Pandey, Varun Agiwal, and Christophe Chesneau. Weighted lindley multiplicative regression frailty models under random censored data. *Computational and Applied Mathematics*, 40:1–24, 2021.

[14] Akram Yazdani, Seyyed Ali Mozaffarpur, Pouyan Ebrahimi, Hoda Shirafkan, and Hamed Mehdinejad. Comorbidities affecting re-admission and survival in covid-19: Application of joint frailty model. *Plos one*, 19(4):e0301209, 2024.