AIDS

DOI: 10.1097/QAD.0000000000001737

Impact of early antiretroviral therapy eligibility on HIV acquisition: Household-level evidence from rural South Africa

Catherine E. OLDENBURG1,2, Jacob BOR3,4, Guy HARLING9, Frank TANSER4,5, Tinofa MUTEVEDZI4, Maryam SHAHMANESH4, George R. SEAGE1, Victor DE GRUTTOLA6, Matthew J. MIMIAGA1,7,8, Kenneth H. MAYER7,9,10, Deenan PILLAY4, Till BÄRNIGHAUSEN4,9,11

1Department of Epidemiology, Harvard T.H. Chan School of Public Health, Boston, MA, USA
2Francis I. Proctor Foundation, University of California, San Francisco, USA
3Department of Global Health, Boston University School of Public Health, Boston, MA, USA
4Africa Health Research Institute, UKZN, Somkhele, South Africa
5Faculty of Health Sciences, University of KwaZulu-Natal, Durban, South Africa
6Department of Biostatistics, Harvard T.H. Chan School of Public Health, Boston, MA, USA
7The Fenway Institute, Fenway Community Health, Boston, MA, USA
8Departments of Behavioral and Social Sciences & Epidemiology; The Institute for Community Health Promotion, Brown University School of Public Health, Providence, RI, USA
9Department of Global Health and Population, Harvard T.H. Chan School of Public Health, Boston, MA, USA
10Department of Medicine, Beth Israel Deaconess Medical Center, Boston, MA, USA
11Heidelberg Institute of Public Health, University of Heidelberg, Heidelberg, Germany
Running Head: ART eligibility and household HIV incidence

Author for Correspondence:
Catherine E. Oldenburg, ScD MPH
Assistant Professor
Francis I. Proctor Foundation
University of California, San Francisco
San Francisco, CA 94143
Email: catherine.oldenburg@ucsf.edu
Phone: 415-502-8843

This is an open access article distributed under the Creative Commons Attribution License 4.0 (CCBY), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.
ABSTRACT

Objectives. We investigate the effect of immediate ART eligibility on HIV incidence among HIV-uninfected household members.

Design. Regression discontinuity study arising from a population-based cohort.

Methods. Household members of patients seeking care at the Hlabisa HIV Treatment and Care Programme in rural KwaZulu-Natal South Africa between January 2007 and August 2011 with CD4 counts up to 350 cells/μl were eligible for inclusion if they had at least two HIV tests and were HIV-uninfected at the time the index patient linked to care (N=4,115). A regression discontinuity design was used to assess the intention-to-treat effect of immediate versus delayed ART eligibility on HIV incidence among household members. Exploiting the CD4-count based threshold rule for ART initiation (CD4<200 cells/μl until August 2011), we used Cox proportional hazards models to compare outcomes for household members of patients who presented for care with CD4 counts just above versus just below the ART initiation threshold.

Results. Characteristics of household members of index patients initiating HIV care were balanced between those with an index patient immediately eligible for ART (N=2,489) versus delayed for ART (N=1,626). There were 337 incident HIV infections among household members, corresponding to an HIV incidence of 2.4 infections per 100 person-years (95% CI 2.5 to 3.1). Immediate eligibility for treatment reduced HIV incidence in households by 47% in our optimal estimate (HR=0.53, 95% CI 0.30 to 0.96), and by 32-60% in alternate specifications of the model.

Conclusions. Immediate eligibility of ART led to substantial reductions in household-level HIV incidence.

Key Words: HIV incidence; South Africa; regression discontinuity; treatment as prevention; antiretroviral therapy


BETWEEN THE COMMUNITY AND INDIVIDUAL RELATIONSHIP LEVEL, THERE MAY ALSO BE UNIQUE PATHWAYS BETWEEN ART UPTAKE AND HIV INCIDENCE WITHIN HOUSEHOLDS. [17] AT A MORE PROXIMAL LEVEL THAN THE COMMUNITY, INDIVIDUALS INITIATING ART IN HOUSEHOLDS MAY
have social influence over HIV-uninfected household members. In addition to direct biological mechanisms via cohabitating sexual partners, individuals who initiate ART may be more willing to disclose their serostatus to their families[18], and may discuss HIV prevention or elements from counseling with family members, which could result in changes in HIV acquisition in households. This could result in spillover effects including changes in sexual behaviors among household members, such as increases in condom use or reductions in number of partners. Previous work has demonstrated a benefit of increasing coverage of ART among opposite-sex household members on HIV transmission[19], however the effect on all household members is unknown.

A critical issue with the identification of effects of ART in population-based surveillance cohorts is the reliance on observational data. A number of techniques meant to improve causal inference in non-randomized studies exist, each of which contain a set of assumptions for making valid inferences.[20-25] Here, we apply a quasi-experimental approach, the regression discontinuity design, to estimate the causal effect of immediate versus delayed ART initiation on HIV incidence in household members.

**METHODS**

**Participants and procedures.** Data for this analysis arose from the population-based longitudinal surveillance program conducted by the Africa Health Research Institute (AHRI).[26] The surveillance program is located in a predominantly rural community of uMkhanyakude district, KwaZulu-Natal, and has been active since 2003. It includes confidential HIV testing, household demographic data, collection of sexual history and behaviors, and relationship status. In addition to longitudinal surveillance,
data are routinely collected from the Hlabisa HIV Treatment and Care Programme, a system of public ART clinics serving the geographic area participating in the surveillance program. As the primary provider of HIV care in the area, this system captures all linkages to ART care, longitudinal CD4 counts (measured every 6 months), and dates of ART initiation. Ethical approval for data collection, linkage, and analysis was obtained from the University of KwaZulu-Natal Biomedical Research Ethics Committee. Written informed consent was obtained from all participants. This analysis was exempted from additional ethical review by the Harvard School of Public Health Institutional Review Board due to use of anonymized secondary data.

As part of the routine demographic surveillance, information is collected about living arrangements of each participant.[19] During each surveillance round, the physical place of resident for the participant (henceforth, “homestead”) is recorded. A participant cannot be a member of two different homestead at the same time. Co-residents of the same homestead were defined as participants who were residents of the same homestead during a given surveillance round. Participants could move homesteads between surveillance rounds.

Participants were eligible for inclusion in this analysis if they were HIV-uninfected and a co-resident of the homestead at the time the first HIV-infected partner linked to HIV care and had their first CD4 count measured, and had more than one HIV test as part of the surveillance program. Participants included in the HIV surveillance program are age 15 and above. Due to uncertainty in the precise timing of HIV seroconversion dates, we calculated the midpoint between the first HIV positive test and the last HIV negative test. Those whose seroconversion date was after the earliest date of
initiation of HIV care for the first HIV-infected household member to link to care were included. We included the HIV-uninfected household members who met these criteria of individuals whose earliest CD4 count date fell between January 1, 2007 and August 1, 2011 and fell between 0 and 350 cells/µl. An upper bound of 350 cells/µl was chosen because during the study period there were other ART eligibility thresholds at 350 cells/µl for pregnant women and tuberculosis patients. We were unable to exclude pregnant and tuberculosis patients from the dataset, because they could not be identified at the time of the earliest CD4 count, and were only identifiable for those who initiated ART. Including patients above 350 cells/µl in this circumstance would bias estimates at the 200 cell/µl threshold used in the regression discontinuity analysis. We did not place any additional restrictions on the age or gender of participants included in the analysis.

**Regression discontinuity design.** The regression discontinuity design is a quasi-experimental study design which can be implemented when an exposure of interest is at least in part determined by a variable measured continuously used to determine treatment or exposure status.[20,24,27,28] Regression discontinuity designs take advantage of clinical or other threshold rules which determine treatment or other exposure status. For example, the regression discontinuity design has been used efficacy of prostate-specific antigen (PSA) screening for detection of prostate cancer.[29] Men with PSA over 4.0 ng/mL are eligible for further prostate cancer workup. The authors used PSA>4.0 ng/mL as a cutoff to determine whether or not a participant received additional prostate cancer workup. The authors found no decrease in prostate cancer-specific or all-cause mortality as a result of increased prostate cancer workup. Other recent examples of applications of
regression discontinuity include the effect of human papillomavirus vaccination on cervical dysplasia and anal warts[30] and sexual behaviors[31], and on immediate versus delayed ART initiation on mortality[20] and retention in care.[32]

We exploit the fact that immediate ART initiation upon engaging in HIV-related care is determined by CD4 count. Prior to August 2011, patients in South Africa were initiated on ART if their CD4 count fell below 200 cells/μl. The standardized monitoring schedule was CD4 count measurement every six months to determine ART eligibility. If individuals presented over the 200 cells/μl threshold, they would not be assessed for eligibility again for six months, which could result in a delay in initiating ART for those who are close to the threshold. CD4 counts are measured with some degree of error. For individuals who engaged in care with CD4 counts of approximately 200 cells/μl, whether or not they presented just above or just below the threshold is approximately random due to the presence of measurement error.[20,24]

Regression discontinuity designs are particularly useful in the setting of unmeasured confounding. Whereas most regression-based confounding adjustment methods require the strong assumption of no unmeasured confounding, regression discontinuity designs require the continuity of potential outcomes assumption to be met. Particularly when the assignment variable is measured with random error, such as CD4 count, whether or not an individual is above or below the threshold will be random within a narrow bandwidth around the threshold. Therefore, the distribution of measured and unmeasured confounders is expected to be similar on either side of the threshold for individuals presenting near the threshold, similar to when exposure is randomized in a randomized controlled trial.[24] In analyses of the effect of ART initiation on HIV
incidence in household members, there may be multiple sources of unmeasured confounding, such as the household’s tendency to seek healthcare, engagement in HIV prevention activities, or self-protective behaviors such as condom use or repeat HIV testing. We therefore chose the regression discontinuity design for the present study.

The assumption of randomness across the threshold may not hold with increasing distance from the threshold. This has two primary implications for the analysis and interpretation of results. First, regression discontinuity allows for the estimation of local effects, which are effects in the CD4 count range close to the 200 cells/μl threshold. This has important generalizability implications, as the effect must be interpreted as the effect for patients who have an earliest CD4 count of approximately 200 cells/μl. To estimate these local effects, regression models are estimated with separate slopes on either side of the threshold and an intercept change at the threshold. The effect estimate is the comparison of regression predictions just above versus just below the threshold (intercept shift).

Second, because regression discontinuity estimates local effects, analyses typically are presented for a range of bandwidths of the continuous variable which is the basis for the eligibility threshold (assignment variable; in this case, CD4 count). Narrow bands represent the least biased effect estimates, because the assumption that individuals immediately above and below the threshold are similar with respect to their baseline characteristics is most likely to hold. In practice, however, there may be a limited number of individuals who are very close to the threshold. Wider bands around the threshold will improve power by including more individuals in the analysis, but will also increase potential for bias if the true relationship is non-linear, as the local linear model
will be a poorer fit to the data and lead to boundary bias at the threshold. Modeling of the assignment variable on either side of the threshold can allow for inclusion of individuals far from the threshold without substantially increasing the risk of bias if the relationship between the log-hazards and the covariates is approximately linear. Presentation of results at multiple thresholds, including narrow thresholds that have reduced power but less risk of bias, and wider thresholds with more power and higher risk of bias can give additional information on the true effect size. Data-driven optimal bandwidth selectors have been derived for regression discontinuity designs using linear regression. In lieu of an optimal bandwidth, best practice is to show sensitivity to a range of bandwidth choices.

Commonly, not all patients will follow treatment assignment as determined by the assignment rule (similar to non-adherence in a randomized controlled trial). Indeed, patients may have been started on therapy with a high CD4 count due to Stage IV HIV illness during this period of study. In this case, the intention-to-treat (ITT) effect in a regression discontinuity design is estimated in a regression model with a term for whether the individual was above or below the threshold and terms for the slope of the assignment variable above and below the threshold. The ITT is equivalent to the effect of presenting just below the threshold, analogous to a randomized controlled trial in which the ITT is the effect of randomized treatment arm, regardless of whether or not the patients actually adhered to their randomized treatment.

Statistical analysis. All analyses accounted for multiple individuals within a homestead by clustering standard errors at the homestead level. We estimated the ITT by fitting a
Cox proportional hazards model to the value of first CD4 count, allowing the hazard to shift at the threshold, and allowing the slope above and below the threshold to differ. Analyses were conducted in a range of bandwidths around the assignment variable as well as the optimal bandwidth as determined by the Imbens-Kalyanaraman algorithm using a linear probability model.[33] This algorithm estimates the optimal bandwidth as a balance of bias and variance. We assessed robustness to modeling the relationship with the assignment variable as a quadratic and as a restricted cubic spline with knots at 100 cells/ul on either side of the threshold, allowing for non-linear relationships between CD4 count and the log-hazards. The restricted cubic splines relax the linearity assumption and provide information on whether our assumption of linearity in the primary models led to bias. We ran an additional sensitivity analysis including baseline covariates. If, as expected, baseline covariates are balanced above and below the threshold, there should be no change in point estimates with the inclusion of additional baseline covariates.[34] Variables included in sensitivity analyses included the age, educational status, and sex of the respondent and the household member linking to care as well as an index of the household’s wealth. Analyses for the ITT were run including multiple specifications of the hazard function, including both exponential and Weibull distributions. We used instrumental variable methods[23] to estimate the effect of immediate initiation of ART on household HIV incidence among individuals who took treatment because they were below the threshold (see Appendix for details). Analyses were run in Stata 14.1 (StataCorp, College Station, TX).
RESULTS

A total of 4,115 individuals were HIV-uninfected at the time the first HIV-infected individual in their household linked to care and had a presentation CD4 count between 0 and 350 cells/μl. Of these, there were 2,490 HIV-uninfected household members among HIV-infected individuals who linked to care below 200 cells/μl (and thus eligible for ART) and 1,626 above 200 cells/μl (and thus ineligible). Baseline characteristics were well balanced between those who started above and below 200 cells/μl (Table 1). Balance tests indicated no difference in baseline characteristics at the threshold.

The Imbens-Kalyanaraman algorithm determined that the optimal bandwidth was 95 cells/μl above and below the threshold. At CD4 count bandwidths of 0-350, 50-350, 105-295 (the optimal bandwidth), 150-250, and 175-225 cells/μl, a total of 4,115, 3,531, 2,356, 1,268, and 615 HIV-uninfected individuals were included for regression discontinuity analyses, respectively. The probability of ART initiation within 6 months of the HIV-infected household member by first CD4 count at the clinic is displayed in Figure 1. The probability of initiation of ART within 6 months was highest among individuals who presented below 200 cells/μl, and there was a strong discontinuity at the threshold. A histogram of baseline CD4 counts (Figure 2) demonstrated no bunching at the threshold, indicating no evidence of manipulation of CD4 counts.

There were 337 HIV seroconversions among 13,785 person-years at risk, with an overall incidence rate of 2.4 seroconversions per 100 person-years (95% CI 2.2 to 2.7). The discontinuity in HIV incidence at the threshold by baseline CD4 count of the first HIV-infected household member to link to care is shown in Figure 3. In the optimal
bandwidth (105-295 cells/μl), immediate initiation of ART reduced HIV incidence by 47% (HR 0.53, 95% CI 0.30 to 0.96), Table 2. Models at wider bandwidths that included more flexible modeling of the functional form of CD4 count were consistent with a 50% reduction in HIV incidence, similar to effect estimates at the narrower bandwidths. Sensitivity analyses modeling CD4 count with restricted cubic splines and squared terms allow for flexible modeling of the relationship between CD4 count and HIV incidence. These models may reduce bias in effect estimates at the widest bandwidth, where individuals are included further from the threshold, by improving model fit. In the widest bandwidth (0 to 350 cells/μl), which includes the most information but is the most sensitive to violations of the assumption of linearity, the hazard ratio with a linear functional form of CD4 count was 0.68 (95% CI 0.46 to 1.02), which decreased to 0.48 (95% CI 0.26 to 0.88) with a restricted cubic spline at 100 cells/μl above and below the threshold and 0.45 (95% CI 0.24 to 0.85) with a squared term for CD4 count. At the narrowest bandwidth (175-225), which has the least power but is the least vulnerable to misspecification, the hazard ratio was 0.40 (95% CI 0.14 to 1.13). These results were robust to alternative specifications of the hazard function and inclusion of baseline covariates in the model (Supplemental Table S1 and S2, http://links.lww.com/QAD/B212). HIV incidence among household members who initiated treatment because they were below the threshold had a 93% reduction in HIV incidence compared to those who did not initiate treatment because they were above the threshold (HR 0.07, 95% CI 0.01 to 0.52; Appendix).
DISCUSSION

We demonstrate a substantial reduction in HIV incidence among HIV-uninfected household members when the first member of the household to seek HIV-related care was eligible for ART immediately compared to delayed ART eligibility. The ITT effect, which represents the effect of ART eligibility for all patients seeking care with CD4 counts close to the threshold, indicated an approximate reduction of 47%, with alternate specifications suggesting reductions in the range of 32-60% decrease in HIV acquisition among household members. In models at the widest threshold, including the entire range of CD4 counts eligible for the study, the effect size increase from a 32% decrease in HIV incidence to 52% in models allowing for a more flexible functional form of CD4 count, indicating that the true effect size is likely closer to 50%. Given that all individuals in the household were included in this analysis, regardless of whether or not two individuals were in a partnership, there may be benefits to entire households with immediate initiation of ART that extend beyond the well-documented benefits in couples.[1,35,36]

As expected, the effect estimates in this study fell between those from the HPTN052 randomized controlled trial among serodiscordant couples[9] and community-level effects of ART on HIV incidence.[14] Similar to prior community-level estimates, the effect estimated in this study includes both biological reduction in transmission risk among sex partners to the person on ART as well as potential changes in behavior that may affect HIV acquisition among HIV-uninfected household members, which may include increased awareness and risk protection. The household spillover effects are expected to be larger than the community spillovers, and smaller than effects in
serodiscordant couples because a smaller share of the population was involved in
serodiscordant sexual relationships with the index patient.

To date, the vast majority of studies assessing HIV infections within households
or families have focused on HIV transmission within couples or mother to child
transmission. In both scenarios, ART has been shown to be highly efficacious in the
prevention of HIV infection.[35,37,38] Evidence from the United States demonstrated
substantial clustering of HIV within households of HIV-infected or high-risk women[17],
with household infection more common among siblings than among intimate partners or
children. Residents of the same physical spaces likely share common characteristics,
including socioeconomic, education, behavioral, and community characteristics that may
influence HIV risk. Within households, individuals who immediately initiate ART may
more frequently disclose their status to household members, which could lead to
behavioral changes among household members. HIV prevention messages from
counseling in the clinic may be more likely to reach household members of those who
immediately uptake ART, which could result in decreases in household HIV acquisition.

This analysis has several important limitations. Because not everyone in the
surveillance system participated in HIV testing every year, the decision to participate in
testing may be affected by the exposure (immediate ART initiation by the HIV-infected
household member). However, data arose from two separate systems (the public sector
ART clinic data and the HIV surveillance system) and these results are unlikely due to
differential non-participation. Without additional untestable assumptions[20,39], the
results of this analysis are only be generalizable to individuals who present close to the
ART initiation threshold. To generalize the results beyond the threshold, additional
assumptions related to the functional form of how the unobserved potential outcome changes with the assignment variable over the distribution of CD4 counts is required, an assumption that is untestable and strong.[39] This is therefore an issue of external, rather than internal, validity as the effect may not be generalizable to a global effect. The use of the regression discontinuity design represents a significant strength, as it does not rely on the assumption of no unmeasured confounding. Furthermore, a strength of this analysis is the ability to link clinic-based data to household data, including HIV surveillance in household members, allowing for estimation of effects within households without relying on self-report from individuals linked to HIV care.

We found a substantial reduction in HIV incidence in households where with immediate eligibility for ART. The results of this study provide further evidence of the importance of immediate initiation of ART to reduce HIV transmission. Furthermore, these results demonstrate that multiple pathways likely exist, including both behavioral and biological pathways, to reduction of HIV incidence upon ART initiation within households. Spillover effects of ART initiation beyond the biological effect of ART uptake on HIV transmission likely play a role in reduction of HIV incidence at the community level that has been previously noted with increasing ART coverage, and ART initiation likely has benefits to members of the social network extending beyond sexual partners.

ACKNOWLEDGEMENTS

Author Contributions

CEO designed the study, conducted analyses, and drafted the report. JB designed the study, conducted analyses, and drafted the report. GH interpreted results and drafted the report. FT
supervised data collection, designed the study, interpreted results, and critically reviewed the manuscript. TM supervised data collection, designed the study, interpreted results, and critically reviewed the manuscript. MS supervised data collection, designed the study, interpreted results, and critically reviewed the manuscript. GRS designed the study, interpreted results, and critically reviewed the manuscript. VDG designed the study, interpreted results, and critically reviewed the manuscript. MJM designed the study, interpreted results, and critically reviewed the manuscript. KHM designed the study, interpreted results, and critically reviewed the manuscript. DP supervised data collection, designed the study, interpreted results, and critically reviewed the manuscript. TB supervised data collection, designed the study, conducted analyses, and drafted the report.

**Funding**

The Hlabisa Human Immunodeficiency Virus Treatment and Care Programme was funded by the United States Agency for International Development and the President's Emergency Plan (grant 674-A-00-08-001-00). This work was partially supported by the National Institutes of Health (grants T32-DA013911 and R25-MH083620 to C. E. O., grant R37-AI51164 to V. D. G., grant R01-HD084233 to T. B., and grants R01-AI124389 and R01-HD084233 to T. B. and F. T). The Africa Health Research Institute, University of KwaZulu-Natal, South Africa receives core funding from the Wellcome Trust (grant 082384/Z/07/Z).
Figure 1. Probability of ART initiation by baseline CD4 count

The probability of antiretroviral therapy (ART) initiation with 6 months of initiating HIV care was calculated as the number of individuals who initiated ART within 6 months of the date of their earliest CD4 count divided by the total number of individuals who had a first CD4 count date. Probabilities were calculated by 10-cells/µl bins of earliest CD4 counts (the first CD4 count the individual had after initiating HIV care).
Figure 2. Regression discontinuity validity test – histogram displaying distribution of baseline CD4 counts. The distribution of frequency of baseline CD4 counts of the first member of the household to link to HIV-related care. This histogram demonstrates no evidence of bunching at the threshold (200 cells/µl), indicating no evidence of manipulation of baseline CD4 counts. Evidence of bunching could be seen if there were substantially more patients in the bin immediately above or below the threshold compared to the other side. Manipulation of baseline CD4 counts, which could occur for example if clinicians reported a CD4 count below 200 when the true CD4 count was above 200 so the patient could access ART, would result in bias in estimates of the effect of ART eligibility on outcomes.
**Figure 3.** HIV incidence by baseline CD4 count HIV incidence in HIV-uninfected household members by the baseline CD4 count of the first HIV-infected member of their household to link to HIV-related care. Orange dots indicate the raw HIV incidence for each 10 cell/μL bin. Green lines indicate fitted regression lines estimating the incidence of HIV as a function of earliest CD4 count above and below the threshold (red line). The dotted green line is the projection for the curve below the threshold, which is the estimate of what HIV incidence would be for individuals above the threshold (and thus not eligible for immediate ART initiation) would have been if they had actually been eligible for ART immediately. The discontinuity at the threshold is the estimate of the effect of ART eligibility on incidence.
Table 1. Baseline characteristics of study sample by household member ART eligibility (N=4,115)

<table>
<thead>
<tr>
<th></th>
<th>Below Threshold</th>
<th>Above Threshold</th>
<th>Balance Test P-value¹</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>(N=2,489)</td>
<td>(N=1,626)</td>
<td></td>
</tr>
<tr>
<td>Age when first household member linked to care, years, median (IQR)</td>
<td>20 (16 to 48)</td>
<td>20 (16 to 47)</td>
<td>0.15</td>
</tr>
<tr>
<td>Female sex</td>
<td>1,529 (61.4%)</td>
<td>1,015 (62.4%)</td>
<td>0.70</td>
</tr>
<tr>
<td>Number of HIV-uninfected individuals in household, median (IQR)</td>
<td>3 (2 to 4)</td>
<td>3 (2 to 4)</td>
<td>0.58</td>
</tr>
<tr>
<td>Highest education attainment</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Less than 7 years</td>
<td>1,026 (41.2%)</td>
<td>682 (41.9%)</td>
<td>0.38</td>
</tr>
<tr>
<td>7 to 12 years</td>
<td>1,394 (56.0%)</td>
<td>891 (54.8%)</td>
<td></td>
</tr>
<tr>
<td>More than 12 years</td>
<td>63 (2.5%)</td>
<td>51 (3.1%)</td>
<td></td>
</tr>
<tr>
<td>Knows HIV status</td>
<td>481 (19.3%)</td>
<td>295 (18.1%)</td>
<td>0.77</td>
</tr>
<tr>
<td>Household location</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Urban</td>
<td>42 (1.7%)</td>
<td>42 (2.6%)</td>
<td>0.78</td>
</tr>
<tr>
<td>Peri-Urban</td>
<td>904 (36.3%)</td>
<td>536 (33.0%)</td>
<td></td>
</tr>
<tr>
<td>Rural</td>
<td>1,544 (62.0%)</td>
<td>1,048 (64.5%)</td>
<td></td>
</tr>
<tr>
<td>Household distance to clinic, km (median, IQR)</td>
<td>2.7 (1.5 to 3.9)</td>
<td>2.7 (1.5 to 4.1)</td>
<td>0.28</td>
</tr>
<tr>
<td>Household wealth (quintile)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Lowest quintile</td>
<td>420 (16.7%)</td>
<td>260 (16.0%)</td>
<td></td>
</tr>
<tr>
<td>Second lowest</td>
<td>481 (19.3%)</td>
<td>333 (20.5%)</td>
<td></td>
</tr>
<tr>
<td>Middle</td>
<td>585 (23.5%)</td>
<td>375 (23.1%)</td>
<td></td>
</tr>
<tr>
<td>Second highest</td>
<td>496 (19.9%)</td>
<td>356 (21.9%)</td>
<td>0.31</td>
</tr>
<tr>
<td>Highest</td>
<td>419 (16.8%)</td>
<td>233 (14.3%)</td>
<td></td>
</tr>
<tr>
<td>Missing</td>
<td>89 (3.6%)</td>
<td>69 (4.2%)</td>
<td></td>
</tr>
</tbody>
</table>

¹Regression discontinuity model using each baseline covariate as the outcome model to test if there is a discontinuity at the CD4 count threshold in each baseline covariate. A linear regression model was used for continuous variables, logistic regression for dichotomous variables, and ordinal logistic regression for ordinal variables.
<table>
<thead>
<tr>
<th>Range</th>
<th>N</th>
<th>Cox</th>
<th>Cox, Quadratic</th>
<th>Cox, Restricted Cubic Splines</th>
</tr>
</thead>
<tbody>
<tr>
<td>0-350</td>
<td>4,115</td>
<td>0.68 (0.46 to 1.02)</td>
<td>0.45 (0.24 to 0.85)</td>
<td>0.48 (0.26 to 0.88)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>0.999 (0.996 to 1.003)</td>
<td>0.993 (0.979 to 1.007)</td>
<td>0.996 (0.987 to 1.006)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>0.997 (0.993 to 1.002)</td>
<td>0.997 (0.980 to 1.014)</td>
<td>1.006 (0.994 to 1.018)</td>
</tr>
<tr>
<td>50-350</td>
<td>3,531</td>
<td>0.55 (0.35 to 0.86)</td>
<td>0.55 (0.28 to 1.08)</td>
<td>0.57 (0.31 to 1.05)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>0.999 (0.996 to 1.003)</td>
<td>0.993 (0.979 to 1.007)</td>
<td>0.996 (0.987 to 1.006)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>0.994 (0.989 to 0.999)</td>
<td>1.007 (0.986 to 1.028)</td>
<td>0.990 (0.971 to 1.010)</td>
</tr>
<tr>
<td>105-295**</td>
<td>2,356</td>
<td>0.53 (0.30 to 0.96)</td>
<td>0.46 (0.21 to 1.00)</td>
<td>0.48 (0.24 to 0.98)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>0.998 (0.991 to 1.006)</td>
<td>0.973 (0.949 to 0.997)</td>
<td>0.983 (0.968 to 0.999)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>0.995 (0.984 to 1.005)</td>
<td>1.035 (0.996 to 1.076)</td>
<td>0.941 (0.854 to 1.037)</td>
</tr>
<tr>
<td>150-250</td>
<td>1,268</td>
<td>0.47 (0.23 to 0.98)</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>0.986 (0.969 to 1.002)</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>1.014 (0.988 to 1.040)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>175-225</td>
<td>615</td>
<td>0.40 (0.14 to 1.13)</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>0.980 (0.938 to 1.023)</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>1.001 (0.938 to 1.080)</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

*Optimal bandwidth per Imbens-Kalyanaraman algorithm; ¹Difference in slope of CD4 count above the 200 cell/μl threshold; ²Difference in slope of CD4 count below the 200 cell/μl threshold.
REFERENCES


