Medicine®

OBSERVATIONAL STUDY

OPEN

# Effects of Referral Bias on Estimates of Anal Intraepithelial Neoplasia Progression and Regression Rates in a 3-State Markov Model

*William Christopher Mathews, MD, MSPH, Edward Rafael Cachay, MD, MAS, Wollelaw Agmas, MD, MAS, and Christopher Jackson*

**Abstract:** The study aim is to compare anal intraepithelial neoplasia (AIN) progression and regression rates in a cytology inception cohort to estimates based on the subcohort referred for ≥1 high-resolution anoscopies (HRAs).

A cytology-based retrospective cohort was assembled including the anal cytology histories and invasive anal cancer (IAC) outcomes of all HIV-infected adults under care between 2001 and 2012. A 3-state Markov model (<HSIL↔HSIL→IAC) was estimated separately for all patients and for the subcohort undergoing ≥ 1 HRAs with biopsy. Cytology was adjusted for misclassification. State transition rates (per person-year) and covariate hazard ratios were estimated using the R package *msm*.

Of 2804 eligible patients in the inception cohort, 629 (22%) were in the HRA subcohort and 2175 (78%) in the non-HRA subcohort. Patients in the HRA subcohort were more likely to have baseline CD4<350, viral load >400, and to have HSIL at baseline and thereafter. They also had more anal cytology examinations (median 6 vs 3) and longer follow-up (median 5.5 vs 3.6 years). State transition rates were overestimated in the HRA subcohort relative to inception cohort, but the degree of discordance varied by transition: for <HSIL to HSIL (0.44 vs 0.04); for HSIL to <HSIL (0.56 vs 0.17); and for HSIL to IAC (0.014 vs 0.011). Beneficial covariate effects on the <HSIL to HSIL transition were concordant ($P < 0.05$) for time-updated HIV viral load, CD4 count, and antiretroviral therapy. The observed effects of HRA-triage bias may be relevant to estimates of AIN state transitions from other cohorts subject to referral bias.

(*Medicine* 94(35):e1476)

**Abbreviations**: AIN = anal intraepithelial neoplasia, ASC-H = atypical squamous cells, can't rule out high grade, ASCUS = atypical squamous cells of uncertain significance, HIV = human immunodeficiency virus, HPV = human papillomavirus, HRA = high-resolution anoscopy, HSIL = high-grade squamous intraepithelial lesion, IAC = invasive anal cancer, IANS = International Anal Neoplasia Society, IQR = interquartile range, IRC = infrared coagulation, LSIL = low-grade squamous intraepithelial lesion, MSM = men who have sex with men, *msm* = multistate model package in R statistical software, $R_{HRA}$ = transition rate in the HRA referral cohort, $R_{IC}$ = transition rate in the inception cohort.

## INTRODUCTION

Among the 10 classic criteria of Wilson and Jungner to justify implementing a screening program is an adequate understanding of the natural history of the index disease.[1] However, a number of biases may distort accurate estimation of key transition rates along the natural history pathway.[2] Referral bias, a form of selection bias, may occur when natural history is modeled using a referred study cohort with a prognostic factor distribution that differs from that of the source population at risk of disease. Modeling of natural history and the impact of prognostic factors is therefore ideally based on the experience of inception cohorts so as to minimize referral bias. Screening programs for anal intraepithelial neoplasia (AIN) typically limit referral for high-resolution anoscopy (HRA) to those with abnormal anal cytology results (high-grade squamous intraepithelial lesion [HSIL], low-grade SIL [LSIL], atypical cells of uncertain significance [ASCUS] or atypical cells, cannot rule out high grade [ASC-H]) or other clinical abnormalities.[3] The aim of this study was to compare estimates of AIN progression and regression rates in a cytology inception cohort to estimates based on the subset of patients referred for ≥ 1 HRA with biopsy procedures. We propose the term *HRA-triage bias* to designate the form of referral bias demonstrated in the following analysis.

## METHODS

A clinical care cytology-based retrospective cohort was assembled including the anal cytology histories and invasive anal cancer (IAC) outcomes of all HIV-infected adults under care at the UCSD Owen Clinic between 2001 and 2012. Eligibility criteria, screening program characteristics, and study measure definitions were as previously reported.[4] A 3-state Markov model (<HSIL↔HSIL→IAC) was estimated separately for all patients (inception cohort) and for the subset of patients undergoing at least one HRA with biopsy (HRA subcohort). Bidirectional transitions were allowed between the <HSIL and HSIL states, and IAC was considered an absorbing state that could not regress in the absence of treatment. Cytology

Editor: Akhilanand Chaurasia.
Received: June 19, 2015; revised: July 20, 2015; accepted: July 24, 2015.
From the Department of Medicine, University of California San Diego, San Diego, California, United States of America (WCM, ERC, WA); and Medical Research Council Biostatistics Unit, Cambridge Institute of Public Health, Cambridge, United Kingdom (CJ).
Correspondence: William Christopher Mathews, UCSD Medical Center, 8681, 200W, Arbor Dr., San Diego, CA 92103 (e-mail: cmathews@ucsd.edu).
This work was presented in part at the International Anal Neoplasia Society (IANS) Scientific Meeting, 13 March 2015, Atlanta, Georgia, USA.
Sources of Support: This work was supported by the following: Clinical Investigation and Biostatistics Core of the UC San Diego Center for AIDS Research (AI036214) and CFAR Network of Integrated Clinical Systems-CNICS, an NIH funded program (R24 AI067039). The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

ISSN: 0025-7974
DOI: 10.1097/MD.0000000000001476

states were adjusted for misclassification using within-cohort estimates of sensitivity (0.66) and specificity (0.90), with concurrent HRA-directed biopsy as the reference standard. Entry time for both study groups was defined as the date of the first screening anal cytology obtained at clinic entry (for those who entered care after inauguration of the screening program in 2001) or as the date of the first anal cytology (for those already under care when the screening program was started). All evaluable cytology results (baseline and subsequent) were included irrespective of their temporal association with an HRA procedure. State transition rates (per person-year) and prognostic covariate hazard ratios were estimated using the R package *msm*.[5] We present absolute and relative differences in state transition rates between the inception cohort ($R_{IC}$) and HRA referral cohort ($R_{HRA}$) where the relative difference is defined as $(R_{HRA} - R_{IC})/R_{IC}$. The study was approved by the UCSD Human Research Protection Program (Project 071931).

## RESULTS

Of 2804 eligible patients in the inception cohort, 629 (22%) were in the HRA subcohort and 2175 (78%) in the non-HRA subcohort. Patient characteristics are presented in Table 1. Patients in the HRA subcohort were: more likely to be male (91.6% vs 88.2%, $P = 0.02$), white (64.7% vs 60.8%, $P = 0.01$), men having sex with men (MSM) (82.7% vs 76.3%, $P = 0.002$), to have baseline CD4<350 cells/mm$^3$ (50.6% vs 42.4%, $P < 0.0001$) and HIV plasma viral load >400 copies/mL (58.7% vs 49.6%, $P < 0.0001$), and more likely to have HSIL at baseline (32.0% vs 5.3%, $P < 0.0001$) and thereafter (98.9% vs 22.6%, $P < 0.0001$). They also had more anal cytology examinations (median 6 vs 3, $P < 0.0001$) and longer follow-up (median 5.5 vs 3.6 years, $P < 0.0001$) than the non-HRA subcohort. The groups did not differ by age at entry (median 40.2, interquartile range 34.1 – 46.4), use of antiretroviral therapy at entry (75%), or smoking (29.9%). Of the 23 confirmed incident IAC cases, all occurred in the HRA subcohort. One or more infrared coagulation (IRC) treatments of HSIL lesions were documented in 26% of HRA subcohort patients.

Table 2 presents estimates of *state transition rates* (per person-year) for the inception cohort and for the HRA subcohort. State transition rates were overestimated in the HRA subcohort relative to the inception cohort, but the degree of discordance varied by transition: for <HSIL to HSIL (0.44 vs 0.04); for HSIL to <HSIL (0.56 vs 0.17); and for HSIL to IAC (0.014 vs 0.011). When expressed as *2-year transition probabilities*, the comparable estimates of discordance (HRA subcohort vs inception cohort) are: for <HSIL to HSIL (0.38 vs 0.07); for HSIL to <HSIL (0.47 vs 0.28); and for HSIL to IAC (0.019 vs 0.019).

Covariate effects on state transition rates, estimated separately for the inception cohort and for the HRA subcohort are presented in Table 3. Beneficial covariate effects on the <HSIL to HSIL transition were concordant ($P < 0.05$) for time-updated HIV viral load, CD4 count, and antiretroviral therapy. But the favorable effect of IRC on the HSIL to <HSIL transition was significant only in the inception cohort. Smoking was not a significant covariate in either model.

## DISCUSSION

An inception cohort has been defined as ''a group of individuals identified and assembled for subsequent study at an early and uniform point in the course of the specified health condition.''[6] Failure to assemble an inception cohort can have unpredictable and often important effects on estimates in natural history studies.[7] In this analysis, we found that Markov model state transition rates were overestimated if cohort membership was conditioned on receipt of one or more HRA procedures as compared to unconditional estimates from the cytology inception cohort. Relative to inception cohort estimates, the effect of HRA referral bias was greatest for the <HSIL to HSIL transition (10-fold), moderate for the HSIL to <HSIL transition (2.3-fold), and minimal for the HSIL to IAC transition (0.3-fold).

What factors may account for the differential effects of conditioning analytic cohort membership on receipt of HRA? We discuss first the observed differential inflation of state transition rates and then effects on covariate effect estimation. Because the development of HSIL cytology was the primary criterion for referral to HRA, it is unsurprising that the transition rate from <HSIL to HSIL would be overestimated when the model was conditioned on receipt of ≥1 HRA procedures. Likewise, because IRC treatment of HSIL was found to augment the regression rate from HSIL to <HSIL in the inception cohort, and because 26% of the HRA subcohort underwent at least one IRC procedure, it is perhaps not unexpected that the rate of regression of HSIL would be augmented when estimated in the HRA subcohort. With regard to the minimal impact of HRA subcohort selection on estimation of the HSIL to IAC transition, at least 2 factors may be involved: (1) all cases of IAC were observed in the HRA subcohort (likely because HSIL is the precursor state to IAC and HSIL patients were preferentially triaged to HRA); and (2) IAC is a rare outcome; it is possible that with longer cohort followup, estimates of the HSIL to IAC transition rate in the HRA subcohort would diverge from that estimated in the inception cohort.

With regard to the impact of HRA subcohort selection on covariate effect estimation, it is noteworthy that there was concordance between both analytic cohorts in the significant protective effects of antiretroviral therapy, HIV viral load suppression, and CD4 + lymphocyte count on the <HSIL to HSIL transition and in the null effect of smoking on all modeled transitions. With regard to the discordant estimates of the effect of IRC on downgrading HSIL to <HSIL (a substantial effect in the inception cohort but no effect in the HRA subcohort), we hypothesize that the failure to detect an effect in the HRA subcohort is related to the coincidence or mixture of at least 2 effects on HSIL regression: putative spontaneous regression augmented by the superimposed effect of IRC on those who were so treated. However, it seems counter-intuitive that, although the unadjusted relative estimate of the HSIL to <HSIL SIL transition rate was 2.3-fold higher in the HRA subcohort (compared to the estimate in the inception cohort), the specific effect of IRC could not be distinguished from the spontaneous regression effect in the HRA subcohort.

Our explanatory speculations highlight the previously noted unpredictability of the effects of failure to assemble an inception cohort owing to limited ability to account for discordant distributions of both measured and unmeasured prognostic factors in the context of referral bias.[7] Although we have identified no other publications examining the effects of referral bias on estimation of AIN clinical evolution, we do call attention to work examining the effect of referral bias on estimation of progression to cirrhosis in hepatitis C infected analytic cohorts. Fu et al found in a simulation study that the estimated 20-year probability of progression to cirrhosis in patients referred to a liver clinic was 4-fold higher (20%) than the estimate for community-based samples (5%). The authors

**TABLE 1.** Patient Characteristics by Study Group

| Characteristic | HRA Subcohort (A) (n = 629) | No HRA Subcohort (B) (n = 2175) | *P* Value (A vs. B)$_1$ | All Patient Inception Cohort (n = 2804) |
|---|---|---|---|---|
| Age at entry (years Median [IQR]) | 40.2 [34.6,46.1] | 40.3 [33.9,46.5] | 0.588 | 40.2 [34.1,46.4] |
| Sex | | | 0.017 | |
| Male | 91.6% | 88.2% | | 89.0% |
| Female | 8.4% | 11.8% | | 11.0% |
| Race/ethnicity | | | 0.012 | |
| White | 64.7% | 60.8% | | 61.7% |
| Hispanic | 18.8% | 17.1% | | 17.5% |
| Black | 9.0% | 13.8% | | 12.7% |
| Other/unknown | 7.5% | 8.3% | | 8.1% |
| HIV risk factor | | | 0.002 | |
| MSM (not IDU) | 82.7% | 76.3% | | 77.8% |
| IDU (not MSM) | 3.7% | 5.0% | | 4.7% |
| Heterosexual (not IDU) | 9.0% | 14.3% | | 13.1% |
| Other/unknown | 4.6% | 4.4% | | 4.4% |
| Baseline CD4 | | | <0.0001 | |
| $<350/mm^3$ | 50.6% | 42.4% | | 44.2% |
| $\geq 350/mm^3$ | 49.4% | 57.6% | | 55.8% |
| Baseline HIV viral load | | | <0.0001 | |
| $\leq 400$ copies/$mm^3$ | 41.3% | 51.4% | | 49.2% |
| $>400$ copies/$mm^3$ | 58.7% | 49.6% | | 50.8% |
| ART at entry$_2$ | | | 0.327 | |
| Yes | 73.3% | 75.2% | | 74.8% |
| No | 26.7% | 24.8% | | 25.2% |
| Smoking | | | 0.785 | |
| Yes | 29.4% | 30.0% | | 29.9% |
| No | 70.6% | 70.0% | | 70.1% |
| Baseline cytology$_3$ | | | <0.0001 | |
| NAMC | 14.8% | 33.9% | | 29.6% |
| ASCUS | 24.6% | 35.1% | | 32.8% |
| LSIL | 22.9% | 23.9% | | 23.6% |
| ASC-H | 5.7% | 1.8% | | 2.7% |
| HSIL | 32.0% | 5.3% | | 11.3% |
| Worst cytology group | | | <0.0001 | |
| HSIL | 98.9% | 22.6% | | 39.7% |
| <HSIL | 1.1% | 77.4% | | 60.3% |
| IAC$_4$ case (n = 23) | | | <0.0001 | |
| Yes | 3.7% | 0% | | 0.8% |
| No | 96.3% | 100% | | 99.2% |
| IRC ($\geq 1$)$_5$ | | | <0.0001 | |
| Yes | 25.9% | 2.5% | | 7.8% |
| No | 74.1% | 97.5% | | 92.2% |
| No. cytology results (Median [IQR]) | 6 [4,8] | 3 [2,5] | <0.0001 | 5 [3,8] |
| Follow-up (years, Median [IQR]) | 5.5 [3.0,8.2] | 3.6 [1.8,6.6] | <0.0001 | 4.0 [2.0,7.1] |
| Year first cytology (Median [IQR]) | 2004 [2002,2007] | 2004 [2002,2007] | 0.347 | 2004 [2002,2007] |

1. Statistical comparison between HRA subcohort and non-HRA subcohort by the chi-square test for categorical variables and the Wilcoxon rank sum test for numerical variables.

2. ART = antiretroviral therapy, ASC-H = atypical squamous cells, cannot rule out high grade, ASCUS = atypical squamous cells of uncertain significance, HRA = high-resolution anoscopy, HSIL = high-grade squamous intraepithelial lesion, IAC = invasive anal cancer, IQR = interquartile range, IRC = infrared photocoagulation. A small number of patients underwent IRC without documented HRA-directed biopsy and are therefore included in the non-HRA subcohort, LSIL = low-grade squamous intraepithelial lesion, MSM = men who have sex with men, NAMC = no atypical or malignant cells.

concluded that ''When attempting to establish the natural history of new diseases with long incubation periods, researchers should be on the lookout for potential biases that result from the way patients are referred into clinical cohorts.''[8]

Although not the primary focus of our analysis, it is notable that all 23 IAC cases were documented in the HRA referral subcohort. We believe that there are several potential explanations for this observation. First, because HSIL was the

**TABLE 2.** Estimates of State Transition Rates (per Person-Year) Adjusted for Cytology Misclassification Assumptions, by Study Group (All Patient Inception Cohort vs HRA Subcohort)

| State Transition$_1$ | (A) All Patients (n = 2804) (rate [95% CI]) | (B) HRA Subcohort (n = 629) (Rate [95% CI]) | Relative Difference$_2$ |
|---|---|---|---|
| <HSIL to HSIL | 0.04 [0.036, 0.054] | 0.44 [0.33, 0.59] | 10 |
| HSIL to <HSIL | 0.17 [0.13, 0.22] | 0.56 [0.43, 0.72] | 2.29 |
| HSIL to IAC | 0.011 [0.007, 0.017] | 0.014 [0.009, 0.021] | 0.27 |

CI = confidence interval, HRA = high-resolution anoscopy, HSIL = high-grade squamous intraepithelial lesion,IAC = invasive anal cancer. Calculated as (B−A)/A.

primary criterion for triage to HRA, the HRA subcohort was greatly enriched with patients at highest risk of progression to IAC. Second, we did not implement IRC treatment of HSIL lesions until 2007 (year 7 of a 12-year study period). Third, although we found in our primary analysis that IRC increased the downgrading of HSIL to <HSIL, we were unable to show that IRC reduced the rate of progression from HSIL to IAC.[4] Finally, even among patients undergoing HRA surveillance during the IRC treatment era (2007 and thereafter), assuring regularity of followup examinations was a challenge because of limited HRA operator availability and patient adherence to followup recommendations.

Our study results are subject to several limitations. (1) Unknown or imprecisely measured factors may have contributed to prognostic differences between the inception and HRA subcohorts. (2) For both analytic cohorts, there may be residual state misclassification after correcting for the fallibility of cytology using HRA-directed biopsy as the reference standard;

**TABLE 3.** Estimated Unadjusted Hazard Ratios (95% CI) of Time-Updated Covariates, by State-Transition and by Study Group

| Covariate | State Transition | Study Group | Hazard Ratio | 95% CI |
|---|---|---|---|---|
| IRC [reference: no IRC] | < HSIL to HSIL | All patients | 2.2 | 0.6, 7.9 |
| | | HRA subcohort | 0.3 | 0.1, 1.4 |
| | HSIL to <HSIL | All patients | **4.2** | **2.0, 8.5** |
| | | HRA subcohort | 1.4 | 0.7, 3.0 |
| | HSIL to IAC | All patients | 2.7 | 0.6, 11.7 |
| | | HRA subcohort | 2.1 | 0.5, 9.3 |
| ART [reference: no ART] | < HSIL to HSIL | All patients | **0.4** | **0.2, 0.6** |
| | | HRA subcohort | **0.2** | **0.04, 0.5** |
| | HSIL to <HSIL | All patients | 0.9 | 0.04, 2.1 |
| | | HRA subcohort | 0.5 | 0.2, 1.8 |
| | HSIL to IAC | All patients | 2.2 | 0.5, 9.4 |
| | | HRA subcohort | 2.5 | 0.6, 10.6 |
| HIV Viral Load [reference: >400 copies/mm$^3$] | < HSIL to HSIL | All patients | **0.3** | **0.2, 0.5** |
| | | HRA subcohort | **0.3** | **0.2, 0.7** |
| | HSIL to <HSIL | All patients | 1.3 | 0.7, 2.3 |
| | | HRA subcohort | 0.9 | 0.4, 1.9 |
| | HSIL to IAC | All patients | 1.6 | 0.7, 3.9 |
| | | HRA subcohort | 1.7 | 0.7, 4.2 |
| CD4 Category [reference: < 350/mm$^3$] | < HSIL to HSIL | All patients | **0.3** | **0.2, 0.5** |
| | | HRA subcohort | **0.4** | **0.2, 0.8** |
| | HSIL to <HSIL | All patients | 0.8 | 0.5, 1.3 |
| | | HRA subcohort | 0.9 | 0.5, 1.7 |
| | HSIL to IAC | All patients | 1.4 | 0.6, 3.3 |
| | | HRA subcohort | 1.5 | 0.6, 3.4 |
| Smoking [reference: not smoking] | < HSIL to HSIL | All patients | 1.1 | 0.7, 1.7 |
| | | HRA subcohort | 1.2 | 0.6, 2.2 |
| | HSIL to <HSIL | All patients | 0.8 | 0.5, 1.4 |
| | | HRA subcohort | 1.0 | 0.6, 1.9 |
| | HSIL to IAC | All patients | 1.2 | 0.5, 2.9 |
| | | HRA subcohort | 1.4 | 0.6, 3.2 |

ART = antiretroviral therapy, CI = confidence interval; HIV = human immunodeficiency virus, HRA = high-resolution anoscopy, HSIL = high-grade squamous intraepithelial lesion, IAC = invasive anal cancer, IRC = infrared photocoagulation. A small number of patients underwent IRC without documented HRA-directed biopsy and are therefore included in the non-HRA subcohort.

however, HRA-directed biopsy is itself a fallible reference standard.[9] (3) The <HSIL state is heterogeneous, including some human papillomavirus (HPV) uninfected patients; nonetheless, similar HIV-infected cohorts have demonstrated high incidence and persistence of oncogenic HPV.[10] (4) Because of the composition of our analytic cohorts, model inferences are most robust for HIV-infected MSM. (5) Power to detect covariate effects on the transition from HSIL to IAC is limited by the small number of IAC endpoints.

## CONCLUSIONS

Modeling AIN state transition rates in a noninception cohort defined by differential HRA referral resulted in substantial overestimation of the <HSIL to HSIL and HSIL to <HSIL transitions. There was less bias in estimation of the HSIL to IAC transition. The effects of HRA-triage bias are both challenging to predict and are very likely relevant to estimates of AIN state transitions from other cohorts subject to any form of referral bias.

## REFERENCES

1. Andermann A, Blancquaert I, Beauchamp S, et al. Revisiting Wilson and Jungner in the genomic age: a review of screening criteria over the past 40 years. *Bull World Health Organ.* 2008;86:317–319.
2. Morrison AS. Screening in Chronic Disease. New York: Oxford University Press; 1992:2nd ed..
3. Park IU, Palefsky JM. Evaluation and management of anal intraepithelial neoplasia in HIV-negative and HIV-positive men who have sex with men. *Curr Infect Dis Rep.* 2010;12:126–133.
4. Mathews WC, Agmas W, Cachay ER, et al. Natural history of anal dysplasia in an HIV-infected clinical care cohort: estimates using multi-state Markov modeling. *PloS One.* 2014;9:e104116.
5. Jackson C. Multi-state modelling with R: the msm package. 2014. http://127.0.0.1:11362/library/msm/doc/msm-manual.pdf. Accessed May 31, 2015.
6. Porta M. A Dictionary of Epidemiology. New York: Oxford University Press; 2014:6th ed..
7. Ales KL, Charlson ME. In search of the true inception cohort. *J Chronic Dis.* 1987;40:881–885.
8. Fu B, Tom BD, Delahooke T, et al. Event-biased referral can distort estimation of hepatitis C virus progression rate to cirrhosis, and of prognostic influences. *J Clin Epidemiol.* 2007;60:1140–1148.
9. Mathews WC, Cachay ER, Caperna J, et al. Estimating the accuracy of anal cytology in the presence of an imperfect reference standard. *PloS One.* 2010;5:e12284.
10. de Pokomandy A, Rouleau D, Ghattas G, et al. Prevalence, clearance, and incidence of anal human papillomavirus infection in HIV-infected men: the HIPVIRG cohort study. *J Infect Dis.* 2009;199:965–973.