Does **being employed reduce your probability of starting antiretroviral therapy (ART)** one month after testing HIV-positive?

Methods

Under a retrospective cohort study design, in order to test for the causal association between our exposure - employment and outcome - the probability of an individual from Western Cape, South Africa starting ART one month after testing HIV-positive, the main method of analysis will be to apply a logistic regression analysis between both variables, fully adjusted for well-considered confounders, and further stratified for appropriately identified effect modifiers.

To identify confounders, the following two methods where employed. Firstly, an odds ratio for the independent relationship between the variable under review with respect to exposure and then separately the outcome would be generated. As part of this any variable which may be strongly associated close to the 5% significance value ($P<0.05$) for both exposure and outcome would be considered as a potential confounder. Secondly, the variable under review would be further tested using the Mantel Haenszel method (MH method), with the initial identification of a variable as a potential confounder if found to alter the odds to a measureable difference beyond the crude odds ratio for the relationship between exposure and outcome (OR 0.77 95% CI:0.56-1.01). The chi-squared p-value for this test may also be consulted to support this decision. As part of this confounding investigation missing observations for four variables (referred to in Table 1) were also taken into consideration. For variables which there was a strong independent association with the exposure and the outcome but a very weak change to the MH method odds ratio **(**Mhodds), when compared to the crude odds ratio (OR) for the relationship between exposure and outcome (OR 0.77 95% CI:0.56-1.01), these variables would still be considered as confounders and included in the fully adjusted logistic regression model.

To identify variables as effect modifiers, the MH method would be adopted and the test for homogeneity consulted. A $p-value<0.05$ would be a marker with which to conclude a strong likelihood that such variable under review demonstrates some effect modification for the association between our study exposure and outcome.

Considerations to regroup variables were appropriately considered. CD4 count at enrolment was the variable with the greatest potential to be regrouped. On review the CD4 count variable was not regrouped for two reasons. The first reason because CD4 count was one of the four variables that presented with missing observations (N=45). Secondly the difference of the outcome result between the class <=100 and 101-350 suggest that groups are not

similar enough with respect to outcome. Further reducing the confidence with which to combine both groups.

The intention with this method therefore is by controlling for the strongly identified confounders and avoid multicollinearity, a full adjusted logistic regression model can be built that is able to effectively access the effect of employment on the the probability of an individual starting ART one month after testing HIV-positive.

Results

From our fully adjusted logistic regression model, employment was strong associated with a reduction in one's probability of starting ART one month after testing HIV-positive (P=0.005). Those who were employed were approximately 34% less likely when compared to those who were unemployed to self-report starting ART one month after HIV diagnosis (OR 0.66 95% CI: 0.49-0.88).

Via a univariable analysis of the dataset as described in Table 1 few observations may be made. First is that for four variables data was discovered to be missing for a number of individuals. The second observation is that a number of other variables, including participants age group (P = 0.0001), CD4 count at enrolment (P<0.0001), participant's marital status (P=0.04) and social capital status (P=0.0001) suggest a strong association with our study outcome of interest. Such potential association was further considered as a part of the our confounding effect investigation, as described above in our methods.

Via access to prior knowledge and understanding of the background for this study a number of potential confounders were suggested. Sex and age for example were considered as confounders as both may influence an individual's ability to find employment but also an individual's ability to start ART one month after HIV diagnosis because roles within society differ and often strongly linked to sex and age. Also neither sex nor age can be considered as part of the causal pathway between our exposure and outcome. At this point it is important to recognise further consultation with experts between such relationships would need to be consulted to validate these thoughts.

Marital status and rural site demonstrated a strong association independently with respect to the exposure and outcome at or below the 5% significance value (p<0.05). Yet sensitivity

analysis for our exposure and outcome via the Mhodds, for evidence of a confounder, were both weak as the odds ratio remained very close to the crude odds ratio for the association between our study exposure and outcome (OR 0.77 95% CI:0.56-1.01). Both therefore can be considered as weak confounders and so part of this study where included within the fully adjusted logistic regression odds ratio result (OR:0.66 95%CI: 0.49 – 0.88) but not the partially logistic regression adjusted odds ratio result (OR:0.67 95%CI: 0.50 – 0.89).

Discussions

From the analysis based simply on the crude odd ratio for the association between employment and whether a participant self-reported started ART one month after HIV diagnosis, it is suggestive that those in employment are approximately 23% less likely when compared to those who are unemployed to self-report starting ART one month after HIV diagnosis (OR 0.77 95% CI:0.56-1.01).

After controlling for age group, previous treatment for TB, rural site status, marital status and this 23% likelihood further increased to 34% (OR:0.66 95%CI: 0.49 – 0.88). Therefore it can be concluded that employment does in fact reduce an individual's probability of starting ART one month after testing HIV-positive.

Factoring the interaction associated with education would also suggest that not having secondary/tertiary level education makes it 65% less likely if employed that an individual self-reports starting ART one month after HIV diagnosis. This is in comparison to only 16% if the individual is employed but has an education level at secondary or higher.

Further factoring the interaction associated with age group adds further complexity. It suggest those who that are least likely to self-report starting ART one month after HIV diagnosis are those younger than 25 years old (at 71% less likely).

With reference to the results described above it seems therefore that those with the lowest probability of starting ART one month after HIV diagnosis are those under the age of 25, who are employed and achieve an education level lower than secondary education. Although caution should be taken when attempting to combine such variables it is interesting to consider why individuals with some of these variable may have a lower probability of starting ART one month after HIV diagnosis than those without. Some suggestions might be to better educate the younger and older population. If we are to look at the data again we do find that approximately 51% of 45+ have an education status lower than secondary,

compared to 16% for 30-34 and 10% for those under the age of 25 years old. Therefore suggesting the results for those under 25 may not be so related to education, completely. Potentially turning one's attention to the main household income of those under 25 years old provides more answers as the grant/pension percentage from the government is highest for the youngest age group when compared to all other age groups, 26%, 18%, 14%, 15% and 22% respectively from the youngest to oldest age group. This might also provide some suggestion to the type of employment that is linked to the younger population but not the population between the age 30-34. Equally we must not forget this association may be distorted by the impact of missing data, residual confounding and the potential impact of selection bias.

Although strongly associated it is questionable to suggest this association between employment and a participant self-reporting starting ART is generalizable to other populations dissimilar to Western Cape, South Africa. It is likely there are number of cultural nuances unseen in the data that make this association specific to this region.

To conclude therefore, it seems attempts to cater public health clinics for those in employment have not worked to reduce the probability of starting ART one month after testing HIV-positive. Recommendations therefore are potentially to conduct a prospective study to better understand this association; doing more to collect a complete dataset.

Table 1 – Univariable analysis for all study variables with crude odds ratio with respect to the study outcome (Number of people that self-reported yes for starting ART one month after HIV diagnosis) (n=1,769)

Variable		Number of individuals (%)	Number individuals for which data is missing	Number of people that self-reported yes, for starting ART one month after HIV diagnosis (%)	Crude odds ratio with respect to study outcome	p-value*
Employment status	Unemployed	763 (43)		115 (49)	1	
	employed	1,006 (57)	0	121 (51)	0.77	0.06
Sex	Female	1,081 (61)	0	150 (64)	1	
	Male	688 (39)		86 (36)	0.89	0.41
Age in years	<25	295 (17)		29 (12)	1	
	25-29	349 (20)		33 (14)	0.96	
	30-34	357 (20)	0	43 (18)	1.26	
	35-44	465 (26)		76 (32)	1.79	
	>=45	303 (17)		55 (23)	2.03	0.0001
Previous HIV test	No	627 (35)	0	86 (36)	1	
	Yes	1,142 (65)		150 (64)	0.95	0.73
Previously treated for TB	No	1, 614 (96)	90	212 (95)	1	
	Yes	65 (4)		11 (5)	1.35	0.38
Participant knows partner's HIV status	No	1,586 (90)		210 (89)	1	
	Yes	183 (10)	0	26 (11)	1.09	
						0.72
CD4 count at enrolment	<=100	59 (3)		20 (9)	1	
	101-350	552 (32)		155 (68)	0.76	
	351-500	459 (27)	45	30 (13)	0.14	
	>500	654 (38)		24 (10)	0.07	<0.0001
Participant lives in a rural area	No	1,096 (62)		133 (56)	1	
	Yes	673 (38)	0	103 (44)	1.31	
						0.06
	Formal	1,067 (64)		155 (69)	1	0.12

Participant dwelling type	Informal	589 (36)	113	69 (31)	0.78	
Participant education	None/primary	400 (23)		65 (28)	1	
	Secondary/Tertiary	1,369 (77)	0	171 (72)	0.74	0.05
Transport generally used by participant to each their local public health clinic	Minibus/taxi/bus/other	1,078 (61)		142 (60)	1	
	Walk	691 (39)	0	94 (40)	1.04	
						0.79
Participant's reason for HIV test	Routine	1,427 (81)		184 (78)	1	
	Other (felt unwell/pregnant)	342 (19)	0	52 (22)	1.21	
						0.26
Number of people participant lives with	0	260 (16)		34 (15)	1	
	1-2	596 (36)	122	87 (39)	1.14	
	3-5	571 (35)		68 (30)	0.89	
	>5	220 (13)		34 (16)	1.21	0.86
Participant's marital status	Single	1,103 (62)		138 (58)	1	
	Married/cohabiting	534 (30)		71 (31)	1.07	
	Divorced/widowed	132 (8)	0	27 (11)	1.80	0.04
Participant social capital status	Low	858 (49)		86 (36	1	
	High	991 (51)	0	150 (64)	1.77	
						0.0001
Main income of household	Full-time work	806 (46)		102 (43)	1	
	Informal employment	632 (35)	0	82 (35)	1.03	
	Grant/pension from government	331 (19)		52 (22)	1.29	
						0.21

*using Mantel Haenszel method chi-squared test

Table 2 – Mhodds, partially adjusted and fully adjusted estimates of the odds ratio for employment status with respect to whether a participant self-reported starting ART one month after HIV diagnosis estimated by logistic regression (n = 1,679)

		Mhodds adjusted for age (95% CI)	p-value*	Mhodds adjusted for previous treatment for TB[1] (95% CI)	p-value*	Logistic **partially** adjusted odds ratio (95% CI)[2]	p-value	Logistic **fully** adjusted odds ratio (95% CI)[3]	p-value
Employment status	Unemployed	1		1		1		1	
	Employed	**0.69** (0.52 – 0.92)	0.01	**0.74** (0.56 – 0.99)	0.04	**0.67** (0.50 – 0.89)	0.06	**0.66** (0.49 – 0.88)	0.005

*using chi-squared test
[1] taking into account missing data
[2] logistic regression adjusted for age group, previous treated for TB
[3] logistic regression adjusted for age group, previous treated for TB, **rural site status, marital status**

Table 3 – Stratified fully adjusted estimates of odds ratio for employment status with respect to whether a participant self-reported starting ART one month after HIV diagnosis estimated by logistic regression (n = 1,679)

	Employment status	Stratum specified adjusted OR for employment status (95% CI)	P-value for interaction
Stratified by highest level of education			
None/Primary	Unemployed	1	
	Employed	**0.35** (0.20 – 0.62) [1]	<0.001
Secondary/tertiary	Unemployed	1	
	Employed	**0.84** (0.60 – 1.19) [1]	0.33
Stratified by age group (in years)			
<25	Unemployed	1	
	Employed	**0.29** (0.11 – 0.74) [2]	0.01
25-29	Unemployed	1	

		Employed	**0.57** (0.26 – 1.25) [2]	0.16
30-34		Unemployed	1	
		Employed	**1.76** (0.86– 3.60) [2]	0.12
35-44		Unemployed	1	
		Employed	**0.55** (0.33 – 0.92) [2]	0.022
>=45		Unemployed	1	
		Employed	**0.63** (0.33 – 1.19) [2]	0.15

[1] logistic regression adjusted for age group, previously treated for TB, rural site status, marital status

[2] logistic regression adjusted for previously treated for TB, rural site status, marital status