Estimating Packed Cell Volume Levels Of HIV Positive Patients By Age, Duration Of Infection And Sex

Oyeka I.C.A, Awopeju K.A, Efobi C.C., Onyiaorah A.A.

Abstract: This paper investigates the impact of some selected factors affecting HIV/AIDS prevalence in Nigeria. Factors consider include sex of patients, age of patients, PCV level of the patients, duration of the infection and length of hospitalization of patients before discharge or date. Dummy variable multiple regression methods were used in which the independent variables were represented with 0s and 1s. It was found that at each age and duration of infection, the PCV levels of HIV/AIDS positive females were statistically lower than that of HIV/AIDS positive males. PCV levels were found to decrease with increase in duration of infection for all sex groups. The mean duration of HIV/AIDS infection was found to be highest for the age group less than 30 for all sex groups. Age group 40 and above tend to have the lowest mean duration of the infection, as well as, length of hospitalization of patients.

Keywords: Dummy regression, p-value, length of hospitalization, level of significance, HIV/AIDS, PCV, prevalence.

1.0 INTRODUCTION

HIV/AIDS is a major disease and cause of death in many parts of Africa. Although Africa is home to about only 14.5% of the world's population, it is estimated to be home to 69% of all people living with HIV and to 72% of all AIDS related deaths in 2009 [6]. Southern Africa is the worst affected region in Africa, as well as, the worst affected region in the world, with the epidemic reaching very high levels in Swaziland, Botswana, Lesotho, South Africa, Zimbabwe, Zambia and Namibia. By contrast, North Africa has low HIV/AIDS rates as can be seeing from table 1.

Table 1: Regional comparisons of HIV in 2009

World region	Adult HIV prevalence (ages 15–49)	Total HIV cases	AIDS deaths in 2005
Sub-Saharan Africa	5.0%	22.5 million	1.3 million
Worldwide	0.8%	33.3 million	1.8 million
North America	0.5%	1.5 million	26,000
Western and Central Europe	0.2%	820,000	8,500

Source: [7]

- *Oyeka I.C.A and Awopeju K.A, Department of Statistics, Nnamdi Azikiwe University, Awka, Anambra State*
- *Efobi C.C., Dept. of Haematology and Blood Transfusion, University of Port-Harcourt Teaching Hospital, Port-Harcourt, Rivers State.*
- *Onyiaorah A.A., Dept of Ophthalmology, Enugu State University Teaching Hospital, Enugu.*

Table 1 shows the distribution of HIV in the region of the world with Sub-Saharan Africa having the highest of 5% according to UNAID in the year 2009. In Nigeria, an estimated 3.6 percent of the population is living with HIV and AIDS [8]. Although HIV prevalence is much lower in Nigeria than in other African countries such as South Africa and Zambia, the size of Nigeria's population (around 162.5 million) means that by the end of 2009, there were an estimated 3.3 million people living with HIV[6 and 8]. Approximately 220,000 people died from AIDS in Nigeria in 2009 [6]. With AIDS claiming so many lives, Nigeria's life expectancy has declined significantly. In 2010, the overall life expectancy was only 52 years; currently it is only about 47years [7]. The state of HIV/AIDS in Nigeria and African countries call for urgent need for control measure and investigation to the right direction, in order to be among the countries in the world which has fought the fight against the deadly disease called HIV/AIDS and winning. This paper investigates the level of HIV/AIDS among the people in Nigeria with particular reference to Port-Harcourt, the capital of River State Nigeria. The factors considered are sex of patients, age of patients, PCV level of the patients, duration of the infection and length of hospitalization of patients before discharge or date.

2.0 Source of Data

The data for the research are secondary data from the record office of the University of Port-Harcourt Teaching Hospital in Port-Harcourt, Nigeria. Sex of patients, age, PCV level and duration of infection were recorded by the hospital and were collected by the researcher for research purpose.

Table 3: Percentage Distribution of HIV Positive Patients by Age, Duration of Infection and Sex

Age	**Male**	**Female**	**Total**
<30	23.1(3)	76.9 (10)	**100(13)**
30-39	58.8 (10)	41.2 (7)	**100(17)**
40+	46.7 (7)	53.3 (8)	**100(15)**
Duration			
< 2yrs	33.3(4)	66.7(8)	100.0(12)
2-3yrs	58.3(7)	41.7(5)	100.0(12)
3-4yrs	40.0(4)	60.0(6)	100.0(10)
4yrs +	45.5(5)	54.5(6)	100.0(11)
Total	**44.4(20)**	**55.6(25)**	**100.0(45)**

Observation

Table 2 shows percentage distribution of affected population with respect to sex and age. From the table, it can be observed that for age less than 30 and more than 40, more female; 76.9% and 53.3%were affected than male, 23.1% and 46.7% respectively but the percentage of male affected within the age bracket of 30 to 39 is more than that of female, 58.8% and 41.2% respectively. For the age groups, age group 30-39 has the highest number of affected people closely follow by 40+ which is an indication that age bracket 30-39 are prone to the disease than other groups. In terms of duration of infection, females have longer affected duration than males as shown in the above table. For the period less than 2 years, 66.7% females were involve with 33.3% male and for the period of 2years to 3 years, 41.7% females were involve with 58.3% male. Also, the period of 3 years to 4 years, 60% female were involve and for 4 years and above, the percentage of female was 54.4%.

TABLE 4: Mean P.C.V levels of HIV positive patients by Age, Duration of illness and sex

	Mean PCV level			**Mean Age**			**Mean Duration**		
Age	Male	Female	Total	Male	Female	Total	Male	Female	Total
<30	34.7	27.4	29.1	27	26.2	26.4	3.8	3.1	3.3
30-39	36.1	31.3	34.1	37.1	28.7	33.7	2.8	2.6	2.7
40+	30.67	29.3	29.9	52.6	54.5	53.6	3.5	3.2	3.4
Duration									
< 2years	38.3	29.6	32.5	37.5	34.4	35.42	1.33	1.48	1.43
2-3 years	33.6	29.0	31.7	39.4	53.2	45.2	2.24	2.46	2.33
3-4 years	35.3	27.7	30.7	37.5	29.5	32.7	3.53	3.03	3.23
4years +	30.0	29.9	29.9	42.8	35.2	38.6	5.7	5.3	5.5
Total	34.0	29.1	31.24	39.5	37.2	38.2	3.2	2.97	3.06

Note: The PCV level of age bracket 30-39 is the highest among the three groups, closely follow by that of age group 40+.

Table 5: Mean Age, Mean Duration at Infection and Mean PCV Level by Age and Sex of HIV Positive Patients

	Current Mean Age			**Mean Duration of Infection**			**Likely Mean Age at Infection**			**Mean PCV Level**		
Age	Male	Female	Total	Male	Female	Total	Male	Female	Total	Male	Female	Total
<30	27.0	26.2	26.4	3.8	3.1	3.3	23.2	23.1	23.14	34.7	27.4	29.1
30-39	37.1	28.7	33.7	2.8	2.6	2.7	34.35	30.44	31.0	36.1	31.3	34.1
40+	52.6	54.5	53.6	3.5	3.2	3.4	49.0	51.32	50.3	30.6	29.3	29.9
Total	39.5	37.2	38.2	3.19	3.0	3.1	36.31	34.19	35.1	34.0	29.1	31.2

	Current Mean Age			**Mean Duration of Infection**			**Likely Mean Age at Infection**			**Mean PCV Level**		
Duration of Infection	Male	Female	Total	Male	Female	Total	Male	Female	Total	Male	Female	Total
< 2years	37.5	34.4	35.4	1.3	1.5	1.4	36.2	32.9	34.0	38.3	29.6	32.5
2-3 years	39.4	53.2	45.2	2.2	2.5	2.3	37.2	50.7	42.8	33.6	29.0	31.7
3-4 years	37.5	29.5	32.7	3.5	3.0	3.2	34.0	26.5	29.5	35.3	27.7	30.7
4years +	42.8	35.2	38.6	5.7	5.3	5.5	37.1	29.9	33.13	30.0	29.8	29.9
Total	39.5	37.5	38.2	3.2	3.0	3.1	36.3	34.2	35.1	34.0	29.1	31.2

3.0 Methodology

Here, we used the method of dummy variable multiple regression to model the relationship between PCV level, age, duration of infection and sex of HIV positive patients. For this purpose patients will be categorized into three age groups: (1) less than 30 years (2) 30-39 years and (3) 40 years or more; four durations of infection (1) less than 2 years (2) 2years but less than 3 years (3) 3years but less than 4 years, and (4) 4 year or more for both male and female patients. Consistent with the dummy variable coding system (Boyle 1970, Nater and Wasserman 1974, Oyeka

1992) each parent independent variable in the model is represented by one dummy variable of 1s and 0s less than the number of its categories. Thus, age with three categories will be represented by two dummy variables x_{i1} and x_{i2}; A, for age groups less than 30years and 30-39 year respectively, duration of infection with four categories or groups will be represented by three dummy variables $x_{i1}D$, x_{i2} and x_{i3}; D for durations less than 2 years, 2years but less than 3 years, and 3years but less than 4 years respectively, and sex with two classes male and female, is represented by one dummy variable x_{i1};S for the male gender for i= 1,2,…,45. Specifically, if y_i is the PCV level of the i^{th} HIV positive patient, I = 1, 2, … , 45, we let

$$x_{i1A} = \begin{cases} 1, if\ the\ i^{th}\ patient\ is\ in\ age\ group\ 1\ (age\ less\ than\ 30\ years) \\ 0, otherwise \end{cases}$$

$$x_{i2A} = \begin{cases} 1, if\ the\ i^{th}\ patient\ is\ in\ age\ group\ 2\ (age\ \ 30-39\ years) \\ 0, otherwise \end{cases}$$

$$x_{i1D} = \begin{cases} 1, if\ the\ i^{th}\ patient\ is\ in\ group\ 1\ of\ duration\ (D \le 2 years) \\ 0, otherwise \end{cases}$$

$$x_{i2D} = \begin{cases} 1, if\ the\ i^{th}\ patient\ is\ in\ group\ 2\ of\ duration\ (2 \le D \le 3 years) \\ 0, otherwise \end{cases}$$

$$x_{i3D} = \begin{cases} 1, if\ the\ i^{th}\ patient\ is\ in\ group\ 3\ of\ duration\ (3 \le D \le 4 years) \\ 0, otherwise \end{cases}$$

$$x_{i1S} = \begin{cases} 1, if\ the\ i^{th}\ patient\ is\ male \\ 0, otherwise \end{cases} \quad \ldots . (1)$$

$for\ i = 1, 2, \ldots, 45.$

Now, a dummy variable multiple regression model expressing the dependence of PCV level y, of HIV positive patients on age duration of infection and sex represented by their dummy variables $x_{i1}; A, x_{i2}; A; x_{i1}D, x_{i2}D, x_{i3}D\ and\ x_{i1}S$ may be written as;

$$y_i = \beta_0 + \beta_1 : Ax_{i1A} + \beta_2 : A_{xi2} : A + \beta_1 : D_{xi1} : D + \beta_2 : D_{xi2} D + \beta_3 : D_{xi3} : D + \beta_1 : \beta_{xi1} : S + ei \quad \ldots\ldots\ldots\ldots (2)$$

Where

$\beta_{iS}\ are\ partial\ regression\ co - efficients\ and\ e_i\ is\ the\ error\ term$

Use of the usual Least Squares Method yields the estimated regression model

$$\hat{y}_i = b_0 + b_1 ; Ax_{i1A} + b_2 : A_{xi2} : A + b_1 : D_{xi1} : D + b_2 : D_{xi2} ; D + b_3 : D_{xi3} : D + b_1 ; S_{xi1} ; S \quad \ldots\ldots\ldots\ldots (3)$$

Where b_j is the least square estimate of $\beta_j ; j =$ 0,1,2,3. Applying equation 1 to the HIV Data on Table 1 yields the design matrix X of 1's and 0's (Table 5) required to fit equation 2 to the sample data on HIV.

Table 5: Design Matrix for the HIV Data of Table

S/N patient	PCV Level y_i	Age <30 x_{i1A}	Age 30-39 x_{i2A}	Dur<2years x_{i1D}	Dur2≤D<3 x_{i2D}	Dur3≤D<4 x_{i3D}	Sex Male x_{i1S}	Age A	Dur D	Sex S
1	27	1	0	1	0	0	0	1	1	0
2	30	0	1	0	0	0	1	2	4	1
3	32	0	0	0	0	0	0	3	4	0
4	33	1	0	0	0	0	1	1	4	1
5	36	0	1	1	0	0	1	2	1	1
6	24	0	0	0	1	0	1	3	2	1
7	29	0	0	0	1	0	0	3	2	0
8	24	0	0	1	0	0	0	3	1	0
9	27	0	0	0	1	0	0	3	2	0
10	32	1	0	0	0	1	0	1	3	0
11	27	0	0	0	0	0	0	3	4	0
12	35	0	0	0	1	0	0	3	2	0
13	36	0	1	1	0	0	0	2	1	0
14	46	0	1	1	0	0	1	2	1	1
15	27	1	0	1	0	0	0	1	1	0
16	28	0	1	0	0	1	0	2	3	0
17	30	0	1	0	1	0	1	2	2	1
18	35	0	0	0	0	1	1	3	3	1
19	38	0	1	0	1	0	1	2	2	1
20	28	0	1	0	1	0	0	2	2	0
21	30	0	1	0	0	0	1	2	4	1
22	30	0	0	0	1	0	1	3	2	1
23	28	0	1	1	0	0	0	2	1	0
24	32	0	0	0	0	1	1	3	3	1
25	42	0	1	0	1	0	1	2	2	1
26	36	0	0	1	0	0	1	3	1	1

27	31	1	0	0	0	1	0	1	3	0
28	24	1	0	0	0	1	0	1	3	0
29	34	0	0	1	0	0	0	3	1	0
30	27	0	0	0	0	0	1	3	4	1
31	35	0	1	0	0	0	0	2	4	0
32	34	0	1	0	1	0	1	2	2	1
33	17	1	0	0	0	1	0	1	3	0
34	40	0	1	0	0	1	1	2	3	1
35	30	0	1	1	0	0	0	2	1	0
36	30	1	0	0	0	0	0	1	4	0
37	37	1	0	0	1	0	1	1	3	1
38	26	1	0	0	1	0	0	1	2	0
39	35	0	1	1	0	0	1	2	1	1
40	34	1	0	0	0	1	1	1	3	1
41	34	0	1	0	0	1	0	2	3	0
42	28	1	0	0	0	0	0	1	4	0
43	27	1	0	0	0	0	0	1	4	0
44	31	1	0	1	0	0	0	1	1	0
45	30	0	0	0	0	0	1	3	4	1

Using the data of Table 5, we fit the dummy variable regression model consistent with equation 3 to obtain the estimated regression equation for the HIV patient as

$$\hat{y}_i = 27.1 - 0.10_{xi1}: A + 3.28_{xi2}: A + 2.36_{xi1}: D + 0.71_{xi2}; D + 0.95_{xi2}; D + 4.34_{xi1} S \quad \ldots\ldots\ldots\ldots (4)$$

To estimate the PCV level of a randomly selected HIV positive patient at a specified level of any parent independent variable, we set the dummy variable, representing that level of the parent independent variable equal to and all other dummy variables equal to 0 in equation 4. Thus, the estimated PCV level of a male HIV positive patient aged less than 30 years who has had the infection for less than 2 years is obtained by setting.

$x_{1i}, A = x_{i1}; D = x_{i1}; S = 1 \; and \; x_{i2}; A = x_{i2}; D = x_{i3}; D = 0$ in equation 4 yielding

$$\hat{y}_i = 27.1 - 0.10 + 3.26 + 4.31 = 34.57$$

The marginal PCV level of all HIV positive male patients who have had the infection for less than two years irrespective of the age is obtained by setting $x_{i1}; A = x_{i2}; A = x_{i1}; D = x_{i1}; S = 1 \; and \; x_{i2}; D = x_{i3}; D = 0$ in equation 4 yielding $\hat{y}_i = 27.1 - 0.10 + 3.28 + 3.26 + 4.31 = 37.85$ Other joint and marginal PCV levels are similarly estimated and the results are shown in the table 7

Table 7: Estimated Joint and Marginal PCV Levels of HIV Positive Patients by Age, Duration of Infection and Sex

Duration of Infection

	<2yrs		**2yrs <3yrs**		**3yrs<4yrs**		**≥4yrs**		**Marginal**	
Age	M	F	M	F	M	F	M	F	M	F
<30 yrs	33.7	29.36	32.05	27.71	32.29	27.95	31.34	27	35.36	31.02
30-39	37.08	27.74	35.43	31.09	35.67	31.33	34.72	30.38	38.74	34.40
≥40yrs	33.8	29.46	32.15	27.81	32.39	28.05	31.44	27.1	35.46	31.12
Marginal	36.98	32.64	35.33	30.99	35.51	31.23	34.62	30.28		

It can be observed from table 6 that at each age and duration of infection, the estimated PCV level of females is statistically lower than that of males, PCV levels also tend to reduce with increasing duration of infection. The age effect is however not clearly linear. It may also be further illuminating to examine the total direct and indirect effects of age, duration of infection and sex on PCV Levels of HIV Positive patients. The total or absolute effect of a parent independent variable on a dependent variable is the simple regression coefficient of the dependent variable regressing on the independent variable. The direct effect is the effect of the parent independent variable through its representative dummy variables on the dependent variable and is the weighted sum of the partial regression coefficients or effects of these representative dummy variables on the dependent variable with the weight being the simple regression coefficients of the dummy variables on the parent independent variable. The indirect effect of a parent independent variable through the mediation of its representative dummy variables in the presence of other parent independent variables in the regression model and is measured as the difference between its absolute and direct effects. Using the data of table 5, we estimate the absolute effect of age on the PCV Levels of HIV positive patients as B_A= 0.643. The direct effect is b_A= (-0.500)(-0.10)+(0.000)(3.28)= 0.050 Hence, the indirect effect of age is B_A-b_A= 0.643-0.050 = 0.593 Thus, if the differential effects of sex and duration of infection were not controlled changing age by one year is on the average likely to change the PCV Level of HIV positive patients by

0.643units. However, holding these other independent variables at constant levels has the effect of changing PCV Level by about 0.05units for every one year change in age. The remaining 0.593units seems to be accounted for by differences in other factors including sex of patient and duration of infection. The absolute direct and indirect effects of sex and duration of infection are similarly estimated and the results are presented in table 8.

Table 8: Estimated Absolute Direct and Indirect effects of Age, Duration of Infection and Sex on PCV Levels of HIV Positive Patients

Estimated Effects			
Factor	**Absolute**	**Direct**	**Indirect**
Age	0.643	0.050	0.593
Duration of Infection	-0.874	-0.691	-0.183
Sex	4.87	4.34	0.53

Expectedly, as may be noticed from table 5, increasing the duration of infection has the effect of reducing the PCV Level of HIV Positive Patients.

4.0 FINDINGS AND CONCLUSION

It was observed that for age less than 30 and more than 40, more female; 76.9% and 53.3% were affected than male, 23.1% and 46.7% respectively but the percentage of male affected within the age bracket of 30 to 39 is more than that of female, 58.8% and 41.2% respectively in the region considered. For the age groups, age group 30-39 has the highest number of affected people closely follow by 40+ which is an indication that age bracket 30-39 are prone to the disease than other groups. In terms of duration of infection, females have longer affected duration than males. For the period less than 2 years, 66.7% females were involve with 33.3% male and for the period of 2years to 3 years, 41.7% females were involve with 58.3% male. Also, the period of 3 years to 4 years, 60% female were involve and for 4 years and above, the percentage of female was 54.4%. In the study of PCV Level, age, duration of infection and sex of HIV positive patients, we used the method of dummy variable multiple regression to model the relationship. It was observed from table 6 that at each age and duration of infection, the estimated PCV level of females is statistically lower than that of males, PCV levels also tend to reduce with increasing duration of infection. The age effect is however not clearly linear.

References

[1]. Barreto, Humberto; Howland, Frank (2005). "Chapter 22: Dummy Dependent Variable Models". Introductory Econometrics: Using Monte Carlo Simulation with Microsoft Excel. Cambridge University Press. ISBN 0-521-84319-7

[2]. Draper, N.R.; Smith, H. (1998) Applied Regression Analysis, Wiley. ISBN 0-471-17082-8 (Chapter 14)

[3]. Health Reform Foundation of Nigeria (HERFON) (2007, August) 'Impact, challenges and long-term implications of antiretroviral therapy programme in Nigeria'.

[4]. National Agency for the Control of AIDS (NACA) (2009, December) 'National HIV/AIDS strategic framework (NSF) 2010-15'

[5]. **UNAIDS** (2010): 'UNAIDS report on the global AIDS epidemic'. http//www.unaids.org/globalreport/Global_report.html

[6]. **UNDP** (2011) 'Human Development Report 2011'. http//www.beta.undp/en/home/librarypage/hdr/humandevelopmentreport2011.html

[7]. **UNGASS** (2010) 'UNGASS Country Progress Report: Nigeria'

[8]. Suits, Daniel B. (1957). "Use of Dummy Variables in Regression Equations". Journal of the American Statistical Association 52 (280): 548–551. JSTOR 2281705.

[9]. World Bank (2008): 'Nigeria receives US million additional funding for HIV/AIDS project program'. http//web.worldbank.org/website/external/extabout us/IDA/O. Accessed 13th September 2012.

[10]. World Bank (2009, August): 'IDA at Work: Expanding Systems to Fight HIV/AIDS'. http//web.worldbank.org/website/external/extabout us/IDA/O. Accessed 13th September 2012

[11]. WHO/UNAIDS/UNICEF (2011): 'Global HIV/AIDS Response: Epidemic update and health sector progress towards Universal. http//www.who.int./hiv/pub/progress_report2011/en/index.html. Access 2011'

Appendix
Data Presentation and Description

TABLE 2: Data on PCV levels Age, Duration of infection and sex of HIV positive patients

MALE				FEMALE				MALE				FEMALE		
PCV	AGE	DURATION INFECTION		PCV	AGE	DURATION INFECTION		PCV	AGE	DURATION INFECTION		PCV	AGE	DURATION INFECTION
30	39	6.3		27	27	1.3		34	32	2.5		24	27	3.2
33	26	5.4		32	40	5.3		40	38	3.5		34	46	1.4
36	31	0.5		29	58	2.0		37	28	2.4		35	32	5.0
24	71	2.0		24	62	1.8		35	30	1.3		17	28	3.0
46	32	1.8		27	63	2.6		34	27	3.6		30	30	1.7
30	35	2.4		32	27	3.0		30	65	5.0		30	28	4.0
35	45	3.8		27	61	7.0						26	45	2.6
38	33	2.2		35	61	2.7						34	30	3.0
30	39	5.0		36	32	1.8						28	25	4.6
30	45	2.1		27	26	1.8						27	25	6.0
32	40	3.2		28	36	3.0						31	20	1.0
42	32	2.0		28	39	2.4								
36	57	1.7		28	32	1.0								
27	45	7.0		31	29	3.0								