In-Silico Studies and Property Model to Investigate the Binding Effect of Selected Ligands on HIV Integrase

Maradesa Adeleke, Adegoke Afeez Babatunde

Federal University of Technology Akure, P.M.B 704, Ondo State Nigeria

*Corresponding authors: Maradesa Adeleke, e-mail: maradesasta121378@futa.edu.ng
Adegoke Afeez Babatunde, e-mail: adegokeche120848@futa.edu.ng

Doi 10.29072/basjs.202023

Abstract

This research proposes a new weighted Halfnormal Distributuion (W-HND); the model is built by using the weighting function to add the weighted parameter to the classical Halfnormal distribution. The W-HND has been characterized with heavily tail and leptokurtic, with increasing value of the weighted parameter; the distribution is tending to symmetry thus making the W-HND accurate in modelling both heavily tailed and lightly tailed data. The W-HND is fitted to the data obtained via molecular dynamics simulation, in which two drug candidates (Ligands) called Streptomycin and Abacavir were subjected to molecular docking (in-silico study); their binding effect on HIV protein receptors was studied, and the mean binding affinity for Streptomycin (6.61708 kcal/mol) shows a great binding effect as compared to Abacavir (6.246396 kcal/mol). Performance criteria such as Akaike information (AIC) and Bayesian information criteria (BIC) were used as established standard for selecting best performing model; in this case, WHND performs efficiently as compared to Lognormal and Weibull Model

Article inf.

Received: 27/5/2020

Accepted 29/6/2020

Published 31/8/2020

Keywords:

Weighted Halfnormal Distribution (W-HND), Ligands, Hazard Function, Binding Affinity, Protein Receptor.

1. Introduction

The process of designing and discovering new drugs with their biological activities from wet experiment is very challenging, expensive and even time consuming as compared to computer aided drug design (CADD). The process has been accelerated due to development of bioinformatics tools and using probability models to probe the efficacy and effectiveness of most useful drugs today. Over the last few years, computer aided drug design (CADD) which is also known as *in silico* screening has become a powerful technique because of its utility in various phases of drug discovery and development through various advanced features, Joseph et al. [1]. This is an emerging field that uses computational approach in simulating drug - receptor interactions. This approach has actually increased the rate of discovering new drugs, reduced cost of wet experiment and also changed the ways drugs are being designed. In pharmaceutical, medicinal as well as in other scientific research; a computer plays a very important role, even in development of new compound in quest for better therapeutic agents, Gilda et al. [2]. Combination of rational drug design and structure biology leads to discovery of novel therapeutic agents. Rational Drug Design (RDD) which helps to facilitate and speed up the drug designing process, which involves variety of method to identify novel compounds. One of such method is docking of drug-like (ligand) molecule with a receptor (target), Luu et al. [3].

Human immunodeficiency virus (HIV) is a deadly virus which is of two types (HIV 1 and 2). Both types were originated from simian immunodeficiency viruses (SIVs) of primates, the origin of the two are zoonotic which eventually spread from one human to the other, World Health Organization [4]. Type 1 was first isolated from in the year 1983 and the type two in 1986 which are of two different epidemics. SIV from chimpanzees and sooty mangabey monkey gave rise to type 1 and type 2 respectively in human beings, Weiss et al. [5]. The exact transmission of the SIVs to human is yet to be confirmed but was attached to hunting of these animals for food by indigenous people where most of these animals' species live, Kalish et al. [6]. Ever since its existence in humans, many drugs have been discovered and used to combat the disease in which the virus have shown a resistible activity towards many of those drugs. Hence, combinations of drugs have recently been employed in other to overcome the resistance, Mak et al. [7]. This work is designed to investigate the inhibitory activities of two selected drug targets (Streptomycin and Abacavir) against HIV protein receptor using both computation approach and probability model

to establish efficacy of Streptomycin and Abacavir on HIV and to determine the performance of newly built convoluted distribution over some existing probability distributions.

Statistician are so curious about biases and they are working assiduously to finding every means of removing or reducing biasness in data so that the parameter of any fitted model will be significant, the weighing distribution method is a special method of handling data that has been influenced by bias due to the different time in which the data was collected; this type of data is called length-biased data and it becomes imperative to build a distribution with utmost robustness to some of biasness in dataset, Patil & Rao[8]. Also, the truncated distributions and damaged observations can give rise to weighted distributions. Biased data of this type arises in all disciplines of science and many spheres of life and statisticians have discovered and have been working tirelessly to give convoluted distribution as the better alternative to handling the problems which, in turn leads to corrections of the biases, Patil and Rao[8]. In this research, the Weighted Halfnormal distribution (W-HND) will be built, the model will be fiited to data obtained via simulation of molecular dynamics and comparison would be made between W-HND and other candidate model using the performance comparison criteria

2.1 Derivation of Weighted Halfnormal Distribution(W-HND)

Patil and Rao[8] developed method of weighting classical probability distribution which add weighted parameter to the baseline distribution in order to improve its flexibility in modeling length-biased data. Many researchers have employed patil et al's method to build some other convoluted distributions: Sanku et al [9] used [8] to construct new weighted weibull distribution (WWD), studied the model and applied WWD to both simulated and lifetime data, Abd el-Monsef & Ghoneim[10] proposed Weighted kumaraswamy distribution and was used to model biological data. Also N.M. Kilany[11] proposed Weighted-Lomax distribution by using the method of weighting distribution[8]. The Weighted Halfnormal is proposed in this research work using [8].

Given that $f(x;\sigma) = \frac{\sqrt{2}}{\sigma\sqrt{\pi}} e^{-\frac{x^2}{2\sigma^2}} \quad x > 0$ and $F(x;\sigma) = erf\left(\frac{x}{\sigma\sqrt{2}}\right)$ are the Probability Density Function (PDF) and Cumulative Distribution Function (CDF) of Halfnormal distribution respectively, Let X be a non negative random variable with its pdf $f(x;\sigma)$, then $g(x;c,\sigma)$, distribution is a weighted version of $f(x;\sigma)$, if its PDF is given (Patil et al.[8]) by:

$$g(x;c,\sigma) = \frac{w(x;c)f(x;\tau)}{E(w(x;c))} \quad (1)$$

Then the Weighted-Halfnormal Distribution can be viewed as:

$$= E\big(w(x;c)\big) = \int_0^\infty x^c \frac{\sqrt{2}}{\sigma\sqrt{\pi}} e^{-\frac{x^2}{2\sigma^2}}\, dx; \quad x > 0 \quad (2)$$

Let y = $\frac{x^2}{2\sigma^2}$; $dx = \frac{\sigma^2}{x}dy$; $x = \sigma\sqrt{2y}$;

$$= \frac{\sqrt{2}}{\sigma\sqrt{\pi}}\int_0^\infty \big(\sigma\sqrt{2y}\big)^{c-1} e^{-y}\sigma^2 dy = \frac{\sigma\sqrt{2}}{\sqrt{\pi}}\int_0^\infty \sigma^{c-1}\, 2^{\frac{c}{2}-\frac{1}{2}} y^{\frac{c}{2}-\frac{1}{2}} e^{-y} dy \quad (3)$$

$$= \frac{\sigma^{c-1}\sigma 2^{\frac{1}{2}} 2^{\frac{c}{2}-\frac{1}{2}}}{\sqrt{\pi}}\int_0^\infty y^{\frac{c}{2}-\frac{1}{2}} e^{-y} dy = \frac{\sigma^c 2^{\frac{c}{2}}}{\sqrt{\pi}}\,\Gamma\left(\frac{c}{2}+\frac{1}{2}\right) = \frac{(\sigma\sqrt{2})^c}{\sqrt{\pi}}\,\Gamma\left(\frac{c}{2}+\frac{1}{2}\right) \quad (4)$$

$$g(x;c,\sigma) = \frac{x^c \frac{\sqrt{2}}{\sigma\sqrt{\pi}} e^{-\frac{x^2}{2\sigma^2}}}{\frac{(\sigma\sqrt{2})^c}{\sqrt{\pi}}\Gamma\left(\frac{c}{2}+\frac{1}{2}\right)} = \frac{x^c \frac{\sqrt{2}}{\sigma\sqrt{\pi}} e^{-\frac{x^2}{2\sigma^2}}}{\Gamma\left(\frac{c}{2}+\frac{1}{2}\right)} \cdot \frac{\sqrt{\pi}}{(\sigma\sqrt{2})^c} = \frac{\sqrt{2}}{\sigma\Gamma\left(\frac{c}{2}+\frac{1}{2}\right)}\left(\frac{x}{\sigma\sqrt{2}}\right)^c e^{-\frac{x^2}{2\sigma^2}} \quad (5)$$

The equation (5) above is the PDF of Weighted Halfnormal Distribution (W-HND)

Then for deriving the CDF we proceed as follows:

$$= \int_0^y \frac{\sqrt{2}}{\sigma\Gamma\left(\frac{c}{2}+\frac{1}{2}\right)}\left(\frac{x}{\sigma\sqrt{2}}\right)^c e^{-\frac{x^2}{2\sigma^2}}\, dx \;;\; y = \left(\frac{x}{\sigma\sqrt{2}}\right)^c;\; dx = \sigma\sqrt{2}\, dy \quad (6)$$

$$= \frac{\sqrt{2}}{\sigma\Gamma\left(\frac{c}{2}+\frac{1}{2}\right)}\int_0^y y^c\, e^{y^2}\sigma\sqrt{2}\, dy\,; \quad z = y^2 \;;\; dz = 2y\, dy \quad (7)$$

$$= \frac{2}{\Gamma\left(\frac{c}{2}+\frac{1}{2}\right)}\int_0^y \big(\sqrt{z}\big)^c e^{-z}\frac{dz}{2\sqrt{z}} = \frac{1}{\Gamma\left(\frac{c}{2}+\frac{1}{2}\right)}\int_0^y z^{\frac{c}{2}-\frac{1}{2}} e^{-z} dz = \frac{1}{\Gamma\left(\frac{c}{2}+\frac{1}{2}\right)}\,\gamma\left(\frac{c}{2}+\frac{1}{2}, z\right) \quad (8)$$

$$G(x,c,\sigma) = \frac{\gamma\left(\frac{c}{2}+\frac{1}{2}, y^2\right)}{\Gamma\left(\frac{c}{2}+\frac{1}{2}\right)} = \frac{\gamma\left(\frac{c}{2}+\frac{1}{2}, \left(\frac{x}{\sigma\sqrt{2}}\right)^2\right)}{\Gamma\left(\frac{c}{2}+\frac{1}{2}\right)} \quad (9)$$

The equation (9) is the CDF of Weighted Halfnormal Distribution (W-HND)

2.2 Moments

Let X be a random variable with PDF g(*x*,c,σ) , that is X~W-HND(c,σ), where c and σ are the weighted and shape parameter respectively, then the moment of W-HND is defined as:

$$E(X^r) = \int_0^\infty x^r\, g(x;c,\sigma)\, dx \;; where\;\; g(x;c,\sigma)\; is\; the\; PDF\; of\; W-HND$$

$$= EX^r = \int_0^\infty x^r \frac{\sqrt{2}}{\sigma\Gamma\left(\frac{c}{2}+\frac{1}{2}\right)}\left(\frac{x}{\sigma\sqrt{2}}\right)^c e^{-\frac{x^2}{2\sigma^2}}\,dx = \frac{\sqrt{2}}{\sigma\Gamma\left(\frac{c}{2}+\frac{1}{2}\right)}\int_0^\infty x^r\, y^c e^{-y^2}dx \qquad (10)$$

Let $y = \frac{x}{\sigma\sqrt{2}}$; $\frac{dy}{dx} = \frac{1}{\sigma\sqrt{2}}$; $dx = \sigma\sqrt{2}dy$; $x = y\sigma\sqrt{2}$, then it follows that,

$$= \frac{\sqrt{2}}{\sigma\Gamma\left(\frac{c}{2}+\frac{1}{2}\right)}\int_0^\infty \left(y\sigma\sqrt{2}\right)^r y^c e^{-y^2}\sigma\sqrt{2}dy = \frac{2(\sigma\sqrt{2})^r}{\Gamma\left(\frac{c}{2}+\frac{1}{2}\right)}\int_0^\infty y^{r+c}\, e^{-y^2}\, dy \qquad (11)$$

$$= \frac{2(\sigma\sqrt{2})^r}{\Gamma\left(\frac{c}{2}+\frac{1}{2}\right)}\int_0^\infty \left(\sqrt{z}\right)^{r+c} e^{-z}\frac{dz}{2\sqrt{z}} = \frac{(\sigma\sqrt{2})^r}{\Gamma\left(\frac{c}{2}+\frac{1}{2}\right)}\int_0^\infty z^{\frac{r}{2}+\frac{c}{2}-\frac{1}{2}}e^{-z}dz = \frac{(\sigma\sqrt{2})^r}{\Gamma\left(\frac{c}{2}+\frac{1}{2}\right)}\int_0^\infty z^{\frac{r}{2}+\frac{c}{2}+\frac{1}{2}-1}e^{-z}dz \qquad (12)$$

$$= \frac{(\sigma\sqrt{2})^r}{\Gamma\left(\frac{c}{2}+\frac{1}{2}\right)}\,\Gamma\left(\frac{r}{2}+\frac{c}{2}+\frac{1}{2}\right) \qquad (13)$$

$$E(X) = \frac{\sigma\sqrt{2}}{\Gamma\left(\frac{c}{2}+\frac{1}{2}\right)}\Gamma\left(\frac{c}{2}+1\right) = \frac{\frac{c\sigma\sqrt{2}}{2}\Gamma\left(\frac{c}{2}\right)}{\Gamma\left(\frac{c}{2}+\frac{1}{2}\right)} \quad r = 1 \qquad (14)$$

$$E(X^2) = \frac{(\sigma\sqrt{2})^2}{\Gamma\left(\frac{c}{2}+\frac{1}{2}\right)}\,\Gamma\left(1+\frac{c}{2}+\frac{1}{2}\right) = \frac{(\sigma\sqrt{2})^2}{\Gamma\left(\frac{c}{2}+\frac{1}{2}\right)}\cdot\left(\frac{c}{2}+\frac{1}{2}\right)\Gamma\left(\frac{c}{2}+\frac{1}{2}\right) = (\sigma\sqrt{2})^2\left(\frac{c}{2}+\frac{1}{2}\right);\, r=2 \qquad (15)$$

$$E(X^3) = \frac{(\sigma\sqrt{2})^3}{\Gamma\left(\frac{c}{2}+\frac{1}{2}\right)}\,\Gamma\left(\frac{3}{2}+\frac{c}{2}+\frac{1}{2}\right) = \frac{(\sigma\sqrt{2})^3}{\Gamma\left(\frac{c}{2}+\frac{1}{2}\right)}\Gamma\left(\frac{c}{2}+2\right) = \frac{c(\sigma\sqrt{2})^3\left(\frac{c}{2}+1\right)\Gamma\left(\frac{c}{2}\right)}{2\Gamma\left(\frac{c}{2}+\frac{1}{2}\right)} \;;\; r=3 \qquad (16)$$

$$E(X^4) = \frac{(\sigma\sqrt{2})^4}{\Gamma\left(\frac{c}{2}+\frac{1}{2}\right)}\Gamma\left(\frac{4}{2}+\frac{c}{2}+\frac{1}{2}\right) = \frac{(\sigma\sqrt{2})^4}{\Gamma\left(\frac{c}{2}+\frac{1}{2}\right)}\,\Gamma\left(\frac{c}{2}+\frac{1}{2}+2\right);\, r=4 \qquad (17)$$

$$= \frac{(\sigma\sqrt{2})^4}{\Gamma\left(\frac{c}{2}+\frac{1}{2}\right)}\left(\frac{c}{2}+\frac{1}{2}+1\right)\left(\frac{c}{2}+\frac{1}{2}\right)\Gamma\left(\frac{c}{2}+\frac{1}{2}\right) = \frac{\left(\frac{c}{2}+\frac{1}{2}+1\right)\left(\frac{c}{2}+\frac{1}{2}\right)(\sigma\sqrt{2})^4}{\Gamma\left(\frac{c}{2}+\frac{1}{2}\right)} \quad r=4 \qquad (18)$$

2.3 Moment About the Mean

$$= \mu_2 = var(X) = E(X^2) - (EX)^2 = (\sigma\sqrt{2})^2\left(\frac{c}{2}+\frac{1}{2}\right) - \left(\frac{\frac{c\sigma\sqrt{2}}{2}\Gamma\left(\frac{c}{2}\right)}{\Gamma\left(\frac{c}{2}+\frac{1}{2}\right)}\right)^2 \qquad (19)$$

$$= \mu_3 = Ex^3 - 3\mu Ex^2 + 3\mu^2 Ex - \mu^3$$

$$\mu_3 = \frac{c(\sigma\sqrt{2})^3\left(\frac{c}{2}+1\right)\Gamma\left(\frac{c}{2}\right)}{2\Gamma\left(\frac{c}{2}+\frac{1}{2}\right)} - 3\frac{\frac{c\sigma\sqrt{2}}{2}\Gamma\left(\frac{c}{2}\right)}{\Gamma\left(\frac{c}{2}+\frac{1}{2}\right)}(\sigma\sqrt{2})^2\left(\frac{c}{2}+\frac{1}{2}\right) + 2\left(\frac{\frac{c\sigma\sqrt{2}}{2}\Gamma\left(\frac{c}{2}\right)}{\Gamma\left(\frac{c}{2}+\frac{1}{2}\right)}\right)^3 \qquad (20)$$

$$\mu_4 = Ex^4 - 4\mu Ex^3 + 6\mu^2 Ex^2 - 4\mu^3 Ex + \mu^4$$

$$\mu_4 = \frac{\left(\frac{c}{2}+\frac{1}{2}+1\right)\left(\frac{c}{2}+\frac{1}{2}\right)(\sigma\sqrt{2})^4}{\Gamma\left(\frac{c}{2}+\frac{1}{2}\right)} - 4\frac{\frac{c\sigma\sqrt{2}}{2}\Gamma\left(\frac{c}{2}\right)}{\Gamma\left(\frac{c}{2}+\frac{1}{2}\right)}\frac{c(\sigma\sqrt{2})^3\left(\frac{c}{2}+1\right)\Gamma\left(\frac{c}{2}\right)}{2\Gamma\left(\frac{c}{2}+\frac{1}{2}\right)} + 6\left(\frac{\frac{c\sigma\sqrt{2}}{2}\Gamma\left(\frac{c}{2}\right)}{\Gamma\left(\frac{c}{2}+\frac{1}{2}\right)}\right)^2(\sigma\sqrt{2})^2\left(\frac{c}{2}+\frac{1}{2}\right) - 3\left(\frac{\frac{c\sigma\sqrt{2}}{2}\Gamma\left(\frac{c}{2}\right)}{\Gamma\left(\frac{c}{2}+\frac{1}{2}\right)}\right)^4$$

(21)

2.4 Skewness and kurtosis

This is a measure of departure from symmetry (asymmetry). It enables us to know the asymptotic properties of the distribution (W-HND). That is, the coefficient of skewness helps to diagnose if W-HND is positively skewed or negatively skewed. Kurtosis (K) measures the combined sizes of the two tails of the distribution (W-HND). If the coefficient (K>3), then data to be modeled has heavier tails than normal distribution, K<3 and K=3 show that the data has lighter tail and mesokurtic respectively.

$$\gamma_1(x) = \frac{(\mu_3)^2}{(\mu_2)^3} = \frac{\left(\frac{c(\sigma\sqrt{2})^3\left(\frac{c}{2}+1\right)\Gamma\left(\frac{c}{2}\right)}{2\Gamma\left(\frac{c}{2}+\frac{1}{2}\right)} - 3\frac{\frac{c\sigma\sqrt{2}}{2}\Gamma\left(\frac{c}{2}\right)}{\Gamma\left(\frac{c}{2}+\frac{1}{2}\right)}(\sigma\sqrt{2})^2\left(\frac{c}{2}+\frac{1}{2}\right) + 2\left(\frac{\frac{c\sigma\sqrt{2}}{2}\Gamma\left(\frac{c}{2}\right)}{\Gamma\left(\frac{c}{2}+\frac{1}{2}\right)}\right)^3\right)^2}{\left((\sigma\sqrt{2})^2\left(\frac{c}{2}+\frac{1}{2}\right) - \left(\frac{\frac{c\sigma\sqrt{2}}{2}\Gamma\left(\frac{c}{2}\right)}{\Gamma\left(\frac{c}{2}+\frac{1}{2}\right)}\right)^2\right)^3} \quad (22)$$

Kurtosis

$$\mathbf{K} = \gamma_2(x) = \frac{\mu_4}{(\mu_2)^2} = \frac{\left[\frac{\left(\frac{c}{2}+\frac{1}{2}+1\right)\left(\frac{c}{2}+\frac{1}{2}\right)(\sigma\sqrt{2})^4}{\Gamma\left(\frac{c}{2}+\frac{1}{2}\right)} - 4\frac{\frac{c\sigma\sqrt{2}}{2}\Gamma\left(\frac{c}{2}\right)}{\Gamma\left(\frac{c}{2}+\frac{1}{2}\right)}\frac{c(\sigma\sqrt{2})^3\left(\frac{c}{2}+1\right)\Gamma\left(\frac{c}{2}\right)}{2\Gamma\left(\frac{c}{2}+\frac{1}{2}\right)} + 6\left(\frac{\frac{c\sigma\sqrt{2}}{2}\Gamma\left(\frac{c}{2}\right)}{\Gamma\left(\frac{c}{2}+\frac{1}{2}\right)}\right)^2(\sigma\sqrt{2})^2\left(\frac{c}{2}+\frac{1}{2}\right) - 3\left(\frac{\frac{c\sigma\sqrt{2}}{2}\Gamma\left(\frac{c}{2}\right)}{\Gamma\left(\frac{c}{2}+\frac{1}{2}\right)}\right)^4\right]}{\left((\sigma\sqrt{2})^2\left(\frac{c}{2}+\frac{1}{2}\right) - \left(\frac{\frac{c\sigma\sqrt{2}}{2}\Gamma\left(\frac{c}{2}\right)}{\Gamma\left(\frac{c}{2}+\frac{1}{2}\right)}\right)^2\right)^2} \quad (23)$$

2.6 Reliability

$$\mathbf{R(x)} = \mathbf{1}- \frac{\gamma\left(\frac{c}{2}+\frac{1}{2}, \left(\frac{x}{\sigma\sqrt{2}}\right)^2\right)}{\Gamma\left(\frac{c}{2}+\frac{1}{2}\right)} \quad (24)$$

2.7 Hazard

$$H(x) = \frac{f(x)}{R(x)} = \frac{\frac{\sqrt{2}}{\sigma\Gamma\left(\frac{c}{2}+\frac{1}{2}\right)}\left(\frac{x}{\sigma\sqrt{2}}\right)^{c} e^{-\frac{x^2}{2\sigma^2}}}{1-\frac{\gamma\left(\frac{c}{2}+\frac{1}{2},\left(\frac{x}{\sigma\sqrt{2}}\right)^2\right)}{\Gamma\left(\frac{c}{2}+\frac{1}{2}\right)}} \quad (25)$$

2.8 Renyi Entropy

Let X follows Weighted Halfnormal Distribution, $X \sim W-HND(c,\sigma)$ then the measure of variation of the uncertain situation for the random variable X is shown by the Renyi Entropy:

$$= \frac{1}{1-r}\log\frac{\sigma}{\sqrt{2r}}\left[\frac{\sigma^{(r-2)}}{\Gamma\left(\frac{c}{2}+\frac{1}{2}\right)}\left(\sqrt{\frac{2}{r}}\right)^{r}\right]^{c}\Gamma\left(\frac{rc}{2}+\frac{1}{2}\right) \quad (26)$$

Proof:

$$=\delta_r = \frac{1}{1-r}\log\int_0^{\infty}\left(\frac{\sqrt{2}}{\sigma\Gamma\left(\frac{c}{2}+\frac{1}{2}\right)}\left(\frac{x}{\sigma\sqrt{2}}\right)^{c} e^{-\frac{x^2}{2\sigma^2}}\right)^{r} dx = \frac{1}{1-r}\log\left(\frac{\sqrt{2}}{\sigma\Gamma\left(\frac{c}{2}+\frac{1}{2}\right)}\right)^{rc}\frac{1}{(\sigma\sqrt{2})^{rc}}\int_0^{\infty} x^{rc} e^{-\frac{rx^2}{2\sigma^2}} dx$$

$$= \frac{1}{1-r}\log\left(\frac{\sqrt{2}}{\sigma\Gamma\left(\frac{c}{2}+\frac{1}{2}\right)}\right)^{rc}\frac{1}{(\sigma\sqrt{2})^{rc}}\cdot\int_0^{\infty} x^{rc} e^{-y} dx \quad (27)$$

$$y = \frac{rx^2}{2\sigma^2}\ ;\ rx^2 = 2\sigma^2 y\ ;\ x = \sigma\sqrt{\frac{2y}{r}}\ ;\ \frac{dy}{dx} = \frac{2rx}{2\sigma^2} = \frac{rx}{\sigma^2}\ ;\ \frac{\sigma^2 dy}{rx} = dx$$

$$= \frac{1}{1-r}\log\left(\frac{\sqrt{2}}{\sigma\Gamma\left(\frac{c}{2}+\frac{1}{2}\right)}\right)^{c}\frac{1}{(\sigma\sqrt{2})^{rc}}\int_0^{\infty}\left(\sigma\sqrt{\frac{2y}{r}}\right)^{rc-1} e^{-y}\frac{\sigma^2 dy}{r} \quad (28)$$

$$= \frac{1}{1-r}\log\left(\frac{\sqrt{2}}{\sigma\Gamma\left(\frac{c}{2}+\frac{1}{2}\right)}\right)^{c}\frac{2^{\frac{rc}{2}-\frac{1}{2}}}{(\sigma\sqrt{2})^{c}}\frac{\sigma^2}{r}\frac{\sigma^{rc-1}}{r^{\frac{rc}{2}-\frac{1}{2}}}\int_0^{\infty} y^{\frac{rc}{2}-\frac{1}{2}} e^{-y} dy \quad (29)$$

$$= \frac{1}{1-r}\log\left(\frac{\sqrt{2}}{\sigma\Gamma\left(\frac{c}{2}+\frac{1}{2}\right)}\right)^{c}\frac{2^{\frac{rc}{2}-\frac{1}{2}}}{(\sigma\sqrt{2})^{c}}\frac{\sigma^2}{r}\frac{\sigma^{rc-1}}{r^{\frac{rc}{2}-\frac{1}{2}}}\Gamma\left(\frac{rc}{2}+\frac{1}{2}\right) = \frac{1}{1-r}\log\left(\frac{\sqrt{2}}{\sigma\Gamma\left(\frac{c}{2}+\frac{1}{2}\right)}\right)^{c}\frac{2^{\frac{rc}{2}-\frac{1}{2}}}{(\sigma\sqrt{2})^{c}}\frac{\sigma^{rc+1}}{r^{\frac{rc}{2}+\frac{1}{2}}}\Gamma\left(\frac{rc}{2}+\frac{1}{2}\right)$$

$$= \frac{1}{1-r}\log\frac{2^{\frac{c}{2}+\frac{rc}{2}-\frac{1}{2}2-\frac{c}{2}}}{\left(\Gamma\left(\frac{c}{2}+\frac{1}{2}\right)\right)^{c}}\frac{\sigma^{rc+1-2c}}{r^{\frac{rc}{2}+\frac{1}{2}}}\Gamma\left(\frac{rc}{2}+\frac{1}{2}\right) = \frac{1}{1-r}\log\frac{2^{\frac{rc}{2}-\frac{1}{2}}}{\left(\Gamma\left(\frac{c}{2}+\frac{1}{2}\right)\right)^{c}}\frac{\sigma^{rc+1-2c}}{r^{\frac{rc}{2}+\frac{1}{2}}}\Gamma\left(\frac{rc}{2}+\frac{1}{2}\right) \quad (30)$$

$$= \frac{1}{1-r}\log\frac{2^{\frac{rc}{2}-\frac{1}{2}}}{\left(\Gamma\left(\frac{c}{2}+\frac{1}{2}\right)\right)^{c}}\frac{\sigma^{c(r-2)+1}}{r^{\frac{rc}{2}+\frac{1}{2}}}\Gamma\left(\frac{rc}{2}+\frac{1}{2}\right) = \frac{1}{1-r}\log\frac{2^{\frac{rc}{2}-\frac{1}{2}}}{\left(\Gamma\left(\frac{c}{2}+\frac{1}{2}\right)\right)^{c}}\frac{\sigma^{c(r-2)}\sigma}{r^{\frac{rc}{2}+\frac{1}{2}}}\Gamma\left(\frac{rc}{2}+\frac{1}{2}\right) \quad (31)$$

$$= \frac{1}{1-r}\log\frac{\sigma^{c(r-2)}}{\left(\Gamma\left(\frac{c}{2}+\frac{1}{2}\right)\right)^{c}}\frac{\sigma 2^{\frac{rc}{2}-\frac{1}{2}}}{r^{\frac{rc}{2}+\frac{1}{2}}}\Gamma\left(\frac{rc}{2}+\frac{1}{2}\right) = \frac{1}{1-r}\log\left(\frac{\sigma^{(r-2)}}{\Gamma\left(\frac{c}{2}+\frac{1}{2}\right)}\right)^{c}\left(\sqrt{\frac{2}{r}}\right)^{rc}\frac{\sigma}{\sqrt{2r}}\Gamma\left(\frac{rc}{2}+\frac{1}{2}\right) \quad (32)$$

$$= \frac{1}{1-r}\log\frac{\sigma}{\sqrt{2r}}\left[\frac{\sigma^{(r-2)}}{\Gamma\left(\frac{c}{2}+\frac{1}{2}\right)}\left(\sqrt{\frac{2}{r}}\right)^{r}\right]^{c}\Gamma\left(\frac{rc}{2}+\frac{1}{2}\right) \quad (33)$$

2.9 Parameter Estimation

$$= \mathrm{L}\left(f(x;c,\sigma)\right) = \prod_{i=1}^{n}\frac{\sqrt{2}}{\sigma\Gamma\left(\frac{c}{2}+\frac{1}{2}\right)}\left(\frac{x}{\sigma\sqrt{2}}\right)^{c}e^{-\frac{x^2}{2\sigma^2}} = \frac{2^{\frac{n}{2}(1+c)}}{\sigma^{n(1+c)}\left(\Gamma\left(\frac{c}{2}+\frac{1}{2}\right)\right)^{n}}e^{-\sum_{i=1}^{n}\frac{x^2}{2\sigma^2}}\prod_{i=1}^{n}x^{c}{}_{i} \quad (34)$$

=

$$\log \mathrm{L}\left(f(x;c,\sigma)\right) = n\left(\frac{1+c}{2}\log 2 - (1+c)\log\sigma - \log\left(\Gamma\left(\frac{c}{2}+\frac{1}{2}\right)\right)\right) - \sum_{i=1}^{n}\frac{x^2}{2\sigma^2} + c\sum_{i=1}^{n}\log(x)$$

(35)

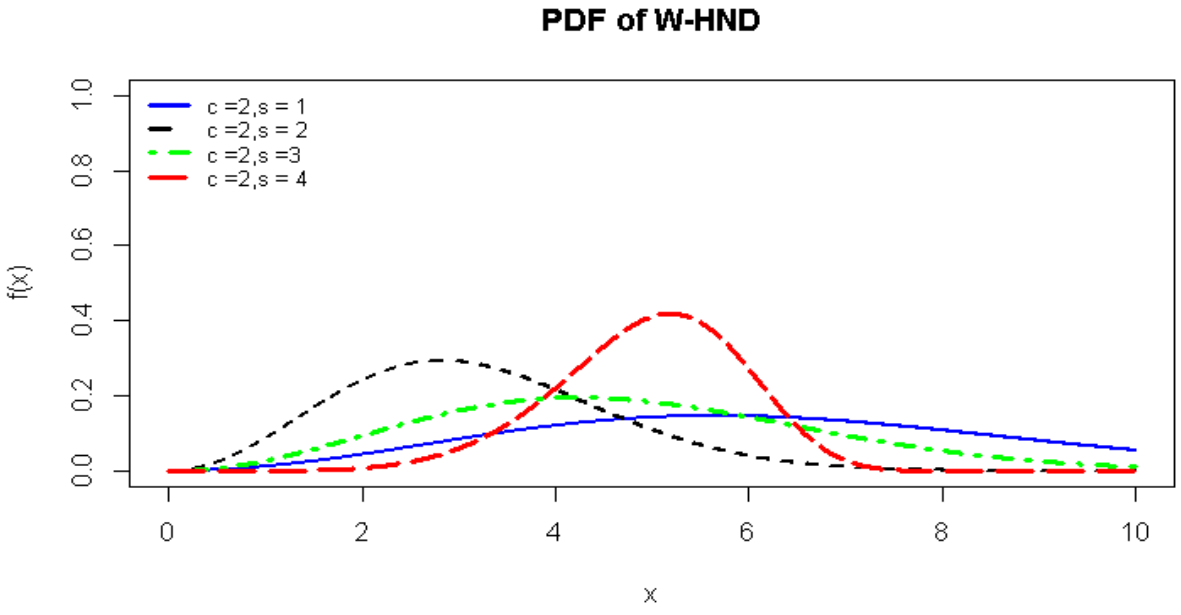


Fig. 1: PDF of Weighted-Halfnormal at $\hat{c}$ =2

From the figure 1 above, we can say that the distribution is heavily tail and highly skewed, and the higher the value of σ , the lower the mode and the more the distribution is tending to symmetry..

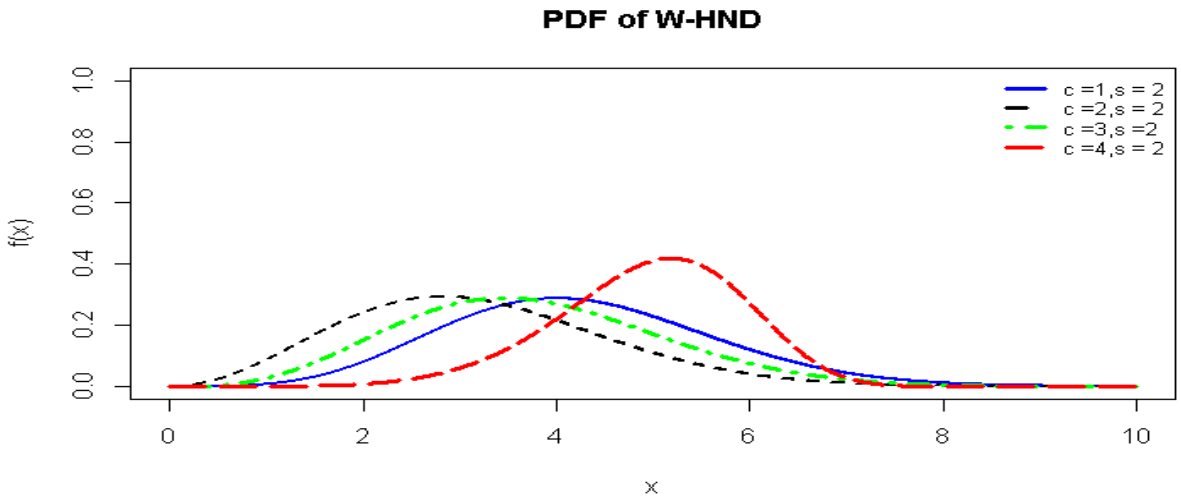


Fig. 2: **PDF of W-HND at** $\hat{\sigma}$ **=2**

From the fig 2 above, we can say that the distribution assumes symmetry as $(\hat{c})$ increases. We can deduce that at $\hat{c} = 4$ and $\hat{\sigma} = 2$ (fixed), the tail is reduce and the distributing tending to symmetry and behave like an approximate normal distribution. The $\hat{c}$ (weighted parameter) therefore influence the behavior of the distribution

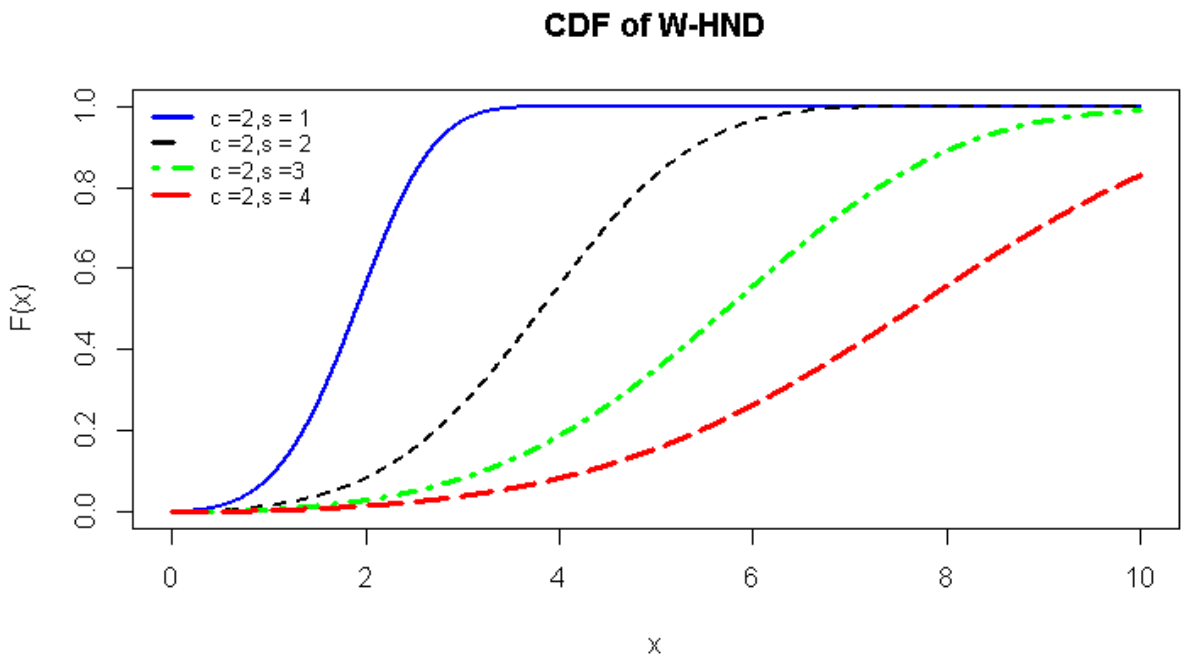


Fig. 3: The CDF of Weighted-HalfNormal at $\hat{c}$=2

With c (weighted parameter) fixed and the shape parameter $\hat{\sigma}$ increasing, the more there is significance stochastic dominance. The weighted parameter lead to more stochastic dominance at its large value.

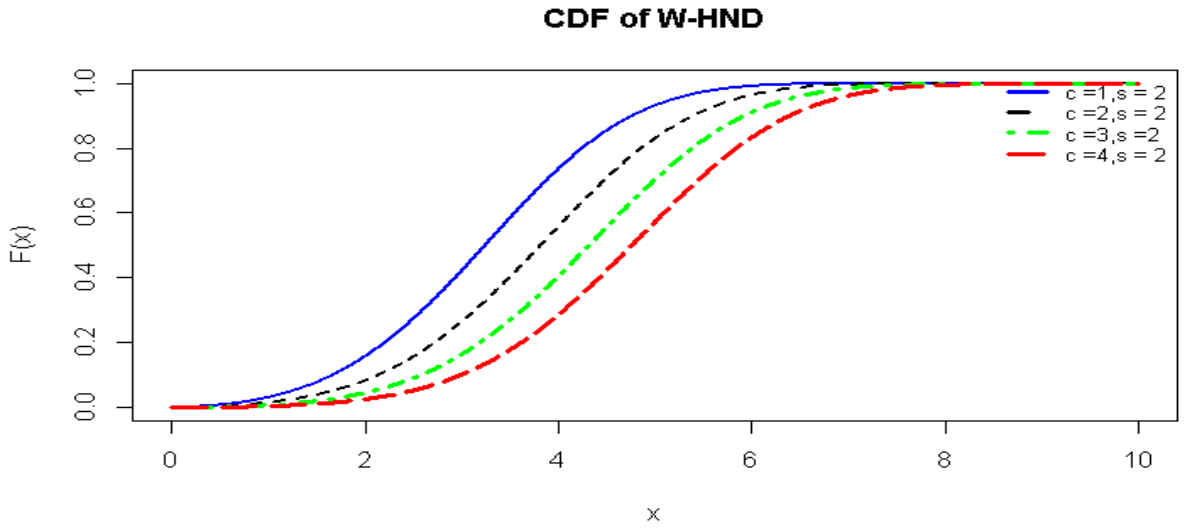


Fig. 4: The CDF of Weighted-HalfNormal at $\hat{\sigma}$ =2

At $\hat{\sigma}$ =2 and increasing$\hat{c}$, the CDF exhibits the property of second order stochastic dominance. With fixed shape parameter ($\hat{\sigma}$), the stochastic dominance is not that significant as the case with fixed$\hat{c}$.

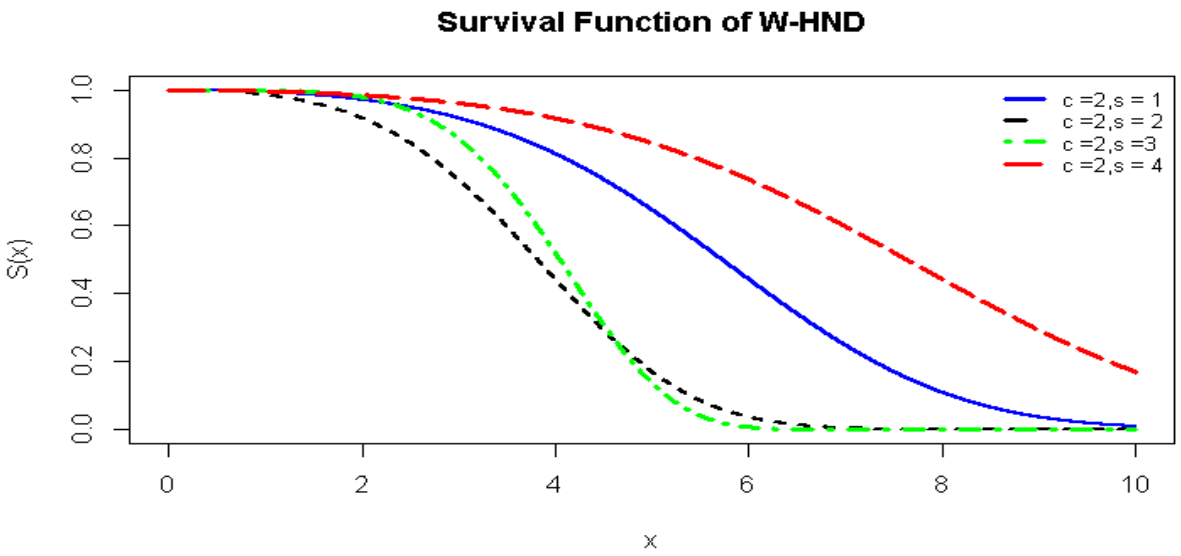


Fig. 5: Survival Function at fixed ($\hat{c}$ =2)

From figure 5 above, as the value higher value of $\hat{\sigma}$, the higher the survival probability. We can therefore conclude that, at time 6hrs, the probability of survival is approximately 80% ($\hat{c}$ **=2,** $\hat{\sigma}$ **= 4**), approximately 58% ($\hat{c}$ **=2,** $\hat{\sigma}$ **= 1**) and it is tending to zero for the model with parameter $\hat{c}$ **=2 and** $\hat{\sigma}$ **= 2**

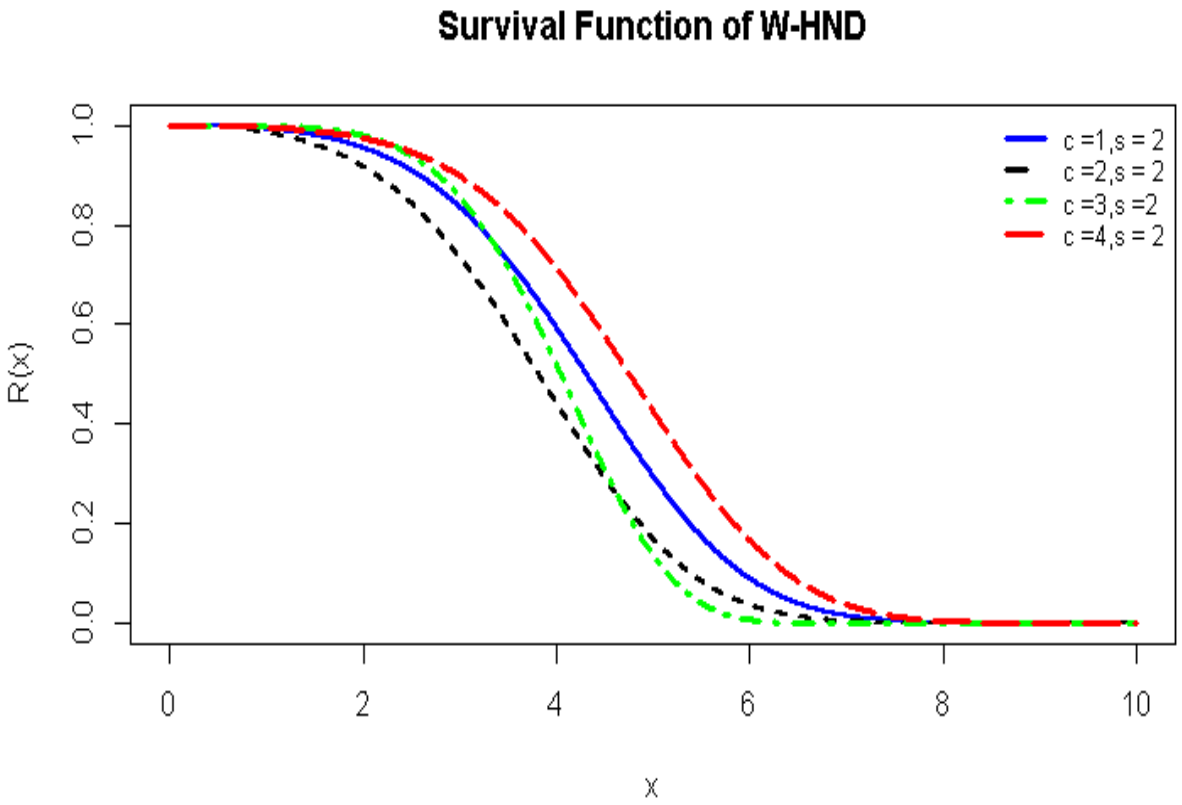


Fig. 6: Survival Function at fixed ($\hat{\sigma}$=2)

From figure 5 above, as the value higher value of $\hat{c}$, the higher the survival probability. We can therefore conclude that, at time 6hrs, the probability of survival is approximately 20% ($\hat{c}$ **=4,** $\hat{\sigma}$ **= 2**), less than 20% ($\hat{c}$ **=1,** $\hat{\sigma}$ **= 2**) and it is tending to zero for other parameter value in the models. We can establish that the survival probability increases with higher value of $\hat{\sigma}$ and moderate value of $\hat{c}$

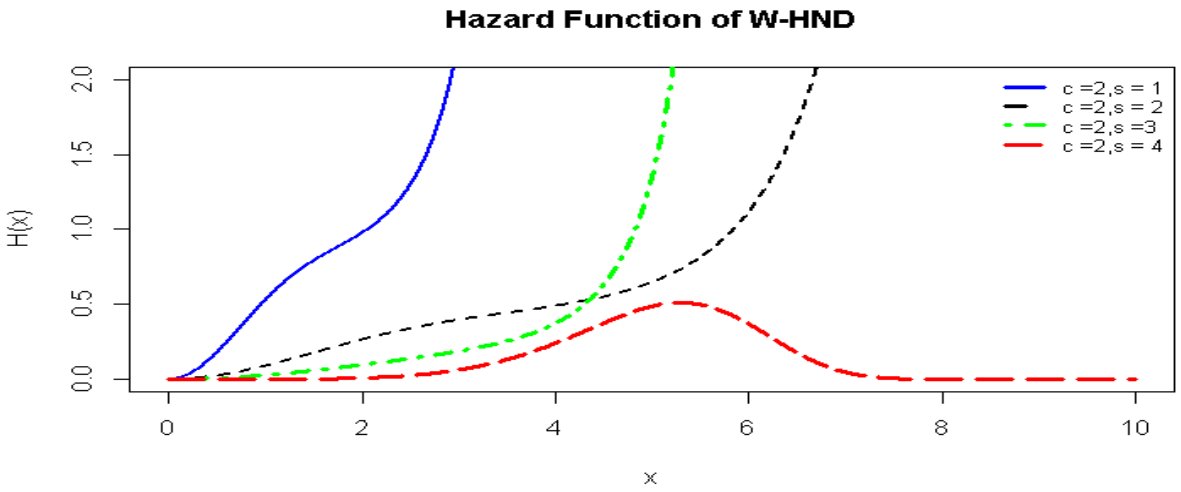


Fig. 7: The Hazard Function of W-HND at fixed ($\hat{c}$ =2)

In the Fig. 7 above the models ($\hat{c}$ **=2,** $\hat{\sigma}$ **= 1**) , ($\hat{c}$ **=2,** $\hat{\sigma}$ **= 3**) and ($\hat{c}$ **=2,** $\hat{\sigma}$ **= 2**) shows slightly increasing failure rate or increasing expected death within the first 6hrs and ($\hat{c}$ **=2,** $\hat{\sigma}$ **= 4**) shows only high failure rate at approximately 3.5 hrs and its curve flatten out as time passes.

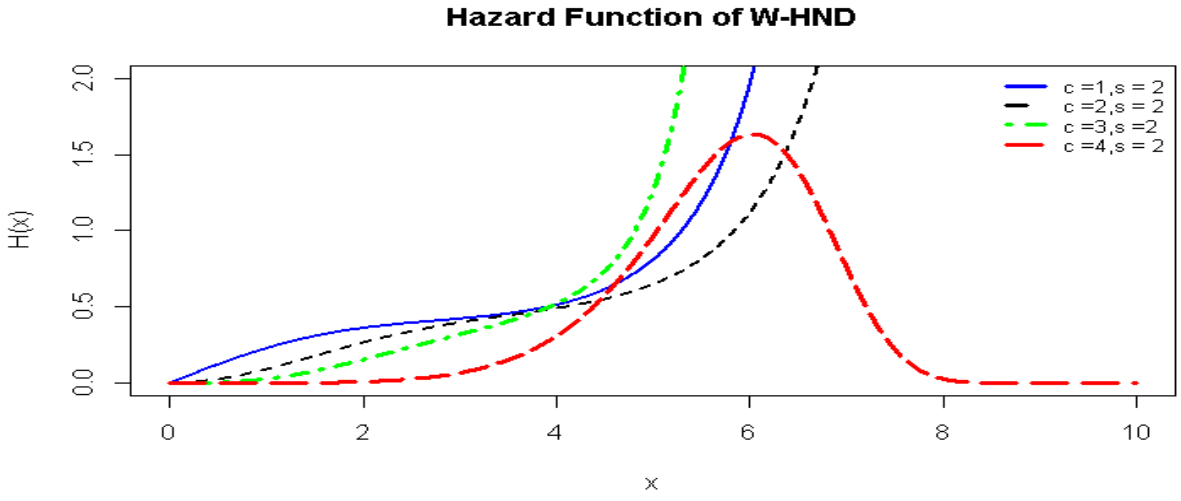


Fig. 8: The Hazard Function of W-HND at fixed ($\hat{\sigma}$=2)

The model with parameter ($\hat{c}$ =**4,** $\hat{\sigma}$ = **2**) shows that the failure rate is very high as compared to the model ($\hat{c}$ =**2,** $\hat{\sigma}$ = **4**)

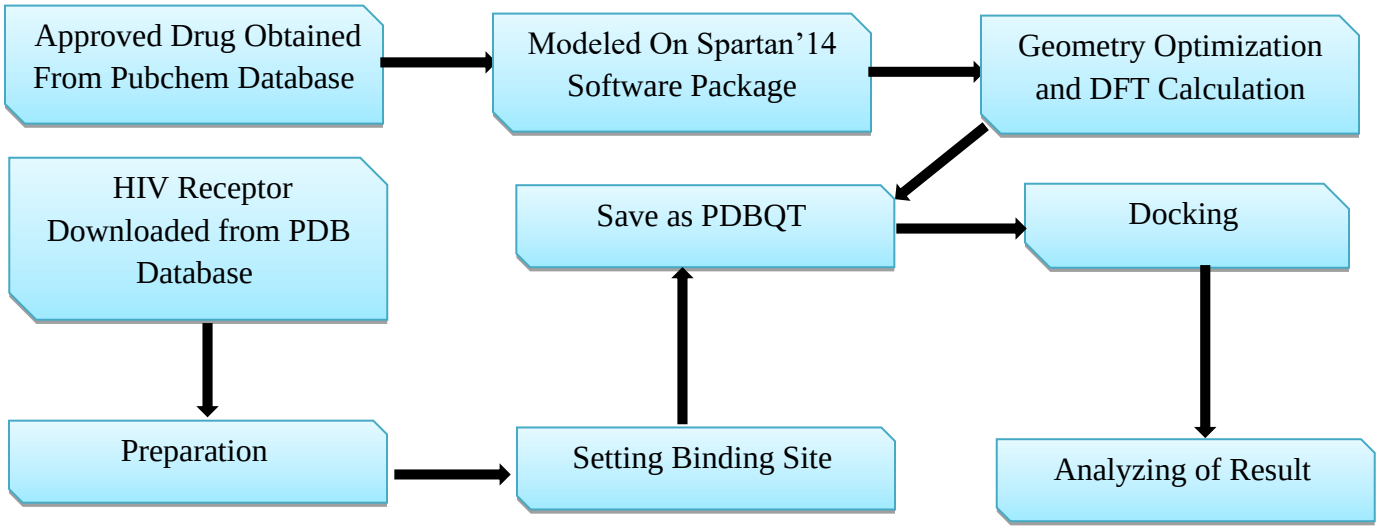
3.2 Sources of Data and Computation procedure

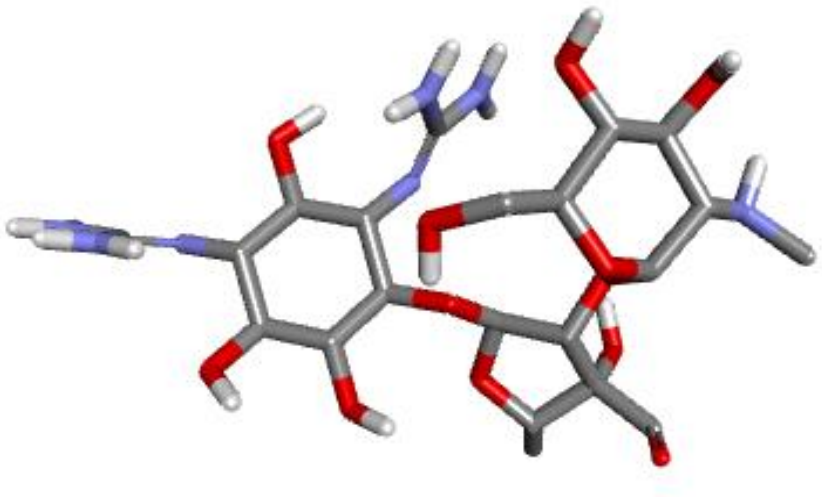
The two approved drugs (streptomycin and Abacavir) for HIV and the protein receptor with code 4CEE responsible for HIV Integrase were obtained from drug bank and protein data bank (PDB) respectively. The drugs were modeled and optimized using Spartan 14 version 1.18 molecular software package and the molecular properties were calculated using density functional theory (DFT). The calculated properties of the drugs were the molecular weight, HOMO (Highest Occupied Molecular Orbital), LUMO (Lowest Unoccupied Molecular Orbital), LogP (Partition Coefficient), HBA (Hydrogen Bond Acceptor), HBD (Hydrogen Bond Donor). In chemical and biological sciences, all these properties accounted for the toxicity and stability of drug-like compounds when binds to a particular protein receptor. This receptor was clean to remove any debris such as water molecules and ligands. The two drugs and the protein receptor were prepared by setting the binding site of the receptor and then converted to pdbqt file format before subjected to docking using simulation process. The softwares used for docking study were AutoDock Tools -1.5.6 and AutoDockVina version 1.1.2, Trott and Olson[12]. The vina binds the prepared drugs, Abacavir and Streptomycin, to the active sites of the receptor and the simulated results in nine different binding pose were obtained in kcal/mol (Table 1) which is the binding energy of the drug to the HIV protein receptor. Discovery studio and Edupymol were used to view the interactions of the drug receptor complex. The simulated results were then used to calculate the inhibition constant (k_i) using the following equations.

$$k_i = e^{-\Delta G/RT} \tag{36}$$

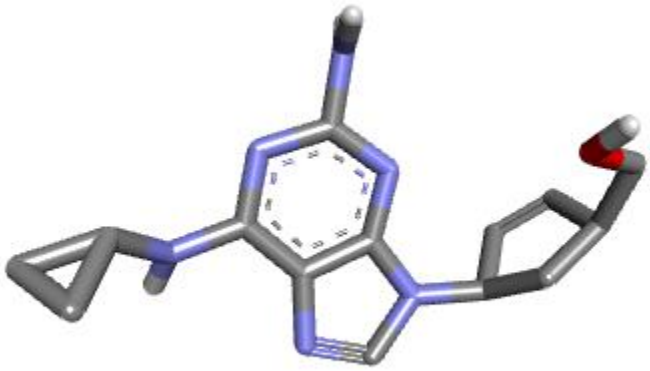
Where k_i is the inhibition value of the drug to the protein receptor (4CEE) measured in μM, ΔG is the binding energy in kcal/mol, R is ideal gas constant and T is room temperature in Kelvin.

Docking Algorithm





(a) Optimized Model for Streptomycin

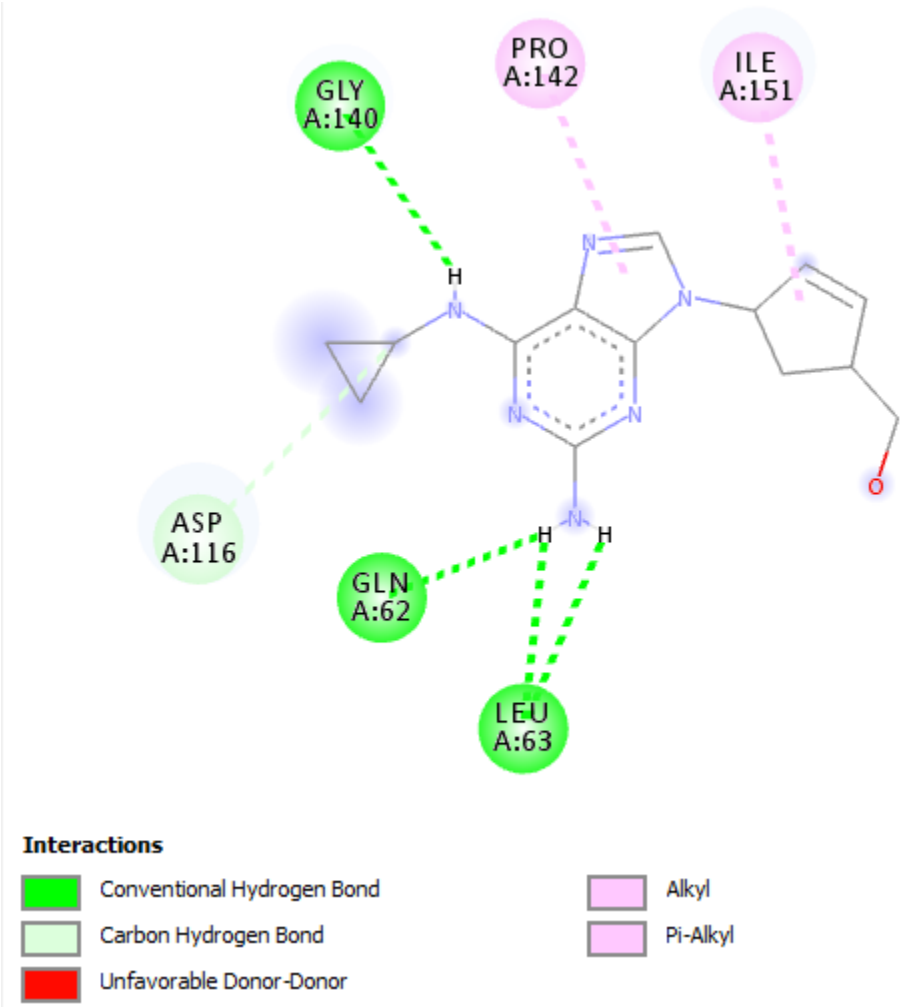


(b)Optimized Model for Abacavir

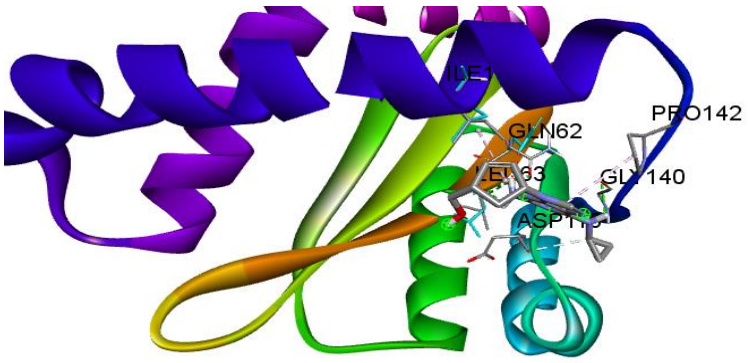
Fig. 9: The (a) and (b) above represent Ligands (Streptomycin and Abacavir) used for Molecular Docking

3.2.1 Result Visualization

Within the context of Drug – receptor interaction complex, Hydrogen bonding is very important in predicting the specificity and binding potential of the drug, Elekofehinti et al.[13] and Adegoke et al. [14]. From the 3D model (Fig 10) of Abacavir – HIV receptor complex, we could observe the binding of the drug to the active site and the 2D representation shows the amino acids at the binding pocket that contributed to the interactions. Gly140, Gly62 and Leu63 interact with hydrogen bonding to form a complex with the drug in Abacavir – HIV receptor while in Streptomycin – receptor complex as shown in figure 11, Glu152, Asn155, Cys65, Asp64, Asp116, Gly140, Pro142, Leu63, Gln62 were all involved in hydrogen bond interaction showing more of hydrogen bonding formation in streptomycin drug complex as compared to Abacavir. Vander waals, Carbon hydrogen, Alkyl bonding type also participate in the interaction leading to the potency of the two drugs on HIV integrase receptor.

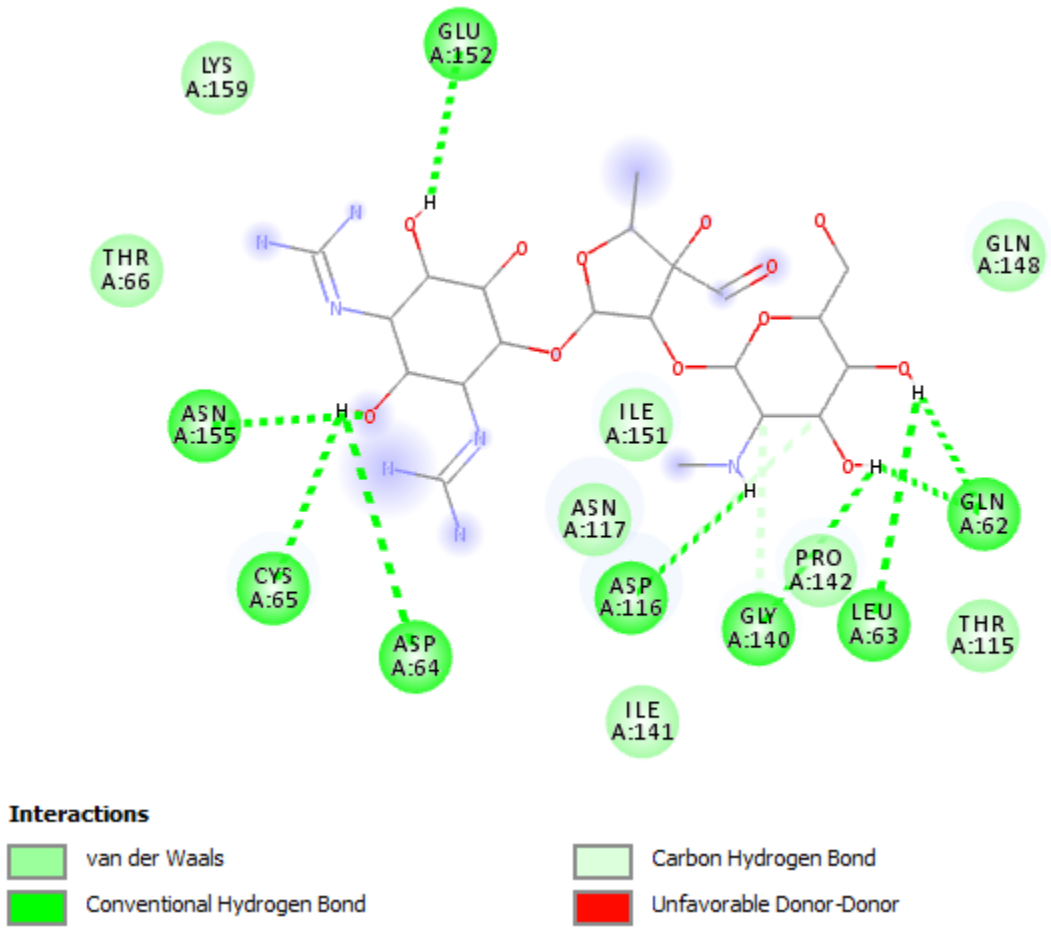


(a) 2D Interaction of Abacavir – 4CEE Complex

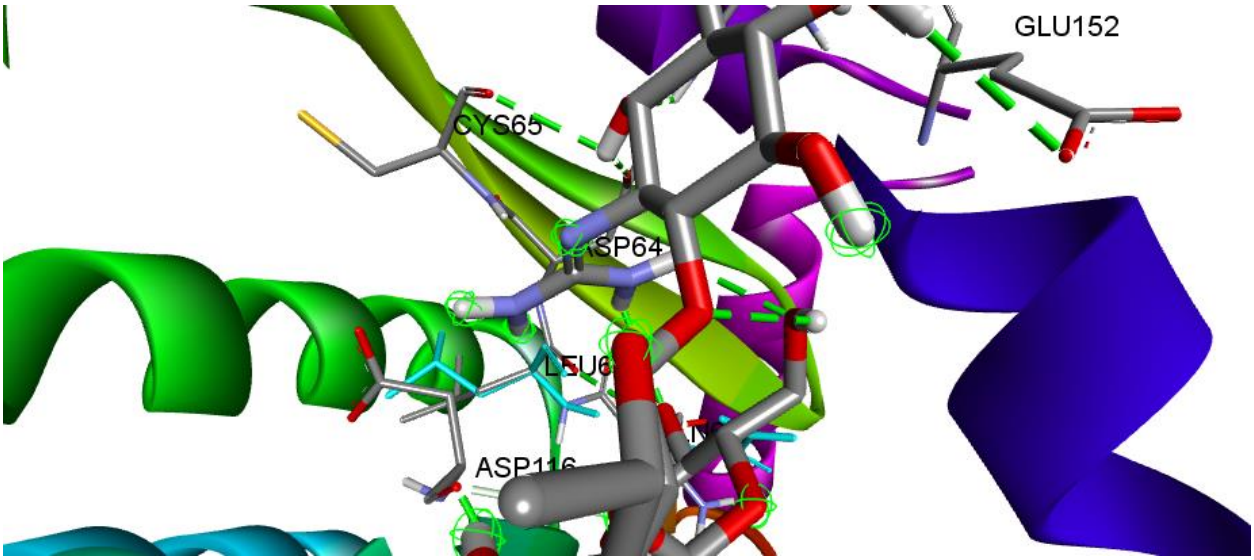


(b) Abacavir at the Binding Pocket

Fig. 10: the (a) and (b) show their respective interaction with HIV protein receptor



(a) 2D Interaction of Streptomycine – 4CEE Complex



(b) Streptomycin at the Binding Pocket

Fig. 11: the (a) and (b) show their respective interaction with HIV (4CEE) protein receptor

Table 1: Docking Score Results

	Streptomycin		Abacavir	
Receptor Code	E_1 (kcal/mol)	k_i(μM)	E_1 (kcal/mol)	k_i(μM)
4CEE	-8.0	1.0136	-6.3	1.0107
	-7.9	1.0134	-6.1	1.0103
	-7.4	1.0126	-6.0	1.0102
	-7.4	1.0126	-6.0	1.0102
	-7.3	1.0124	-5.9	1.0100
	-6.9	1.0117	-5.6	1.0095
	-6.8	1.0115	-5.6	1.0095
	-6.7	1.0114	-5.5	1.0093
	-6.6	1.0112	-5.5	1.0093

From Table1, the binding affinity (kcal/mol) for Streptomycin and Abacavir show that Streptomycin has better binding Affinity than Abacavir. Since the inhibition constants of Streptomycin are all larger than that of Abacavir at all the binding modes, then we can conclude that Streptomycin will serve as a good inhibitor against HIV receptors (Protein Receptors)

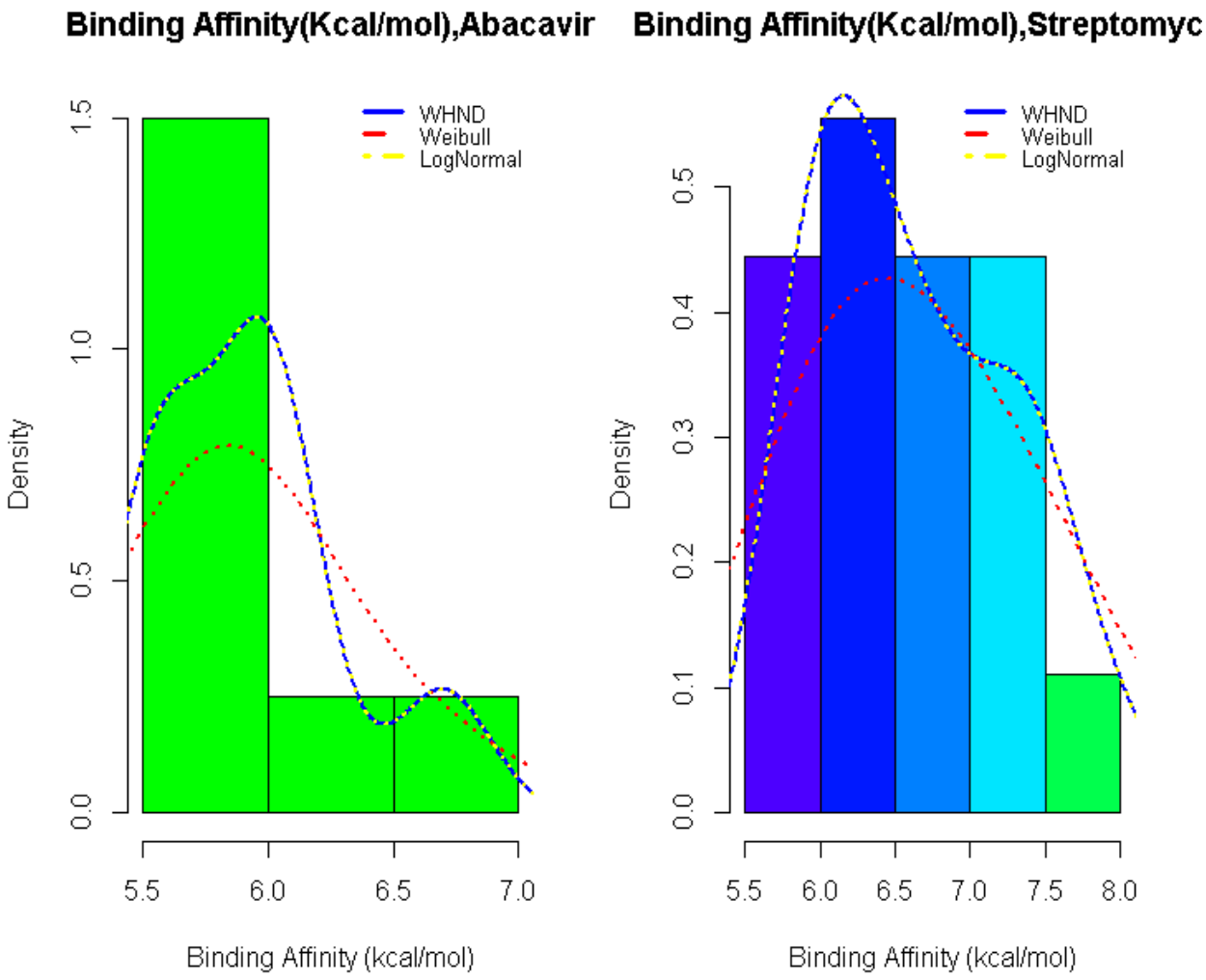


Fig. 9: Accessing distribution of the data obtained via molecular dynamics simulation

From the fig 12 above, we can say that the data (Abacavir and Streptomycin) approximately follow the distribution under investigation; we can therefore fit those distributions to the data. We can conclude from Figure 9 that W-HND and other two selected models reasonably follow the data and we can fit those models(W-HND, Lognormal and Weibull) to the data on binding affinity obtained via in-silico study of Streptomycin and Abacavir.

Table 2: Parameter Estimates

ABACAVIR		
W-HND:	Performance Criteria	Descriptives
$\hat{c} = 59.7212(6.8792)$ $\hat{\sigma} = 0.80491\ (0.1261)$	AIC=8.0244 BIC =6.2044	Mean=6.246396 Std = 0.5679749 CV= 9.092842%
LogNormal:		
$\hat{\mu}_l$= 1.7773936(0.020981) $\hat{\sigma}_l$= 0.0593426(0.01481670)	AIC=9.90476 BIC = 10.10936	Mean= 5.924844 Std=0.3519056 Cv = 5.9395%
Weibull:		
$\hat{\alpha}$= 15.273579(3.8245020) $\hat{\beta}$= 6.104649(0.1504712)	AIC =12.55563 BIC = 12.71451	Mean=5.898374 Std=0.474083 Cv = 8.03752%
STREPTOMYCIN		
W-HND:		
$\hat{c} = 56.29181(18.9812)$ $\hat{\sigma} = 0.87804\ (0.1467)$	AIC = 37.1863 BIC = 33.8163	Mean=6.61708 Std = 0.619500 CV= 9.362157%
LogNormal:		
$\hat{\mu}_l$=1.885238(0.0218828) $\hat{\sigma}_l$= 0.0928386(0.015465)	AIC = 39.3923 BIC = 35.3923	Mean= 6.616378 Std=0.6155 Cv = 9.3039%
Weibull:		
$\hat{\alpha}$= 10.98110(1.9402360) $\hat{\beta}$= 6.90999(0.157454)	AIC = 40.47537 BIC = 42.25613	Mean=6.599139 Std = 0.726667 CV= 11.01154%

From the Table 2 above, we can deduce that Weighted Halfnormal Distribution (W-HND) provide the best fit to both data (Abacavir and Streptomycin) in capturing the binding

effect of the two selected Ligands on HIV receptor. The lowest value of both performance comparison criteria, that is, AIC (Akaike Information Criterion) and Bayesian Information Criterion (BIC) validated W-HND as the best candidate model as compared to LogNormal and Weibull distribution. In the case of Abacavir (binding affinity data) obtained via molecular docking and simulation of molecular dynamics, we observed that AIC (8.0244) and BIC (6.2044) for W-HND are lower than that of Weibull ((AIC, 12.55563) and (BIC, 12.71451)) and LogNormal ((AIC, 9.90476) and (BIC, 10.10936)). Also, in the case of Streptomycin data, the performance comparison criteria ((AIC, 37.1863), (BIC, 33.8163)) for W-HND is lower than that of LogNormal ((AIC, 39.3923), (BIC, 35.3923)) and Weibull distribution ((AIC, 40.47537) , (BIC, 42.25613)). For this, W-HND outperforms Weibull and LogNormal in modeling the molecular dynamic simulated data. Streptomycin has the highest average binding effect 6.61708 kcal/mol, then we can say that the binding effect of streptomycin has dominated Abacavir (6.246396 kcal/mol) from W-HND point of view. Then Streptomycin has provided utmost docking effect and impacted greatly in lowering the activities of HIV protein receptor: Streptomycin serves as good inhibitor as compared to Abacavir. For this reason, Streptomycin is a better drug candidate as compared to Abacavir.

4. Conclusions

Since Streptomycin has higher mean binding effect than Abacavir, this showed that Streptomycin has higher binding affinity than Abacavir and we can conclude that this research showed that Streptomycin has better inhibition activities as compared to Abacavir, thus it has more efficacy when docking HIV integrase. The best candidate model (W-HND) is validated using performance comparison criteria (AIC and BIC); the W-HND is a better model as compared to Lognormal and Weibull model. We therefore conclude that W-HND is a better model of all the candidate models and Streptomycin showed better efficacy and higher binding effect as compared to Abacavir.

References

[1] T.L. Joseph, V. Namasivayam, V. Poongavanam, S. Kannan, In silico Approaches for Drug Discovery and Development, Bentham Science Publishers, 3(2017)3-74.

[2] K. Gilda, J. Yuan, P. Lai, Biopharmaceutical Drug Design and Development, Biotech .adv. J., 33(2013) 86-96.

[3] K. Luu, E. Kraynov, B. Kuang, P. Vicini, W. Zhong, Modeling, Simulation and Translation Framework for the Preclinical Development of Monoclonial Antibodies, AAPS J., 15(2013) 551-558.

[4] World Health Organization, Consolidated Guidelines on the Use of Antiretroviral Drugs for Treating and Preventing HIV Infection, 2nd edition, 2016.

[5] R.A. Weiss, A.G. Dalgleish, C. Loveday, D. Pillay, Human Immunodeficiency Viruses. In: Zuckerman AJ, Banatvala JE, Pattison JR, Griffiths PD, Schoub BD (eds), Principles and Practice of Clinical Virology. 5th ed. Chichester: John Wiley & Sons Ltd, 2004, pp721-57.

[6] M.L. Kalish, N.D Wolfe, C.B. Ndongmo, Central African Hunters exposed to Simian Immunodeficiency Virus, Emerg. Infect. Dis. 11(2005)1928-30.

[7] L. Mak, S. Liggi, L. Tan, K. Kusonmano, J. Rollinger, A. Koutsoukas, Anti-cancer Drug Development: Computational Strategies to Identify and Target Proteins Involved in Cancer Metabolism, Current Pharma. Design J., 19(2013) 532-77.

[8] G.P. Patil, C.R. Rao, Weighted Distribution: A Survey of Their Application, InP.R.Krishnaiah (Ed.), Application of Statistics, North Holland Publishing Company ,1977. Pp. 383-405.

[9] D. Sanku, D. Tanujit, M.Z Anis, Weighted Weibull Distribution: Properties and Estimation, J. Stat. Theory Pract., 00(2014)1:16.

[10] M.M.E Abd El-Monsef, S.A.E. Ghoneim, The Weighted Kumaraswamy Distribution, Information, 18(2015) 3289-3300.

[11] N.M. kilany, Weighted-Lomax Distribution, Springerplus, 5(2018)18-62.

[12] O. Trott, A. Olson, AutoDockVina: Improving the Speed and Accuracy of Docking with A New Scoring Function, Efficient Optimization and Multithreading, J. Comput. Chem., 31(2010)455-461.

[13] O. Elekofehinti, C. Oluwamodipe, P. Jean, B. Oluwaseun, F. Ayodeji, D. Damilare, I. Opeyemi, I. Yetunde, J. Ige, B. Jao, Discovery of Potential Visfatin Activators Using In-silico Docking and ADME Predictions A Therapy for Type-II Diabete, J. Basic App. Sci., 4(2018)35-43.

[14] A.B. Adegoke, A. Maradesa, H. Afolabi, Statistical Analysis of the Inhibitory Activities of Triterpenoid Derivatives against Two Selected Diseases., Int. J. Res. Innovation App. Sci. (IJRIAS), 5(2019) 130-135 .