**Table S1**: Summary estimates of characteristics of recent seroconverters residing in the demographic surveillance area, South Africa (2004-2013), from the 10 datasets where the date of HIV infection was randomly imputed.

|  |  |  |  |
| --- | --- | --- | --- |
| **Variables** | **Median** **Number, Mean, or proportion** | **Minimum** **Number, Mean, or proportion** | **Maximum** **Number, Mean, or proportion** |
| **Sample size** | 1,761 | 1,713 | 1,779 |
| **Gender** |  |  |  |
|  Female | 76.5% | 75.6% | 76.6% |
|  Male | 23.6% | 23.4% | 24.4% |
| **Age (years)** | 26.8 | 26.7 | 26.9 |
| **Education level** |  |  |  |
|  None or less than one year | 5.5% | 5.1% | 6.0% |
|  Some or completed primary | 15.4% | 14.9% | 16.3% |
|  Some or completed secondary | 74.8% | 74.2% | 75.0% |
|  *Missing* | *4.3%* | *3.7%* | *4.9%* |
| **Socio-economic status** |  |  |  |
|  Poorest  | 21.6% | 20.5% | 26.9% |
|  Poor | 26.4% | 25.7% | 28.6% |
|  Rich | 27.2% | 25.8% | 28.5% |
|  Richest | 24.5% | 20.3% | 25.6% |
|  *Missing* | *0.1*% | *0*% | *0.2*% |
| **Food security** |  |  |  |
|  Never (or some months) missed meals | 97.2% | 96.9% | 97.6% |
|  Missing meals almost every month (financial reasons) | 2.3% | 1.8% | 2.7% |
|  *Missing* | *0.6*% | *0.3*% | *1.0*% |
| **Uptake of ART by household members** |  |  |  |
|  No household members on ART | 83.7% | 83.3% | 84.0% |
|  At least one household member on ART | 16.3% | 16.0% | 16.7% |
| **Knowledge of HIV Status\*** |  |  |  |
|  Aware | 63.1% | 62.2% | 63.3% |
|  Unaware / Refused  | 37.0% | 36.7% | 37.8% |
| **Area of residence** |  |  |  |
|  Urban | 3.1% | 2.9% | 3.3% |
|  Peri-urban | 35.7% | 34.9% | 36.4% |
|  Rural | 61.1% | 60.7% | 62.0% |
| **Distance to closest health facility (km)** | 9.2 | 9.0 | 9.4 |
| **Calendar year of HIV infection** |  |  |  |
|  2004 | 4.3% | 3.7% | 4.7% |
|  2005 | 12.5% | 11.2% | 13.6% |
|  2006 | 14.2% | 13.8% | 15.3% |
|  2007 | 13.7% | 12.5% | 14.5% |
|  2008 | 13.4% | 12.3% | 14.1% |
|  2009 | 11.8% | 10.8% | 12.5% |
|  2010 | 10.8% | 9.5% | 11.9% |
|  2011 | 8.9% | 7.8% | 9.7% |
|  2012 | 7.9% | 7.2% | 9.3% |
|  2013 | 2.8% | 2.2% | 3.4% |
| \*The question about knowledge of HIV status from previous testing was not asked in 2004 and 2005 (between 277 and 317 seroconverters were excluded, depending on the imputed datasets). |

**Table S2**: Median CD4 cell count (cells/μL) at linkage, stratified by linkage time since HIV infection, with 95% bootstrapped confidence intervals as estimated using midpoint imputation or random imputation of 10 datasets.

|  |  |  |
| --- | --- | --- |
|  | **Midpoint Imputation** | **Pooled Random Imputations** |
| **Time from HIV infection to linkage-to-care** | **Median CD4 cells/μL****(95% CI\*)** | **Median CD4 cells/μL****(95% CI\*)** |
| Less than one year | 380 (370, 442) | 370 (320, 410) |
| Between one and two years | 380 (360, 430) | 380 (340, 420) |
| Between two and three years | 330 (320, 400) | 360 (290, 430) |
| Between three and four years | 320 (290, 380) | 310 (230, 390) |
| Between four and five years | 360 (310, 440) | 320 (220, 410) |
| More than five years | 260 (250, 410) | 290 (160, 430) |
| **Overall** | **350 (330, 380)** | **350 (330, 380)** |
| \*95%CI=95% confidence intervals. These are based on 9,999 bootstrap replicates. |

**Table S3**: Time from HIV infection (years) to linkage to care stratified by gender and calendar year of HIV infection, as estimated using midpoint imputation or random imputation of 10 datasets.

|  |  |  |
| --- | --- | --- |
|  | **Midpoint Imputation** | **Pooled Random Imputations** |
| **Characteristics** | **Time from HIV infection for 25% of seroconverters to link to care (95% CI\*)** | **Time from HIV infection for 50% of seroconverters to link to care (95% CI\*)** | **Time from HIV infection for 25% of seroconverters to link to care (95% CI\*)** | **Time from HIV infection for 50% of seroconverters to link to care (95% CI\*)** |
| **Gender** |  |  |  |  |
|  Females | 1.8 (1.7-2.1) | 4.1 (3.8-4.5) | 1.7 (1.5-2.0) | 4.2 (3.7-4.6) |
|  Males | 3.4 (1.8-4.5) | NA† | 3.4 (2.4-4.4) | NA† |
| **Calendar year of HIV infection** |  |  |  |
|  2004 | 3.0 (2.8-4.1) | 6.1 (4.1-8.5) | 3.7 (2.4-5.0) | 6.1 (3.7-8.5) |
|  2005 | 2.6 (2.4-3.4) | 4.8 (4.3-5.8) | 3.1 (2.4-3.8) | 5.6 (4.2-7.1) |
|  2006 | 2.1 (1.8-2.6) | 4.3 (3.4-5.4) | 2.3 (1.7-2.9) | 4.6 (3.3-6.0) |
|  2007 | 1.8 (1.5-2.2) | 3.6 (3.0-5.1) | 1.6 (1.0-2.3) | 4.0 (2.7-5.4) |
|  2008 | 1.8 (1.4-2.2) | NA† | 1.5 (1.0-2.0) | NA† |
|  2009 | 1.9 (1.5-3.0) | NA† | 1.6 (1.0-2.2) | NA† |
|  2010 | 1.8 (1.2-3.0) | NA† | 1.4 (0.6-2.2) | NA† |
| **Overall** | **2.1 (1.9-2.3)** | **4.8 (4.3-5.4)** | **2.0 (1.8-2.2)** | **4.9 (4.2-5.7)** |
| \*95%CI=95% confidence intervals. These are based on 9,999 bootstrap replicates.†NA=not available (the follow-up time was not long enough to observe 50% of seroconverters linking to care). |



**Figure S1**. Kaplan-Meier curves of time from HIV infection to linkage-to-care in rural Kwa-Zulu Natal South Africa (2004-2013) for A) all seroconverters and B) stratified by gender using midpoint imputation.

**

**Figure S2.** Pooled Kaplan-Meier estimates of the 10 randomly imputed datasets of time from HIV infection to linkage-to-care in rural Kwa-Zulu Natal South Africa (2004-2013) for A) all seroconverters and B) stratified by gender.

**Table S4**: Univariate and multivariable effect size estimates from Cox proportional hazard models of determinants of time from HIV infection to linkage-to-care. (Midpoint imputation and pooled results from the 10 randomly imputed datasets.)

|  |  |  |
| --- | --- | --- |
|  | **Midpoint Imputation** | **Pooled Random Imputations** |
| **Variables** | **Univariate** | **Multivariable\*** | **Univariate** | **Multivariable\*** |
| **HR (95% CI)** | **aHR (95% CI)** | **HR (95% CI)** | **aHR (95% CI)** |
| **Gender** |  |  |  |  |
|  Male | **0.47 (0.38-0.59)** | **0.44 (0.34-0.57)** | **0.49 (0.39-0.61)** | **0.49 (0.37-0.64)** |
| **Age** |  |  |  |  |
|  15-29 years old | 1.00 | 1.00 | 1.00 | 1.00 |
|  30-39 years old | 1.17 (0.94-1.46) | 1.09 (0.85-1.40) | 1.22 (0.96-1.55) | 1.16 (0.86-1.57) |
|  40-49 years old | **1.89 (1.53-2.43)** | **1.33 (1.00-1.78)** | **1.92 (1.53-2.40)** | **1.54 (1.14-2.08)** |
|  50+ years old | 1.19 (0.84-1.67) | 0.82 (0.52-1.28) | 1.15 (0.80-1.64) | 0.87 (0.53-1.44) |
| **Education level** |  |  |  |  |
|  None or less than one year | 1.00 | 1.00 | 1.00 | 1.00 |
|  Some or completed primary | 0.87 (0.63-1.20) | 0.68 (0.46-1.01) | 0.87 (0.61-1.25) | 0.75 (0.46-1.22) |
|  Some or completed secondary | **0.68 (0.51-0.90)** | **0.56 (0.38-0.83)** | **0.68 (0.50-0.93)** | 0.63 (0.39-1.00) |
| **Food Security** |  |  |  |  |
|  Missing meals almost every month (financial reasons) | 1.60 (0.95-2.68) | 1.40 (0.72-2.73) | 1.56 (0.88-2.79) | 1.37 (0.67-2.80) |
| **Socio-Economic Status**  |  |  |  |  |
|  Poorest | 1.00 | 1.00 | 1.00 | 1.00 |
|  Poor | 1.11 (0.87-1.42) | 1.21 (0.92-1.61) | 1.15 (0.90-1.46) | 1.21 (0.91-1.61) |
|  Rich | 1.14 (0.90-1.44) | 1.39 (1.06-1.82 | 1.14 (0.90-1.44) | 1.19 (0.75-1.90) |
|  Richest | 0.98 (0.78-1.24) | 1.24 (0.95-1.63) | 1.00 (0.78-1.29) | 1.05 (0.62-1.79) |
| **Other household members using ART** |  |  |  |  |
|  At least one (versus ‘None’) | 1.12 (0.93-1.34) | 1.18 (0.97-1.43) | 1.16 (0.96-1.41) | 1.23 (0.98-1.55) |
| **Knowledge of HIV status\*** |  |  |  |  |
|  Yes (versus ‘No / Refused’) | **1.61 (1.33-1.94)** | **1.42 (1.17-1.73)** | **1.52 (1.24-1.87)** | **1.35 (1.09-1.68)** |
| **Area of residence** |  |  |  |  |
|  Urban | 1.00 | 1.00 | 1.00 | 1.00 |
|  Peri-Urban | 0.69 (0.45-1.06) | 0.76 (0.48-1.21) | 0.65 (0.42-1.02) | 0.73 (0.43-1.24) |
|  Rural | 0.83 (0.55-1.26) | 0.83 (0.52-1.33) | 0.77 (0.49-1.21) | 0.76 (0.45-1.30) |
| **Distance to Health Facility** |  |  |  |  |
|  <2 km | 1.00 | 1.00 | 1.00 | 1.00 |
|  2 to 4 km | 1.07 (0.89-1.28) | 1.00 (0.81-1.25) | 1.04 (0.86-1.25) | 1.00 (0.79-1.26) |
|  >4 km | 1.21 (0.99-1.47) | 1.11 (0.85-1.45) | 1.16 (0.93-1.45) | 1.10 (0.84-1.44) |
| **Calendar year of HIV infection** | Strata | Strata | Strata | Strata |