Distribution Agreement

In presenting this Thesis as a partial fulfillment of the requirements for an advanced degree from Emory University, I hereby grant to Emory University and its agents the non-exclusive license to archive, make accessible, and display my Thesis in whole or in part in all forms of media, now or hereafter known, including display on the world wide web. I understand that I may select some access restrictions as part of the online submission of this Thesis. I retain all ownership rights to the copyright of the Thesis. I also retain the right to use in future works (such as articles or books) all or part of this Thesis.

____

Signature of Student Date

HIV TESTING AMONG AT RISK
ADOLESCENT POPULATIONS, HOW CAN WE DO BETTER?
YRBS, 2005-2011

BY

Karen Coeytaux

Master of Public Health
Applied Epidemiology

____

Patrick S. Sullivan, PhD, DVM   December 2013
Committee chair

____

Michael R. Kramer, PhD   December 2013
Field advisor

____

Kevin M. Sullivan, PhD, MPH   December 2013
Track director

____

Melissa Alperin, MPH, MCHES   December 2013
Program chair

HIV TESTING AMONG AT RISK
ADOLESCENT POPULATIONS, HOW CAN WE DO BETTER?
YRBS, 2005-2011

BY

Karen Coeytaux

MPH, Emory University, 2013
Ph D, University Paris XI, 2004

Thesis Committee Chair: Patrick S. Sullivan, PhD, DVM

An abstract of
A Thesis submitted to the Faculty of the
Rollins School of Public Health of Emory University
in partial fulfillment of the requirements of the degree of
Master of Public Health in Epidemiology in the Career MPH program 2013

Abstract

HIV TESTING AMONG AT RISK
ADOLESCENT POPULATIONS, HOW CAN WE DO BETTER?
YRBS, 2005-2011

BY
Karen Coeytaux

Background: HIV testing is crucial among adolescents aged 13-19, as they often report multiple HIV risk behaviors. Currently, most are unaware of their positive HIV status (59.5% of HIV-positive people aged 13-24) and adolescents aged 13-19 account for 5% of new infections in the US.

Purpose: The purpose of this study was to (i) assess the association between HIV testing and related risk behaviors, controlling for sex, grade, race and ethnicity, (ii) examine changes overtime in HIV testing and related risk behaviors at the national and state levels, and (iii) compare 2011 state and national levels of HIV testing and related risk behaviors.

Methods: This study used the biennial state and national Youth Risk Behavioral Survey (YRBS), representative of students in 9th through 12th in the US for the period 2005-2011, and focused on sexually active students. Risk behaviors included in the study were condom use, number of sexual partners in lifetime, age at first intercourse, ever forced sexual intercourse, and ever illegal injection drug use. Assessment was performed through logistic regression and trend analyses.

Results: HIV testing was associated with all analyzed risk behaviors among sexually active high school students, controlling for sex, grade, race and ethnicity. The national level of HIV testing among sexually active high school students was stable around 22% over the 2005-2011 period. In contrast to the national level, evolution of HIV testing and related risk behaviors was observed in some states. Additionally, meaningful differences in HIV testing and related risk behaviors were found across states in 2011.

Conclusions: As of 2011, most sexually active high school students had never been tested for HIV, despite engaging in HIV-related risk behaviors. Although the situation has remained mostly unchanged between 2005 and 2011 at the national level, states display different evolutions and prevention needs. This calls for a more extensive use of YRBS to monitor progress in HIV prevention and testing among adolescents, and ultimately enhance early linkage of HIV positive adolescents to care.

HIV TESTING AMONG AT RISK ADOLESCENT POPULATIONS, HOW CAN WE DO BETTER? YRBS, 2005-2011

BY

Karen Coeytaux
MPH, Emory University, 2013
PhD, Paris XI University, 2004

Thesis Committee Chair: Patrick S. Sullivan, MPH, DVM

A Thesis submitted to the Faculty of the
Rollins School of Public Health of Emory University
in partial fulfillment of the requirements of the degree of
Master of Public Health in Epidemiology in the Career MPH program 2013

ACKNOWLEDGEMENTS

I would like to thank Lisa Whittle from the Division of Adolescent and School Health National Center for HIV/AIDS, Viral Hepatitis, STD, and TB Prevention Centers for Disease Control and Prevention, for helping me collecting the YRBS data from the different states. I would also like to thank the state YRBS coordinators Jean Ajamie, William Owen Blair, Jane Crump, Jennifer Haussler Garing, Pam Hitt, Ava L. Holt, Kim Kovalchick, Chiniqua N. Milligan, Angela Norton, Julia Peek, and Viani Ramirez-Irizarry for granting me access to their state YRBS datasets.

I would also like to thank Michael R. Kramer and Patrick S. Sullivan, from the Epidemiology Department of Emory University for their valuable advice and assistance throughout this project.

Finally, I would like to thank my husband and children for their support.

TABLE OF CONTENTS

INTRODUCTION

Human Immunodeficiency Virus (HIV) remains an important public health issue in the United States and youth aged 13-24 represent a significant share of newly diagnosed cases. It is also estimated to be the age group with the highest share of people unaware about their HIV infection (1, 2). It is important for people to know their HIV status, so that infected people can receive treatment and avoid transmitting the disease (3). Therefore, many prevention efforts focus on promoting HIV testing.

In April 2013, the United States Preventive Services Task Force released a grade A recommendation to screen for HIV in the general population 15-65 years of age, which is aligned with the CDC recommendation from 2006 (4, 5). This implies that all adolescents should be routinely tested for HIV infection, regardless of their risk level. To make sure that these guidelines are properly enforced and to evaluate their impact, it is necessary to monitor the level of HIV testing among adolescents, with a particular attention to subpopulations which are at higher risk. This information, with a view into potential geographic variation and trends over time, can be used in better targeting HIV prevention interventions and screening efforts (e.g., CDC Expanded Testing Initiative).

Data about HIV testing among youth can be retrieved from the Youth Risk Behavioral Survey (YRBS). It is a biennial national survey for students in 9th through 12th grades in public and private schools, conducted by the CDC since 1991. The questionnaire consists of items about health-related risk behaviors. Questions and their response options may differ across and within years and sites. The survey is self-administered and based on anonymous and voluntary participation. National YRBS data files are publicly available online for the 1991- 2011 period. State YRBS data files can be requested and are provided in compliance with YRBS participation, data quality, and data-sharing policies. Some previous studies have shown that rates of HIV testing were rather low among youth, especially among male, including those who are sexually active and reported high-risk behaviors for HIV (6, 7).

This epidemiological study built on this previous work and went into further exploration in order to bring additional elements informing HIV testing progresses. It relied on national YRBS data and state YRBS data, for states capturing information about HIV testing (7states in total: Massachusetts, Connecticut, New Jersey, Arkansas, North Carolina, North Dakota, and South Carolina) over the period 2005-2011. Five HIV-related risk behaviors were selected for this study: ever had sexual intercourse, age at first sexual intercourse, use of condom at last sexual intercourse, ever force to have intercourse, lifetime use of injected illegal drug. This study focused on high school students who had sexual intercourse and addressed the following questions:

1. Did HIV testing and related risk behaviors changed between 2005 and 2011, controlling for sex, grade, race and ethnicity?
2. Was HIV testing associated with related risk behaviors, controlling for sex, grade, race and ethnicity, between 2005 and 2011?
3. How did HIV testing and related risk behaviors compare across states and with national level?

LITERATURE REVIEW

Adolescent and young adult population is at increased risk of HIV infection

HIV infection and AIDS remain a major public health issue in the US. By year end 2010, 888,917 people were living with diagnosed HIV in the US, of which 75% were males and 25% females. Compared to the overall population, adolescents and young adults are disproportionately affected by HIV; hence prevention programs need to bring particular attention to this sub-population. During 2011, 21% of the 50,007 people newly diagnosed for HIV infection were adolescents or young adults, whereas this age group represented 17% of the overall population at the end of 2011 (1). This share of new diagnoses was further broken down into 5% for adolescents (aged 13-19 yrs) and 16% for young adults (aged 20-24 yrs). In addition, the number of yearly new diagnoses in the adolescent and young adult population has grown at an average annual rate of 5% from 2008 to 2011, while it has been decreasing at an average annual rate of -1% in the overall population. Finally, the CDC estimated that in 2009 18.1% of the people living with HIV infection were undiagnosed in the overall US population, while 59.5% of the 13-24 yr old Americans living with HIV were unaware of their condition (2).

Besides bringing a particular focus on adolescents and young adults, prevention measures also need to take into account the specific socio-demographic factors and risk behaviors that are associated with new HIV infections in this age group. In 2011, the states with the highest rates of new diagnoses among the 20-24yr old were mostly in the South. In fact, while the average national rate of new diagnoses in this age group was 36.3 per 100,000 people, Washington DC and 10 Southern states had rates above 40 (Maryland, Georgia, Louisiana, Mississippi, Alabama, South Carolina, Florida, Texas, Tennessee, North Carolina), while only 2 Northern states were above that threshold (New York and Illinois). In addition, Black/African Americans have been disproportionately affected by HIV, representing 60% of the new cases among adolescents and young adults in 2011, although this figure was only 47% in the overall US population and only 15% of all adolescents and young adults were Black/African Americans (1). When looking at risk behaviors, the main cause of new HIV infection was by far male-to-male sex contact among male adolescents and young adults (91% of new diagnoses in 2011 vs. 78% in the overall male population) and heterosexual contact among female adolescents and young adults (92% of new diagnoses in 2011 vs. 86% in the overall female population). Injection drug use was a somewhat less important cause of infection among adolescents and young adults compared to the overall population (2% vs. 6% for males, and 9% vs. 14% for females) (1).

HIV testing is critical for the control of HIV infection

As mentioned above, it was estimated that more than half of the adolescents and young adults living with HIV in the US don't know about their infection. This figure exemplifies how critical it would be to increase the rate of HIV testing in this population to ensure proper care of infected people and reduce transmission. It has been shown that people who initiate antiretroviral therapy (ART) at CD4 counts between 0.200 and 0.500 × 10^9 cells/L have a reduced risk for AIDS-related events or death (8). As patients are likely to be asymptomatic at that stage of infection, the only way to detect them is through screening. There is also convincing evidence that identification and treatment of HIV-positive pregnant women dramatically reduces rates of mother-to-child transmission (9). In addition, testing and detection of HIV infection enables reduction of transmission by potentially avoiding exposure between patients and their partners, but also because there are substantial reductions in transmission when patients receive ART (3). As Branson and colleagues indicate, control and elimination of HIV will be possible only with widespread testing, prompt and accurate diagnosis, and universal access to immediate antiviral therapy.

Although the benefits of HIV testing are well understood, it is difficult to devise the best approach regarding potential populations to target in priority and the frequency at which testing should be conducted. Studies have shown that screening for HIV on the basis of risk factor assessment alone may miss 20% to 25% of HIV-positive individuals who report no risk factors (5, 10). Also, there is insufficient evidence to determine the optimum time intervals for HIV screening. Still, patient populations that would more likely benefit from more frequent testing include those who are known to be at higher risk for HIV infection, those who are actively engaged in risky behaviors, and those who live in a high-prevalence setting. In this context, the USPSTF has recently proposed a pragmatic approach in which groups at very high risk for new HIV infection would be rescreened at least annually and individuals at increased risk at somewhat longer intervals (for example, 3 to 5 years). It also indicated that routine rescreening may not be necessary for individuals who have not been at increased risk since they were found to be HIV-negative. The validity of this approach is further supported by some cost effectiveness analyses. One-time HIV screening of low-risk persons coupled with annual screening of high-risk persons could prevent 6.7% of projected new infections and would cost $22,382 per QALY gained, assuming a 20% reduction in sexual activity after screening (11). Furthermore, routine, one-time rapid HIV testing for all adult patients is cost-effective even when the prevalence of undiagnosed HIV infection is as low as 0.2%, assuming that earlier diagnosis and antiretroviral treatment reduces HIV transmission (12).

Given the clear benefits of testing on the control of HIV, multiple recommendations have been made over the recent years to reinforce screening at large in the adolescent and adult population. In 2006, the CDC recommended routine voluntary HIV screening in all individuals aged 13 to 64 years regardless of recognized risk factors, unless the prevalence of undiagnosed HIV infection had been documented to be less than 0.1% in the sub-population considered (4). The recommendation was to use opt-out HIV testing in order to reduce barriers to testing. This way, unless they specifically decline, all patients should be informed about and undergo testing without a requirement for prevention counseling before screening. This approach was endorsed in 2009 by the American College of Physicians (13). The Infectious Diseases Society of America formulated a similar recommendation, which promotes routine HIV screening for all sexually active adults (14). The American Congress of Obstetricians and Gynecologists has recommended routine screening in all women aged 19 to 64 years, using an opt-out approach. In addition, for women outside of that age range, it has proposed targeted screening based on risk factors (15). The American Academy of Pediatrics has recommended offering routine HIV testing to all adolescents at least once between 16 and 18 years in communities where HIV prevalence is greater than 0.1%. In sub-populations where prevalence is lower than 0.1%, it recommended testing of all sexually active adolescents and those with risk factors (16). The American Academy of Family Physicians has recommended that clinicians screen adolescents and adults aged 18 to 65 years for HIV infection, as well as younger adolescents and older adults who are at increased risk, along with all pregnant women (17). As indicated above, the most recent recommendation came in 2013 from the US Preventive Services Task Force, which concluded with high certainty that the net benefit of screening for HIV infection in adolescents, adults, and pregnant women was substantial. Hence, it recommended that clinicians screen adolescents and adults aged 15 to 65 years for HIV infection, as well as younger adolescents and older adults who are at increased risk. It also recommended that clinicians screen all pregnant women (5).

Despite success of current HIV testing initiative, rates remain low for youth

To translate recommendations into actions, the CDC launched the Expanded HIV Testing Initiative (ETI) in October 2007. It funded 25 health departments to facilitate HIV screening, focusing on U.S. jurisdictions that had reported 140 or more AIDS diagnoses among blacks in 2005. The ultimate goal was to increase diagnoses of HIV infections and linkage to care among populations disproportionately affected by HIV, especially non-

Hispanic blacks. Through this program, substantial numbers of persons previously unaware of their HIV infection were identified and linked to care over the 2007-2010 period (18). In addition, it has been shown that this initiative had a return on investment superior to $1, providing further support for such large-scale programs (19).

Based on the early success of the ETI, it is important to monitor the levels of HIV testing in the overall population in order to assess progress and be able to adjust the focus of the program to maximize its impact. As seen above, adolescents and young adults represent a significant share of newly diagnosed infections, but also have a very high rate of undiagnosed cases; hence they are a critical population for screening programs and the progress of testing need to be monitored over time in this age group. In 2011, states that had collected data on the percentage of students who have been tested for HIV infection in the Youth Risk Behavior Survey (YRBS), had a rather low percentage (13%) of 9th–12th grade students that had ever been tested for HIV (7). HIV testing rates varied according to state, sexual experience, sex, race/ethnicity, grade, and the use of condom during last intercourse. It was particularly low among males. A previous analysis of the 2009 YRBS dataset had shown that most sexually active students had not been tested for HIV, even those who reported high-risk behaviors for HIV (6). Based on these previous studies, there is interest and value in further understanding how socio-demographic factors and risk behaviors can influence low HIV testing in youth over time. Better understanding of these associations will help monitoring whether the most at risk sub-groups are properly tested and thus further inform where additional effort might be needed to strengthen education and prevention programs.

.

MATERIALS AND METHODS

Data source

The CDC's Youth Risk Behavioral Survey (YRBS) is a national survey that measures the prevalence of health risk behaviors among high school students. It is an anonymous paper-and-pencil survey administered every other year to several thousand students in randomly selected public and private high schools (Appendix A). National and states YRBS data are available for the period 1991-2011 with an overall response rate over 60%. State data are released in compliance with YRBS participation, data quality, weighted data availability, and data-sharing policies. National and state YRBS samples are designed to be representative of students in grades 9-12 in the United States and in the considered state.

The study covered the years 2005, 2007, 2009, and 2011. At the national level, survey data were available for each of these years. Upon completion of IRB process, data were requested for the national level and the 50 states. Data was retrieved for 43 of the states and the remaining 7 had either no weighted data available or did not conduct YRBS in the 2005-2011 period (CA, MN, OR, WA), or did not make the data accessible (HI, IN, VT).

At the state level, only 7 states included a question about HIV testing in their YRBS for either all or part of the years in the study period: Massachusetts (MA) included all years; Connecticut (CT) included 2007 to 2011; New Jersey (NJ) included 2005, 2009, and 2011; Arkansas (AR), North Carolina (NC), and North Dakota (ND) included 2009 and 2011, and

South Carolina (SC) included only 2009. In total, 59,793 observations were available at the national level, including students who had sex or not, 12,089 observations for MA, 7,980 observations for NC, 6,522 observations for CT, 4,908 observations for NJ, 3,749 observations for ND, 3,065 observations for AR, and 1,108 observations for SC.

Variables of interest

Demographic characteristics included sex, grade, and race and ethnicity of the subject. These were respectively monitored through the following questions: (i) "What is your sex?"; (ii) "In what grade are you?"; (iii) "Are you Hispanic or Latino?"; (iv) "What is your race?".

The outcome of interest was HIV testing and was assessed through the question "Have you ever been tested for HIV, the virus that causes AIDS?".

Risk behaviors selected for this study included the following questions related to sexual practices and use of injected drug: (i) "Have you ever had sex?"; (ii) "Have you ever been physically forced to have sexual intercourse when you did not want to?"; (iii) "How old were you when you had sexual intercourse for the first time?"; (iv) "During your life, with how many people have you had sexual intercourse?"; (v) "The last time you had sexual intercourse, did you or your partner use a condom?"; (vi) "During your life, how many times have you used a needle to inject any illegal drug into your body?".

Some variables were derived from existing variables for harmonization purposes across states, years, and with previous work (6). This was done as follow:

For NJ 2005 and MA 2005: (i) the combined variable 'race and ethnicity' was derived from the variables 'race' and 'ethnicity', (ii) the information about HIV testing was derived from the broader question on sexually transmitted diseases: "Have you ever been tested for HIV infection or other sexually transmitted diseases (STDs) such as genital herpes, chlamydia, syphilis, or genital warts?" and the answer option "Yes, I have been tested for HIV".

For ND 2009, the use of condom was derived from the broader question about contraception: "The last time you had sexual intercourse, what one method did you or your partner use to prevent pregnancy?".

For all surveys, 'Hispanic/Latino' and 'Multiple - Hispanic' were assigned to a single category 'Hispanic', and 'American Indian or Native Alaskan', 'Asian', 'Native Hawaiian and other Pacific Islander' were assigned to a single category 'other'. Last, the subpopulation 'ungraded' and 'other' under the question "In what grade are you?" was excluded from the analysis, due to small numbers.

The variables 'race and ethnicity' and 'grade' were polytomous. Risk behaviors and sex were dichotomous variables.

The level of missing data was checked and found to be inferior to 10%, except for HIV testing in the national survey in 2011 (24%) and over the entire 2005-2011 period (13%), as well as in NC over the 2009-2011 period (12%). Also, more than 10% of missing data were found for lifetime illegal injection drug use in NC over the 2009-2011 period (29%). All other variables at either national or state level remained under the 10% threshold of missing data.

All analyses accounted for the complex sampling design of YRBS. State analyses took into account the finite population correction (fpc) due to their smaller size. All data used and results presented in this study were weighted. Logistic regression analyses and collinearity tests were performed with SAS 9.3. Collinearity was assessed using a collinearity macro (Appendix B). Trend analyses and t-tests were performed with SAS-callable SUDAAN. Domain analysis was used to perform analyses on the subpopulation of high school students who had sexual intercourse. A significant level alpha=0.05 was used throughout the study.

Descriptive analysis

Descriptive analyses were drawn at national and state levels for the period 2005-2011. Description included (i) the demographics sex, grade, race and ethnicity, (ii) the outcome of interest "ever been tested for HIV", and (iii) the selected HIV-related risk behaviors: "condom use at last sexual intercourse", "four or more sexual partners in life", "first sexual intercourse before age 13", "ever forced to have intercourse", "lifetime illegal injection drug use".

Trend analysis

Trends were drawn to assess evolutions over time in HIV testing practice and related risk behaviors among high school students who had sexual intercourse. National trend analysis covered the years 2005-2011. State trend analyses covered consecutive years available for each state: 2005-2011 for MA, 2007-2011 for CT, 2009-2011 for NJ, AR, NC, and ND. No trend could be assessed for SC, as only 2009 data was available.

Linear changes were assessed through logistic regression analysis when three or more consecutive surveys were available. Quadratic changes could not be assessed as the study covered less than 6 consecutive surveys. T-test was used for changes when data were available for only two consecutive surveys. Trend and t-tests controlled for underlying changes in sex, race/ethnicity, and grade during the same time period.

Logistic regression analysis

The association of HIV screening with related risk behaviors at the national level was modeled using logistic regression. First, logistic regression was used to assess the crude association between HIV testing and each risk behavior and demographic. Then, effect modification by sex, race/ethnicity, grade level, and year of survey was assessed through backward elimination. Multivariable logistic regression model was then used to assess the adjusted association between HIV testing and each independent variable. Multicollinearity between covariates included in the model was tested in the initial and final model. CNI>30 and VDPs>0.5 were used as criteria for multicollinearity diagnosis. Assessed risk behaviors and demographics were kept in the model, independently of their significance in order to align with literature.

RESULTS

Descriptive analysis at national level

The national YRBS conducted from 2005 to 2011 had 28,177 observations corresponding to high school students, who had sexual intercourse, out of a total of 59,793 observations. The distribution of high school students who had sexual intercourse was as follows: 47.9% were female, 52.1% were male, 19.8% were in 9th grade, 23.7 % in 10th grade, 26.8% in 11th grade, and 29.7% in 12th grade. A majority of students were White (55.6%), while 18.7% were Black or African American, 18,7% were Hispanic, and 6.9% were of other race and ethnicity. 21.9% had ever been tested for HIV and among students who had sex, 35.4% declared that they did not use a condom during their last sexual intercourse. 31.1% had four or more sexual partners during their lifetime, 13.5% had their first intercourse before 13 years old, 14.1% had ever been forced to have sexual intercourse, and 3.5% had injected illegal drug in their body at least once (Table 1). Demographics of the population of interest were consistent across the years studied (Fig. 1).

National trends in HIV testing and related risk behaviors, 2005-2011

National trend analysis did not show any change in either HIV testing or related risk behaviors between 2005 and 2011 among high school students who had sexual intercourse. No significant linear change was detected (Fig. 2 and Table 2).

Association of HIV testing and related risk behaviors at national level

Crude analysis among high school students who had sexual intercourse, showed that the odds of HIV testing was higher for females compared to males (2005-2011: cOR = 1.72, 95% CI: 1.59, 1.86). This result was systematically observed across years. Students in 12th grade had greater odds of HIV testing compared to students in 9th grade for the years 2007 (cOR = 1.59., 95% CI: 1.26, 2.00), 2009 (cOR = 1.59, 95% CI: 1.25, 2.02, and the overall period 2005-2011 (cOR = 1.39, 95% CI: 1.22, 1.58). There was no difference found among grades 9, 10, and 11.

For each year studied, Black or African American high school students had greater odds of HIV testing, compared to White high school students (2005-2011: cOR = 1.67, 95% CI: 1.45, 1.92). Students of other race and ethnicity had smaller odds of HIV testing in 2009 (cOR = 0.72, 95% CI: 0.56, 0.94) and higher odds in 2011 (cOR = 1.43, 95% CI: 1.08, 1.90) compared to White high school students, but no significant difference was noticed for 2005 (cOR = 1.30, 95% CI: 0.90, 1.90) and 2007 (cOR = 1.34, 95% CI: 0.98, 1.84). No other difference by race and ethnicity was found. In all years, the odds of HIV testing were greater for high school students engaging in one of the HIV-related risk behaviors considered in this study (Table 3).

After backward elimination, several significant interaction effects were detected: between ever forced sexual intercourse and sex (p-value=0.02), ever forced sexual intercourse and grade (p-value = 0.01), age at first sexual intercourse and sex (p-value = 0.001), use of condom and sex (p-value = 0.001), and having four or more sexual partners in lifetime and race and ethnicity (p-value < 0.001). Meaningful differences were further assessed through the comparison of stratum specific ORs and CIs. Their comparison lead to the conclusion that there was no strongly meaningful difference in effect estimates (Table 4). Hence, no interaction term was kept in the adjusted analysis. Multicollinearity tests showed no sign of collinearity between any of the co-variates included in the model. The final model included the five selected risk factors and adjusted for sex, grade, and race and ethnicity.

The adjusted analysis showed that the odds of HIV testing were higher in females compared to males in the overall period 2005-2011 and for each year analyzed as well (2005-2011: aOR = 1.84, 95% CI: 1.68, 2.01, 2005: aOR = 1.77, 95% CI: 1.46, 2.14, 2007: aOR = 1.78, 95% CI: 1.52, 2.08, 2009: aOR = 1.94, 95% CI: 1.64, 2.29, 2011: aOR = 1.85, 95% CI: 1.55, 2.21). Students in 11th and 12th grade had greater odds of HIV testing compared to students in 9th grade in the overall period 2005-2011 (2005-2011: aOR = 1.22, 95% CI: 1.06, 1.40 for 11th grade and aOR = 1.41, 95% CI: 1.23, 1.62 for 12th grade). This was reflected in 2007 and 2009, but not in 2005 and 2011 though. No difference in HIV testing was found between 10th and 9th graders at any time, as all 95% CIs included the null value. Among high school students who had sexual intercourse, students of Black or African American ethnicity had greater odds of HIV testing compared to students of White ethnicity. This result was consistent throughout the years studied (2005-2011: aOR = 1.60, 95% CI: 1.37, 1.87, 2005: aOR = 1.59, 95% CI: 1.26, 2.00, 2007: aOR = 1.59, 95% CI: 1.31, 1.93, 2009: aOR = 1.41, 95% CI: 1.04, 1.92, 2011:

aOR = 1.90, 95% CI: 1.24, 2.92). Similarly to the crude analysis, results for students of other ethnicity were not consistent across the years.

The odds of HIV testing among students of Hispanic ethnicity were not significantly different from the odds of HIV testing among students of White ethnicity. 95% CIs included the null value.

The odds of HIV testing were higher for students who did not use a condom at their last intercourse compared to the odds for students who used a condom at their last intercourse, while adjusting for sex, grade, race and ethnicity. This was reflected in the overall analysis period and in each year analyzed (2005-2011: aOR =1.30, 95% CI: 1.20, 1.41, 2005: aOR = 1.23, 95% CI: 1.02, 1.49, 2007: aOR = 1.36, 95% CI: 1.17, 1.59, 2009: aOR = 1.28, 95% CI: 1.09, 1.51, 2011: aOR = 1.29, 95% CI: 1.13, 1.48). The odds of HIV testing among students who had four sexual partners or more in their lifetime was more than twice the odds among students who had less than four sexual partners in their lifetime, while adjusting for sex, grade, race and ethnicity. This result was consistent across years analyzed (2005-2011: aOR =2.27, 95% CI: 2.08, 2.47, 2005: aOR = 2.39, 95% CI: 1.95, 2.92, 2007: aOR = 2.11, 95% CI: 1.80, 2.46, 2009: aOR = 2.35, 95% CI: 2.03, 2.72, 2011: aOR = 2.33, 95% CI: 1.94, 2.81). The odds of HIV testing among students who had ever forced intercourse were higher than the odds among students who never had forced intercourse, while adjusting for sex, grade, race and ethnicity. This result was also consistent across years analyzed (2005-2011: aOR =1.50, 95% CI: 1.35, 1.67, 2005: aOR = 1.46, 95% CI: 1.14, 1.86, 2007: aOR = 1.44, 95% CI: 1.13, 1.84, 2009: aOR = 1.45, 95% CI: 1.22, 1.72, 2011: aOR = 1.73, 95% CI: 1.44, 2.08).

The association between HIV testing and having first sexual intercourse before 13, controlling for sex, grade, race and ethnicity, was not significant. CIs included the null value or were borderline (2005-2011: aOR =1.22, 95% CI: 1.09, 1.38, 2005: aOR = 1.19, 95% CI: 0.96, 1.48, 2007: aOR = 1.34, 95% CI: 1.03, 1.74, 2009: aOR = 1.23, 95% CI: 0.98, 1.55, 2011: aOR = 1.13, 95% CI: 0.93, 1.39). The odds of HIV testing among students who ever used injected illegal drug were higher than the odds among students who never used injected illegal drug, controlling for sex, grade, race and ethnicity, in the overall period 2005-2011 and for 2007 and 2009(2005-2011: aOR =1.68 , 95% CI: 1.34, 2.10, 2007: aOR = 1.87, 95% CI: 1.15, 3.04, 2009: aOR = 1.77, 95% CI: 1.20, 2.61). But results were inconclusive for 2005 and 2011 (2005: aOR = 1.54, 95% CI: 0.95, 2.49, 2011: aOR = 1.54, 95% CI: 0.94, 2.54), as 95% CIs included the null value (Table 5).

Descriptive analysis at state level

Among all states, only 7 had available YRBS data including information about HIV testing for at least one of the years 2005, 2007, 2009, and 2011. These states were MA, CT, NJ, AR, NC, ND, and SC (Table 6). The number of observations available and the proportion of students who had sexual intercourse across states were as follow: 4,884 students out of 12,089, had sexual intercourse in MA YRBS, 2,342 out of 6,522 in CT YRBS, 1,972 students out of 4,908 in NJ YRBS, 3,434 out of 7,980 in NC YRBS, 1,445 out of 3,749 in ND YRBS, 1,282 out of 3,065 in AR YRBS, and 529 out of 1,108 in SC YRBS.

The characteristics of high school students who had sexual intercourse showed some similarities and differences across states and in comparison to the national sample (Table 7 and Fig. 3).

Student populations were similar in gender. There were an estimated 48.4% to 51.7% of males depending on states, in comparison to 52.1% of males for the overall period 2005 and 2011 at national level among sexually active students.

State and national samples had also similar populations in terms of grade level. Still, CT, NJ, and ND's populations were slightly shifted toward higher grades with less students in 9th grade and a greater proportion of students in 12th grade, compared to other states and national sample (CT, NJ, ND: from 14.4% to15.2% in 9th grade and from 34.8% to 36.9% in 12th grade, Other states: from 17.8% to 24.3% in 9th grade and from 27.9% to 31.9% in 12th grade, 2005-2011 national sample: 19.8% in 9th grade and 29.7% in 12th grade).

The distribution of race and ethnicity showed important differences between states and compared to the national distribution. The proportion of White varied from 49.0% in SC to 80% in ND compared to 55.6% for the 2005-2011 national sample, the proportion of Black or African American varied from 0.8% in ND to 44.4% in SC compared to 18.7% for the 2005-2011 national sample, the proportion of Hispanic varied from 2.1% in ND to 21.6% in NJ compared to 18.7% for the 2005-2011 national sample, and the proportion of students of other race and ethnicity varied from 2.5% in AR to 17.0% in ND compared to 6.9% for the 2005-2011 national sample.

The proportion of HIV testing among students who had sex varied from 14.2% for NJ to 26.4% for AR, while the National level was at 21.9%.

The percentage of students who had not used a condom at last intercourse showed some variations from 30.1 % for NJ to 38.5% for ND. All states considered had a higher proportion of students who had sexual intercourse and reporting not using a condom at last intercourse compared to the 2005-2011 national level (35.4%), except NJ (30.1%) and CT (35.0%).

The percentage of students who had sexual intercourse with 4 or more partners during their lifetime was the lowest in CT with 26.6% and the highest in SC with 36.6%, while the 2005-2011 national level was 31.1% .

The proportion of students who had sexual intercourse and whose first sexual intercourse occurred before age 13, was the lowest in ND (7.8%) and the highest in AR (17.7%), as compared to 13.5 % for the 2005-2011 national sample.

The proportion of students, who had been forced to have intercourse, ranged from 11.8% in ND to 16.6% in MA, while the proportion was 14.1% in the 2005-2011 national sample.

The proportion of students who had ever used injected drug was low and close across the different populations surveyed. It varied from 2.1% for NC and SC to 4.2% for NJ, and 3.5% for the 2005-2011 national sample.

State trends in HIV testing and related risk behaviors, 2005-2011

State trends analysis showed several significant trends in HIV testing and related risk behaviors between 2005 and 2011 among high school students who had sexual intercourse, while no significant trends were detected at national level as described previously (Table 8).

In MA, significant linear changes were detected for HIV testing (linear trend: β= 0.84762, p-value≤0.001). This indicated an overall significant increase in HIV testing between 2005 and 2007 among high school students who had sexual intercourse. Over the same period, there was a significant increase of students reporting that they did not use condom at their last intercourse (linear trend: β= 0.20252, p-value=0.004). No other significant change was detected for other HIV related risk factors over this period in MA.

In AR, an overall significant decrease in injected drug use was detected between 2009 and 2011 among high school students who had sexual intercourse (t-test p-value=0.01). No change was detected for other risk behaviors or HIV testing.

In NC, there was an overall significant increase between 2009 and 2011 of students reporting that they did not use condom at their last intercourse (t-test p-value<0.01). No change was detected for other risk behaviors or HIV testing.

No significant change in HIV testing or any of the related HIV risk behaviors was detected in CT, NJ, and ND.

Cross-state comparison of HIV testing and related risk behaviors in 2011

The comparison of HIV testing for the year 2011, among high school students who had sexual activity, shows that the prevalence of HIV testing in AR (25.3%, 95% CI: 20.8, 30.4), ND (23.8%, 95% CI: 20.0, 27.9), NC (22.5, 95% CI: 19.1, 26.2), and CT (22.5, 95% CI: 17.8, 27.8) were above the national level (22.2%, 95% CI: 19.7, 24.7) and that the level of HIV testing in MA (21.3%, 95% CI: 17.5, 25.7) and NJ(18.9%, 95% CI: 14.3, 24.4) were below the national level. However, CIs were fairly wide and all overlapped with the national level (Fig. 4).

In 2011, the prevalence of students in CT who had four or more partners in their lifetime was significantly under the national level (CT: 24.9%, 95% CI: 21.5, 28.7; National 32.3%, 95% CI: 30.8, 33.9). The prevalence of students in ND who had their first sexual intercourse before 13 was under the national level (ND: 8.1%, 95% CI: 6.0, 10.9; National 13.2%, 95% CI: 12.2, 14.2). The prevalence of other risk behaviors showed some differences between state and national level, however with overlaps of CIs. MA, NJ, and ND had four out of five risk behaviors with prevalence under the national level and one above. The four risk behaviors surveyed in CT had all average prevalence under the national level, whereas they were all above in NC. Finally, AR had four out of five risk behaviors with prevalence above the national level, and one under. The most important differences compared to the national level were observed for "did not use condom at last intercourse" in AR (+6.2%) and NC (7.0%); "four or more sexual partners in lifetime" in MA (-5.1%), CT (-7.4%), and AR (+5.9%); "first intercourse before 13" in ND (-5.0%).

DISCUSSION

The analysis of the 2005-2011 national YRBS showed that there was an association between HIV testing and related risk behaviors. Overall, students engaging in at least one HIV-related risk behaviors were more likely to have been tested for HIV (between 1.63 and 2.57 times more according to the risk behavior). In addition, among sexually active students, HIV testing was higher for females (1.84 times higher than for males on average), for Black or African Americans (1.60 times higher than for Whites on average), and for higher grade levels (on average 1.13 times higher for grade 10th, 1.22 times higher for 11th grade, 1.41 times higher for 12th grade compared to grade 9th). However, it also highlighted that only a limited proportion (21.9%) of students sexually active, were tested for HIV, although the prevalence of some risk behaviors was higher (35% did not use a condom at last intercourse and 31% had more than 4 sexual partners in their lifetime) . These results aligned with the literature (6). Besides, this analysis, which included the four successive national surveys from 2005 to 2011, showed that there was no significant evolution over time in either HIV testing or risk behaviors that could be detected at the national level.

Although HIV testing is now widely available and routinely recommended among people aged 15-65, there are still significant barriers to widespread HIV testing among adolescents, leading to low testing rates. Peralta et al. showed that low perception of risk and having never been offered a test are main reasons for not having been tested (20). The lack of communication about HIV and sexual life between parents, physicians, school-based health professionals, and youth, lead adolescents to underestimate their own HIV infection risk and limit their access to HIV testing. Being aware of accessible HIV testing sites and knowing the conditions of testing (e.g., confidentiality, parental permission, free services, rapidity of the test and results, parent access to test results, hours of operations) are key elements to enhance HIV testing among adolescents. Peralta et al. showed that increased availability of rapid testing methods, free testing services, and confidentiality are among the factors most highly associated with increased acceptance of HIV testing in a general population of youth (20). In the same line, Lehrer et al. also showed that young people are likely to forgo reproductive health services if parental consent is required (21).

However, teens have low knowledge about consent and confidentiality protections (22), and research suggests this may be in part due to poor communication regarding laws and policies (23, 24). More recently, in 2013, a survey was performed among 164 public HIV testing locations in NYC seeking to determine whether staff at HIV testing locations communicated clearly the policies regarding confidential testing and protection of health information. It showed that for 48% of them, HIV testing was actually not offered or a staff member could not be reached to ask questions about testing options and confidentiality. At the remaining sites, information provided regarding confidentiality, parental consent, and privacy of test results was correct only 69% to 85% of the time.

Additionally, 23% of sites successfully contacted were offering testing exclusively between 9:00 a.m. and 3:00 p.m. weekdays, when most adolescents are in school (25).

In order to break the status quo, school is the best setting to educate adolescents about HIV risk behaviors, the importance of HIV testing, along with its easy and confidential access. Some schools have already taken an active role in this domain.

First, large efforts have been deployed to integrate HIV education in the curriculum. The School Health Policies and Programs Study 2006 indicated that among U.S. high schools, 85% of them teach about HIV transmission as part of required courses, 77% teach how HIV is diagnosed and treated, 76% teach how to find valid information or services regarding HIV or HIV counseling or testing. Some of these programs had a positive impact on knowledge of HIV, HIV prevention and sexual risk behaviors (26).

In addition, school health professionals have a unique opportunity to refer adolescents for HIV testing. In that spirit, school-based HIV testing services have been developed in different places. For example, some school-based health centers across Seattle provide free, on-site clinical services, including HIV and STD counseling and testing. In Philadelphia, all 9th grade and transfer students are offered STD testing at school, in collaboration with the health department. Students who test positive are provided STD treatment at school and referred locally for HIV testing (27).

When moving to the state level, the analysis revealed some variations across geographies and over time in some places. The average levels of HIV testing varied by several percents between some states. In 2011, while HIV testing was 3.3 percents below the national level in NJ (18.9%), it was 3.1 percents above it in AR (25.3%) over the same time period. However, these results should be used with caution as CIs were wide and overlapping. Furthermore, while most of the studied states showed no evolution in either HIV testing or HIV-related risk behaviors, MA, NC, and AR showed some changes overtime. In MA, the level of HIV testing among high schools students who ever had sexual intercourse increased between 2005 and 2011. In addition, a decrease in condom use could be detected in both MA and NC. Additionally, AR showed a decrease in injection drug users between 2009 and 2011.

In MA, NJ, CT, and ND the prevalence of HIV-related risk behaviors was lower compared to the national level. Among these, MA and NJ had a level of HIV testing below the national average, but CT and ND were above. In AR and NC, both prevalence of HIV-related risk behaviors and HIV testing were above the national average. Hence, although the latter two states seem to have promoted HIV testing among adolescents in line with the observed risk behaviors in this population, some further efforts might be needed to reduce risk behaviors in the first place. In particular, condom use was meaningfully lower in AR and NC compared to the national average (respectively 6.2% and 7.0% lower). Also, in AR, the percentage of adolescents who had four or more sexual partners in their lifetime was 5.9% above the national level. As previously mentioned for the level of HIV testing, the CIs for prevalence of HIV-related risk behaviors were wide and overlapping in most cases. In order to be fully conclusive, these analyses would have to be confirmed on larger population samples in order to reduce CIs.

Although current national guidelines recommend routine HIV screening for the population aged 15-65, legislation pertaining to HIV testing is unique for each state (e.g., written consent, confirmatory testing, HIV test results communication, and reporting HIV test results). Therefore, although many state laws have undergone revision since the release in 2006 of the CDC HIV testing recommendations to eliminate conflicting legislation, the full implementation of CDC guidelines varies across states and sites.

For example, the 2006 recommendation of the CDC stipulated that patients in all health care settings be offered opt-out HIV screening without separate written consent and prevention counseling (28). Although many states have revisited legislation to eliminate the need for minors' separate written consent to HIV testing and treatment, Nebraska and New York laws are still not compatible with CDC 2006 guidelines (29).

Along the same lines, state laws for reporting differ from one to another. In some states, HIV positive results may be reported (by name- or code-based reporting) to public health officials, similarly to other reportable conditions. Moreover, some states require that parents and partners be also notified in case of positive results, which may be of concern to adolescents and prevent them from getting tested (29, 30).

The cross-sectional design of YRBS implies some limitations in this study. No temporality between HIV testing and HIV-related risk behaviors can be established. Hence we cannot conclude whether HIV testing is enhanced by the engagement in some risk behaviors, or if HIV testing itself influences the engagement in some of the risk behaviors as well.

In addition, YRBS is a self-reported survey and therefore raises issues of biases related to self reporting (e.g., recall, non-response). However, the release of YRBS data implies a participation rate of at least 60%. Also, several studies evaluated YRBS validity and all led to the same conclusion that its data were reliable (31, 32).

Another limitation relates to the fact that YRBS is limited to students attending high school. Dropouts among high school students varied from 9.4% in 2005 to 7.1% in 2011, with significantly higher dropout among Hispanics (ranged from 22.2% in 2005 to 13.6% in 2011) and Black or African American ethnicities (ranged from 10.4% in 2005 to 7.3% in 2011) (33).

Additionally, the impossibility to distinguish gay, lesbian, and heterosexual subpopulations brings another limitation, as these subpopulations face significantly different HIV transmission risk, with a greater risk for males having sex with males. Therefore, accessing this information would be very informative.

At last, National YRBS data are useful for characterizing HIV testing trends nationwide. However, state data are critical to have a comprehensive picture of population diagnosis needs and thereby plan for adequate interventions in these populations. However, among YRBS data available, some states do not include some questions considered too sensitive (e.g., sexual-related questions, HIV testing). Thus, the information about HIV testing between 2005 and 2011 was accessible only for a very limited number of states with further variations across years. It included only 6 states in 2011, 7 in 2009, and 2 in 2005 and 2007. Adding the HIV testing question to the YRBS questionnaire in all states, would help getting a more complete and accurate picture of HIV testing in the US, identifying potential gaps in testing, and prioritizing interventions accordingly.

In conclusion, this study showed that most sexually active high school students were never tested for HIV between 2005 and 2011. No change over time was observed at the national level in both HIV testing and HIV-related risk behaviors. This study also demonstrated the consistent association between HIV testing and engaging in HIV-related risk behaviors among students who had sexual intercourse. This supported previous findings for the year 2009 (6). No temporality between HIV testing and engagement in risk behaviors could be established though.

In addition, it highlighted that national and state level analyses bring complementary information. State analysis brings into light specific state changes, which are not detectable at the national level. Thus, between 2005 and 2011, an increase of HIV testing was detected in MA, a decrease of condom use was detected in MA and NC, and a decrease of injection drug use was detected in AR. State needs differ from one another and state level analysis brings valuable information to guide local prevention efforts and state policies. For example, promoting condom use is an effective HIV prevention measure, which could be reinforced in AR and NC based on the gap in use between these states and the national level. Thereby, this highlights how rigorous collection and analysis process are cornerstones to inform local interventions and evaluate their impact.

This study also revealed that some states do not include some questions considered too sensitive. However, including questions about HIV testing and related risk behaviors in YRBS, ultimately helps states in their HIV prevention action. It allows measuring progress and prioritizing needs.

Moreover, adding some questions around the type of sexuality (e.g., lesbian, gay, heterosexual) to the current YRBS questionnaire would be of great value in order to trigger needs in subpopulations at higher risk of HIV transmission.

The implementation of routine HIV, as recommended by the CDC, is an efficient strategy to allow early linkage to care of HIV positive adolescents and reduce the rate of new HIV infection. At that time, strategies aiming at enhancing HIV testing among adolescents require to provide students with convenient, confidential, rapid and free access to HIV testing. Last, combined prevention efforts should target specific risk behavior according to state needs (e.g., promotion of condom use in AR and NC).

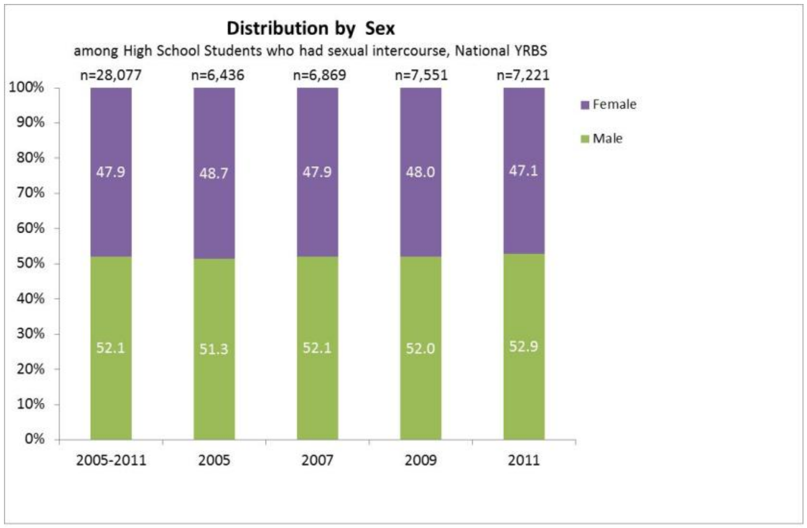
TABLES AND FIGURES

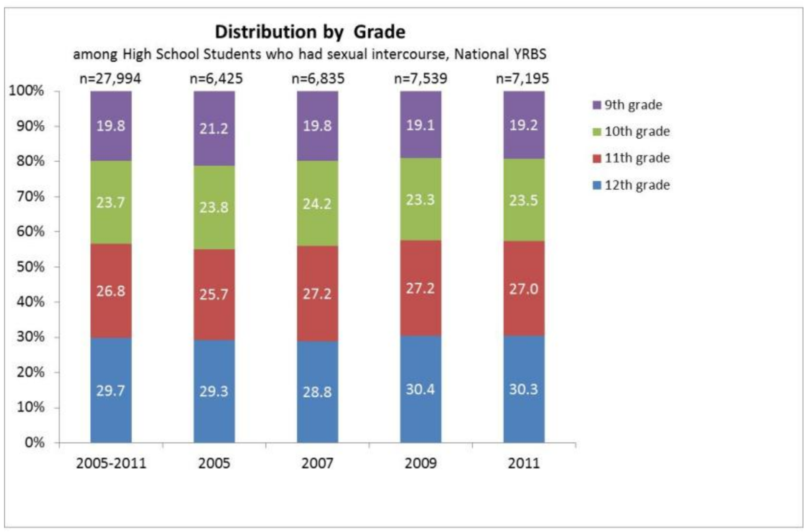
Table 1. Characteristics of High School Students who had sexual intercourse, National YRBS 2005-2011

	Had sexual intercourse, 2005-2011		
	No.	%	95% CI
Overall	28,177		
Sex of the subject			
Female	13,614	47.9	47.1-48.7
Male	14,463	52.1	51.3-52.9
Missing	100		
Grade of the subject			
9th grade	4,749	19.8	18.9-20.6
10th grade	6,147	23.7	23.0-24.4
11th grade	7,944	26.8	26.1-27.5
12th grade	9,154	29.7	28.9-30.6
Missing	183		
Race and ethnicity			
White	10,684	55.6	52.4-58.8
Black or African Americ	6,834	18.7	16.6-21.1
Hispanic	8,106	18.7	16.8-20.7
Other	2,085	6.9	5.9-8.1
Missing	468		
Ever been tested for HIV			
Yes	5,706	21.9	21.0-22.8
No, not sure	18,801	78.1	77.2-79
Missing	3,670		
Used condom at last sexual intercourse			
No	9,919	35.4	34.3-36.4
Yes	17,614	64.6	63.6-65.7
Missing	644		
Had 4 or more sexual partners in life			
Yes	9,116	31.1	30.1-32.0
No	18,771	68.9	68.0-69.9
Missing	290		
First sexual intercourse before 13			
Yes	3,966	13.5	12.8-14.2
No	24,077	86.5	85.8-87.2
Missing	134		
Ever forced to have intercourse			
Yes	3,856	14.1	13.4-14.8
No	24,156	85.9	85.2-86.6
Missing	165		
Lifetime illegal injection drug use			
Yes	930	3.5	3.2-3.9
No	26,630	96.5	96.1-96.8
Missing	617		

Abbreviations: CI, confidence interval; HIV, human immunodeficiency virus.

Figure 1. Distribution of Demographics for High School Students Who Had Sexual Intercourse, National YRBS, 2005-2011





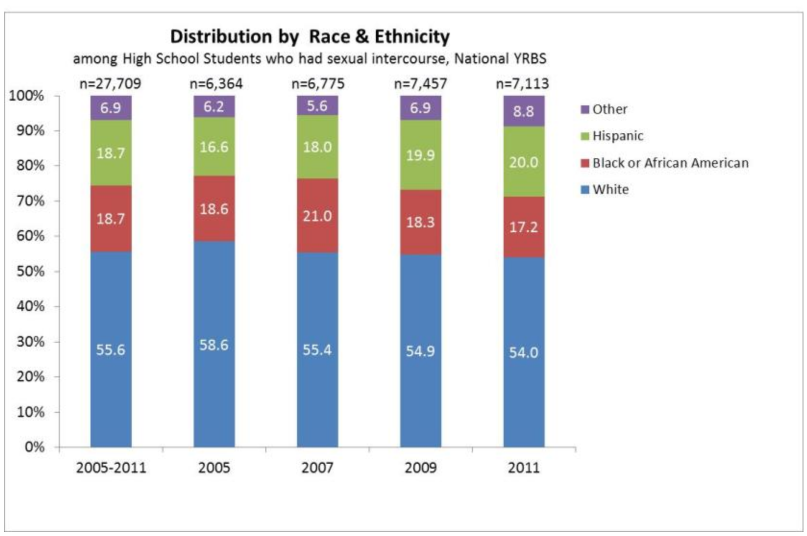


Figure 2. Trends in HIV Testing and Related Risk Behaviors for High School Students Who Had Sexual Intercourse, National YRBS, 2005-2011

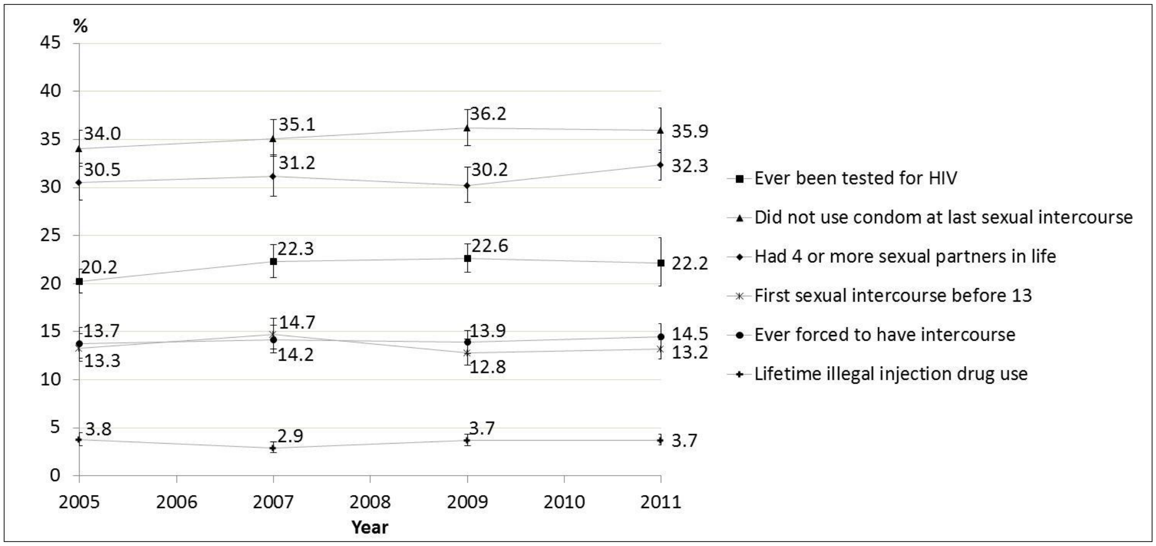


Table 2. Trends in HIV testing and Related Risk Behaviors for High School Students Who Had Sexual Intercourse, National YRBS, 2005-2011

National	2005	2007	2009	2011	Linear change
Ever been tested for HIV					
	20.2	22.3	22.6	22.2	No change, 2005-2011
	19.0-21.5	20.6-24.1	21.2-24.1	19.7-24.8	
Did not use condom at last sexual intercourse					
	34.0	35.1	36.2	35.9	No change, 2005-2011
	32.2-35.9	33.2-37.1	34.4-38.1	33.6-38.3	
4+ sexual partners in life					
	30.5	31.2	30.2	32.3	No change, 2005-2011
	28.7-32.5	29.1-33.4	28.4-32.1	30.8-33.9	
First intercourse before 13					
	13.3	14.7	12.8	13.2	No change, 2005-2011
	11.9-14.8	13.2-16.4	11.5-14.2	12.2-14.2	
Ever forced to have intercourse					
	13.7	14.2	13.9	14.5	No change, 2005-2011
	12.2-15.4	12.8-15.6	12.7-15.1	13.2-15.8	
Injected drug 1+ times in life					
	3.8	2.9	3.7	3.7	No change, 2005-2011
	3.1-4.5	2.4-3.5	3.1-4.3	3.2-4.4	

Table 3. Unadjusted Association of HIV Testing with Related Risk Behaviors, Controlling for Demographics Among High School Students Who Had Sexual Intercourse, National YRBS, 2005-2011

	All		2005		2007		2009		2011	
	OR	95% CI	OR	95% CI	OR	95% CI	OR	95% CI	OR	95% CI
Sex of the subject										
Male	1.00	referent	1.00	referent	1.00	referent	1.00	referent	1.00	referent
Female	1.72	1.59-1.86	1.70	1.43-2.02	1.68	1.48-1.92	1.74	1.50-2.02	1.75	1.47-2.07
Grade of the subject										
9th grade	1.00	referent	1.00	referent	1.00	referent	1.00	referent	1.00	referent
10th grade	1.10	0.96-1.26	1.17	0.87-1.57	1.16	0.90-1.51	1.15	0.92-1.45	0.93	0.68-1.28
11th grade	1.19	1.04-1.36	1.11	0.87-1.42	1.31	0.92-1.85	1.33	1.07-1.65	1.00	0.78-1.28
12th grade	1.39	1.22-1.58	1.29	0.97-1.71	1.59	1.26-2.00	1.59	1.25-2.02	1.10	0.82-1.47
Race and ethnicity										
White	1.00	referent	1.00	referent	1.00	referent	1.00	referent	1.00	referent
Black or African American	1.67	1.45-1.92	1.67	1.34-2.08	1.67	1.41-1.97	1.48	1.13-1.95	1.93	1.31-2.83
Hispanic	1.00	0.90-1.12	1.01	0.82-1.24	0.94	0.76-1.15	1.01	0.81-1.26	1.03	0.84-1.28
Other	1.17	1.01-1.36	1.30	0.90-1.90	1.34	0.98-1.84	0.72	0.56-0.94	1.43	1.08-1.90
Used condom at last sexual intercourse										
Yes	1.00	referent	1.00	referent	1.00	referent	1.00	referent	1.00	referent
No	1.63	1.51-1.76	1.53	1.26-1.84	1.67	1.44-1.94	1.70	1.47-1.96	1.61	1.40-1.86
Had 4 or more sexual partners in life										
No	1.00	referent	1.00	referent	1.00	referent	1.00	referent	1.00	referent
Yes	2.59	2.40-2.79	2.71	2.25-3.26	2.54	2.26-2.86	2.53	2.18-2.94	2.60	2.19-3.08
First sexual intercourse before 13										
No	1.00	referent	1.00	referent	1.00	referent	1.00	referent	1.00	referent
Yes	1.67	1.51-1.85	1.65	1.30-2.09	1.68	1.37-2.06	1.63	1.33-1.99	1.74	1.42-2.14
Ever forced to have intercourse										
No	1.00	referent	1.00	referent	1.00	referent	1.00	referent	1.00	referent
Yes	2.24	2.03-2.47	2.17	1.78-2.63	2.13	1.74-2.60	2.21	1.83-2.67	2.45	2.03-2.96
Lifetime illegal injection drug use										
No	1.00	referent	1.00	referent	1.00	referent	1.00	referent	1.00	referent
Yes	2.51	2.07-3.04	2.08	1.30-3.33	2.81	1.89-4.18	2.52	1.84-3.44	2.71	1.82-4.04

Table 4. Interaction Assessment of HIV-Related Risk Behaviors with Demographics and Years by Strata, National YRBS, 2005-2011

	Used condom at last sexual intercourse		Had 4 or more sexual partners in life		First sexual intercourse before 13		Ever forced to have intercourse		Lifetime illegal injection drug use	
	OR	95% CI	OR	95% CI	OR	95% CI	OR	95% CI	OR	95% CI
Sex of the subject										
Female	1.47	1.32-1,64	-	-	0.97	0.81-1.16	1.45	1.13-1.86	-	-
Male	1.08	0.94-1,24	-	-	1.45	1.23-1.70	1.10	0.81-1.50	-	-
Grade of the subject										
9th grade	-	-	-	-	-	-	1.44	1.24-1.67	-	-
10th grade	-	-	-	-	-	-	1.70	1.24-2.32	-	-
11th grade	-	-	-	-	-	-	2.06	1.50-2.84	-	-
12th grade	-	-	-	-	-	-	1.32	1.01-1,72	-	-
Race and ethnicity										
White	-	-	2.50	2.21-2.82	-	-	-	-		
Black or African American	-	-	1.73	1.49-2.02	-	-	-	-	-	-
Hispanic	-	-	1.99	1.65-2.39	-	-	-	-	-	-
Other	-	-	2.71	2.05-3.58	-	-	-	-	-	-
Year									-	-
2005	-	-	-	-	-	-	-	-	-	-
2007	-	-	-	-	-	-	-	-	-	-
2009	-	-	-	-	-	-	-	-	-	-
2011	-	-	-	-	-	-	-	-	-	-

Abbreviations: CI, confidence interval.

Table 5. Adjusted Association of HIV Testing with Related Risk Behaviors, Controlling for Demographics Among High School Students Who Had Sexual Intercourse, National YRBS, 2005-2011

	All		2005		2007		2009		2011	
	OR	95% CI	OR	95% CI	OR	95% CI	OR	95% CI	OR	95% CI
Sex of the subject										
Male	1.00	referent	1.00	referent	1.00	referent	1.00	referent	1.00	referent
Female	1.84	1.68-2.01	1.77	1.46-2.14	1.78	1.52-2.08	1.94	1.64-2.29	1.85	1.55-2.21
Grade of the subject										
9th grade	1.00	referent	1.00	referent	1.00	referent	1.00	referent	1.00	referent
10th grade	1.13	0.98-1.30	1.16	0.83-1.61	1.22	0.89-1.66	1.20	0.94-1.53	0.95	0.70-1.27
11th grade	1.22	1.06-1.40	1.06	0.79-1.41	1.39	0.95-2.04	1.42	1.14-1.77	0.99	0.80-1.24
12th grade	1.41	1.23-1.62	1.27	0.95-1.69	1.68	1.24-2.27	1.64	1.27-2.11	1.07	0.82-1.38
Race and ethnicity										
White	1.00	referent	1.00	referent	1.00	referent	1.00	referent	1.00	referent
Black or African American	1.60	1.37-1.87	1.59	1.26-2.00	1.59	1.31-1.93	1.41	1.04-1.92	1.90	1.24-2.92
Hispanic	0.99	0.89-1.11	1.03	0.83-1.27	0.91	0.72-1.14	0.99	0.79-1.26	1.03	0.83-1.29
Other	1.10	0.94-1.28	1.26	0.87-1.81	1.27	0.92-1.74	0.65	0.50-0.86	1.38	1.04-1.84
Used condom at last sexual intercourse										
Yes	1.00	referent	1.00	referent	1.00	referent	1.00	referent	1.00	referent
No	1.30	1.20-1.41	1.23	1.02-1.49	1.36	1.17-1.59	1.28	1.09-1.51	1.29	1.13-1.48
Had 4 or more sexual partners in life										
No	1.00	referent	1.00	referent	1.00	referent	1.00	referent	1.00	referent
Yes	2.27	2.08-2.47	2.39	1.95-2.92	2.11	1.80-2.46	2.35	2.03-2.72	2.33	1.94-2.81
First sexual intercourse before 13										
No	1.00	referent	1.00	referent	1.00	referent	1.00	referent	1.00	referent
Yes	1.22	1.09-1.38	1.19	0.96-1.48	1.34	1.03-1.74	1.23	0.98-1.55	1.13	0.93-1.39
Ever forced to have intercourse										
No	1.00	referent	1.00	referent	1.00	referent	1.00	referent	1.00	referent
Yes	1.50	1.35-1.67	1.46	1.14-1.86	1.44	1.13-1.84	1.45	1.22-1.72	1.73	1.44-2.08
Lifetime illegal injection drug use										
No	1.00	referent	1.00	referent	1.00	referent	1.00	referent	1.00	referent
Yes	1.68	1.34-2.10	1.54	0.95-2.49	1.87	1.15-3.04	1.77	1.20-2.61	1.54	0.94-2.54

Table 6. Availability of State Level Datasets Collecting HIV Testing Information, YRBS, 2005-2011

	2005	2007	2009	2011
AR			+	+
CT		+	+	+
NJ	+		+	+
MA	+	+	+	+
NC			+	+
ND			+	+
SC			+	

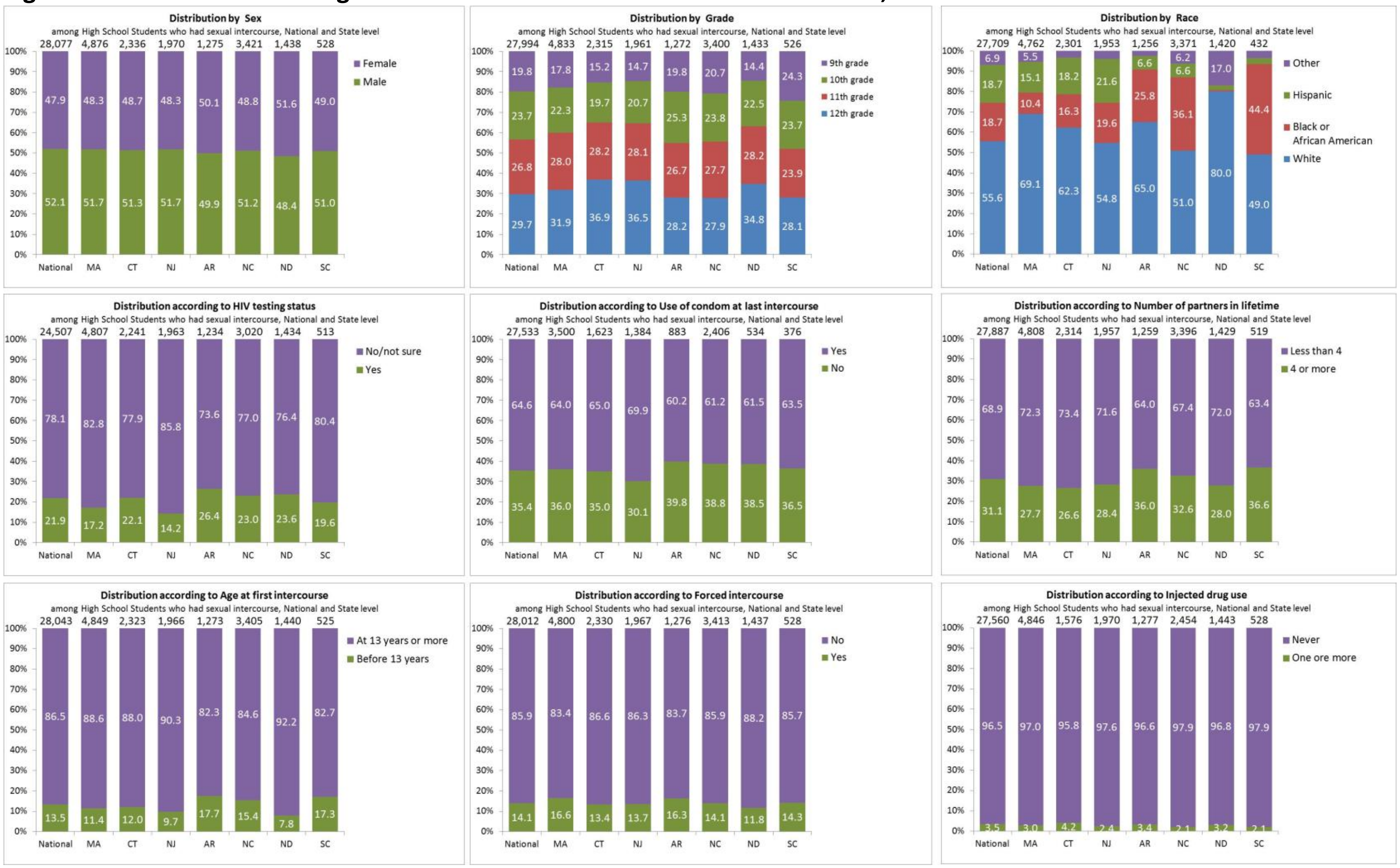
Note: HIV testing question was not included in 2005-2011 YRBS conducted by the following states (n=36): AL, AK, AZ, CO, DE, FL, GA, ID, IL, IA, KS, KY, LA, ME, MD, MI, MS, MO, MT, NE, NV, NH, NM, NY, OH, OK, PA, RI, SD, TN, TX, UT, VA, WV, WI, WY.
CA, MN, OR, and WA had either no weighted data available or did not conduct YRBS in the 2005-2011 period. YRBS datasets for HI, IN, and VT were not accessible.

Table 7. Characteristics of High School Students who Had Sexual Intercourse, State YRBS 2005-2011

	MA 2005-2011			CT 2007-2011			NJ 2005, 2007, 2011			AR 2009-2011			NC 2009-2011			ND 2009-2011			SC 2009		
	No.	%	95% CI	No.	%	95% CI	No.	%	95% CI	No.	%	95% CI	No.	%	95% CI	No.	%	95% CI	No.	%	95% CI
Overall	4,884			2,342			1,972			1,282			3,434			1,445			529		
Sex of the subject																					
Female	2,399	48.3	46.5-50.1	1,168	48.7	46.1-51.2	1,038	48.3	45.2-51.3	641	50.1	46.0- 50.3	1,730	48.8	46.5-51.1	767	51.6	48.3-54.9	268	49.0	44.0- 53.9
Male	2,477	51.7	49.9-53.5	1,168	51.3	48.8-53.9	932	51.7	48.7-54.8	634	49.9	45.7- 54.0	1,691	51.2	48.9-53.5	671	48.4	45.1-51.7	260	51.0	46.1-56.0
Missing	8			6			2			7			13			7			1		
Grade of the subject																					
9th grade	822	17.8	15.4-20.4	384	15.2	11.6-19.6	287	14.7	11.7-18.3	347	19.8	13.0- 29.1	742	20.7	16.5- 25.6	252	14.4	10.8-19.0	123	24.3	16.1-35.0
10th grade	1,101	22.3	19.6-25.2	544	19.7	16.8-23.0	431	20.7	16.7-25.3	390	25.3	20.4-30.9	910	23.8	20.1-27.9	342	22.5	17.7-28.2	90	23.7	15.0-35.2
11th grade	1,438	28.0	24.8- 31.5	662	28.2	23.5-33.5	569	28.1	23.8-32.8	361	26.7	21.5- 32.6	871	27.9	22.4-34.1	413	28.2	22.4-34.8	135	23.9	17.2-32.2
12th grade	1,472	31.9	28.1-36.0	725	36.9	33.1-40.8	674	36.5	30.9-42.6	174	28.2	23.0-34.0	877	27.7	22.7-33.2	426	34.8	28.8-41.4	178	28.1	20.8-36.7
Missing	1,523			752			685			10			34			12			3		
Race and ethnicity																					
White	3,000	69.1	64.1-73.6	1,319	62.3	55.6-68.5	1,037	54.8	45.8-63.4	745	65.0	56.0-73.1	1,932	51.0	43.8-58.1	1,157	80.0	73.6-85.2	230	49.0	38.4-59.7
Black or African American	450	10.4	8.0-13.4	309	16.3	13.1-20.0	297	19.6	13.5-27.7	297	25.8	18.0-35.5	770	36.1	29.3-43.5	22	0.8	0.5-1.2	231	44.4	33.8-55.5
Hispanic	832	15.1	12.6-18.0	540	18.2	14.4-22.7	468	21.6	17.2-26.8	129	6.6	4.9- 8.9	327	6.6	5.5-8.0	62	2.1	1.6- 2.9	28	3.1	2.4-4.1
Other	480	5.5	4.5-6.6	133	3.2	2.7-3.9	151	4.0	3.26-4.9	85	2.5	2.0-3.3	342	6.2	4.4-8.8	179	17.0	12.0-23.6	32	3.4	2.4-4.8
Missing	122			41			19			26			63			25			8		
Ever been tested for HIV																					
Yes	795	17.2	15.5-19.0	489	22.1	19.3-25.1	285	14.2	11.8-17.0	314	26.4	23.5-29.5	686	23.0	20.8-25.3	325	23.6	21.4-26.0	106	19.6	15.7-24.1
No, not sure	4,012	82.8	81.0-84.5	1,752	77.9	74.9-80.7	1,678	85.8	83.0-88.2	920	73.6	70.5- 76.5	2,334	77.0	74.7-79.2	1,109	76.4	74.0-78.6	407	80.4	75.8-84.3
Missing	77			101			9			48			414			11			16		
Used condom at last sexual intercourse																					
No	1,723	36.0	34.1-38.0	783	35.0	32.4-37.6	597	30.1	27.3-33.1	491	39.8	35.9- 43.8	1,285	38.8	36.6-41.0	550	38.5	35.7- 41.4	180	36.5	30.5-42.9
Yes	3,072	64.0	62.0-65.9	1,484	65.0	62.4-67.6	1,322	69.9	66.9-72.7	743	60.2	56.2- 64.1	2,111	61.2	59.0-63.4	869	61.5	58.6-64.3	333	63.5	57.1-69.5
Missing	89			75			53			48			38			26			16		
Had 4 or more sexual partners in life																					
Yes	1,330	27.7	25.56-30.0	586	26.6	24.6-28.8	548	28.4	25.1-31.9	433	36.0	32.7-39.5	1,010	32.6	29.9-35.5	392	28.0	25.3-30.9	197	36.6	32.2-41.2
No	3,478	72.3	70.0-74.4	1,728	73.4	71.2-75.4	1,409	71.6	68.1-74.9	826	64.0	60.5-67.3	2,386	67.4	64.5-70.1	1,037	72.0	69.1-74.7	322	63.4	58.8-67.8
Missing	76			28			15			23			38			16			10		
First sexual intercourse before 13																					
Yes	570	11.4	10.3-12.5	289	12.0	10.0-14.3	183	9.7	8.2-11.5	247	17.7	14.0-22.1	518	15.4	13.6-17.3	115	7.8	6.4- 9.6	91	17.3	13.9-21.2
No	4,279	88.6	87.5-89.7	2,034	88.0	85.7-90.0	1,783	90.3	88.5-91.8	1,026	82.3	77.9-86.0	2,887	84.6	82.7-86.4	1,325	92.2	90.4-93.6	434	82.7	78.8-86.1
Missing	35			19			6			9			29			5			4		
Ever forced to have intercourse																					
Yes	812	16.6	15.5-17.8	317	13.4	12.4-14.5	282	13.7	11.8-15.8	221	16.3	13.4-19.7	534	14.1	12.3-16.1	190	11.8	9.9-13.9	73	14.3	11.1-18.2
No	3,988	83.4	82.2-84.5	2,013	86.6	85.5-87.6	1,685	86.3	84.2-88.2	1,055	83.7	80.3-86.6	2,879	85.9	83.9-87.7	1,247	88.2	86.1-90.1	455	85.7	81.8-88.9
Missing	84			12			5			6			21			8			1		
Lifetime illegal injection drug use																					
Yes	152	3.0	2.4-3.7	63	4.2	2.8-6.3	47	2.4	1.9- 3.0	51	3.4	2.29- 4.9	95	2.1	1.4-3.4	49	3.2	2.2-4.5	13	2.1	1.2- 3.7
No	4,694	97.0	96.3-97.6	1,513	95.8	93.7-97.2	1,923	97.6	97.0-98.1	1,226	96.6	95.-97.7	2,359	97.9	96.6-98.6	1,394	96.8	95.5-97.8	515	97.9	96.3-98.8
Missing	38			766			2			5			980			2			1		

Abbreviations: CI, confidence interval; HIV, human immunodeficiency virus.

Figure 3. Characteristics of High School Students who Had Sexual Intercourse, National and State YRBS 2005-2011



Note: Drawn from available datasets pooled: for each state: NJ: 2005, 2007, and 2011, MA: 2005, 2007, 2009, and 2011, CT: 2007, 2009, and 2011, NC: 2009 and 2011, ND: 2009 and 2011, AR: 2009 and 2011, SC: 2009

Table 8. Trends in HIV Testing and Related Risk Behaviors for High School Students Who Had Sexual Intercourse, State YRBS, 2005-2011

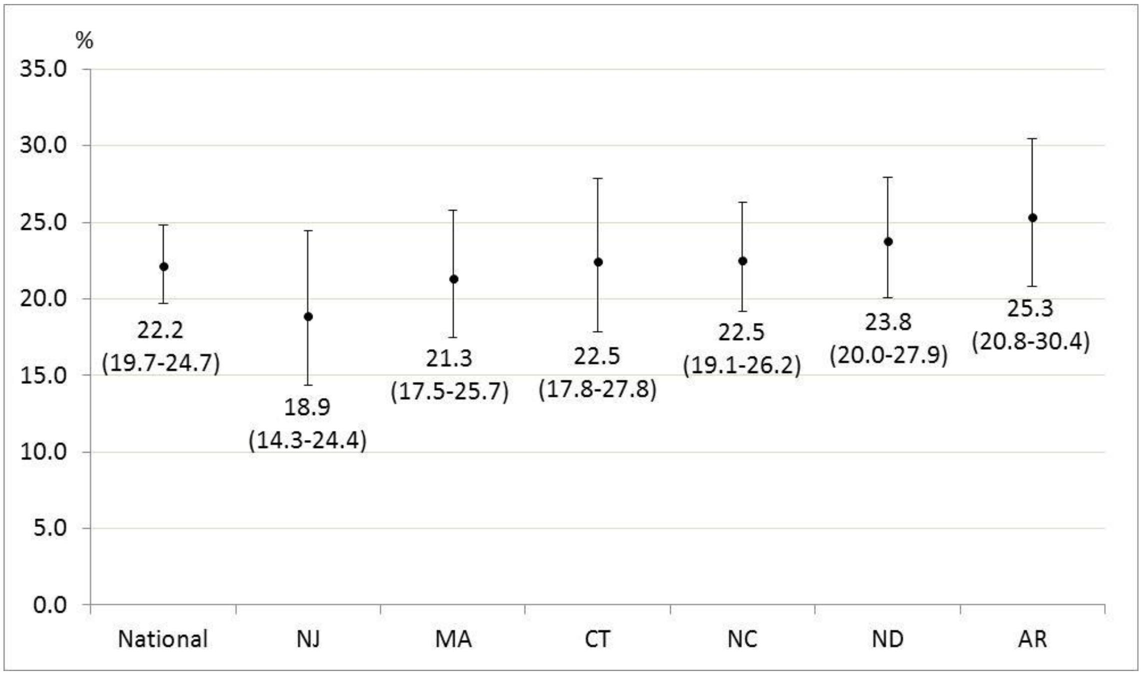
MA	2005	2007	2009	2011	Linear change	t-test
Ever been tested for HIV						
	4.1	22.0	21.5	21.3	Increased, 2005-2011	-
	2.9-5.8	19.2-25.1	18.7-24.6	17.5-25.8		
Did not use condom at last sexual intercourse						
	31.7	35.1	39.1	38.4	Increased, 2005-2011	-
	28.7-34.8	32.2-38.2	36.1-42.1	34.6-42.3		
Had 4 or more sexual partners in life						
	27.9	27.9	27.9	27.2	No change, 2005-2011	-
	23.4-33.0	24.7-31.5	24.6-31.4	22.9-31.9		
First intercourse before 13						
	11.2	13.3	11.2	9.7	No change, 2005-2011	-
	9.1-13.7	11.1-15.9	8.9-14.0	7.6-12.4		
Ever forced to have intercourse						
	15.2	19.0	18.0	14.2	No change, 2005-2011	-
	12.9-17.8	16.5-21.8	15.3-21.0	11.8-17.1		
Lifetime illegal injection drug use						
	2.1	4.3	3.2	2.3	No change, 2005-2011	-
	1.4-3.2	3.1-6.0	2.1-4.8	1.5-3.5		
CT	**2005**	**2007**	**2009**	**2011**	**Linear change**	**t-test**
Ever been tested for HIV						
		22.3	21.5	22.5	No change, 2007-2011	-
		18.8-26.2	17.7-25.8	17.9-27.8		
Did not use condom at last sexual intercourse						
		34.1	35.9	34.9	No change, 2007-2011	-
		28.9-39.8	31.7-40.3	30.8-39.1		
Had 4 or more sexual partners in life						
		29.2	25.7	24.9	No change, 2007-2011	-
		25.2-33.5	21.4-30.5	21.5-28.7		
First intercourse before 13						
		13.5	11.3	11.3	No change, 2007-2011	-
		10.2-17.6	8.5-14.8	8.5-14.8		
Ever forced to have intercourse						
		14.8	13.1	12.3	No change, 2007-2011	-
		12.2-17.7	10.6-16.1	10.4-14.6		
Lifetime illegal injection drug use						
		4.7	3.8		-	No change, 2007-2009
		2.9-7.4	2.3-6.1			
NJ	**2005**	**2007**	**2009**	**2011**	**Linear change**	**t-test**
Ever been tested for HIV						
	4.4		18.9	18.9	-	No change, 2009-2011
	2.4-7.8		15.5-22.8	14.4-24.4		
Did not use condom at last sexual intercourse						
	25.0		31.0	34.2	-	No change, 2009-2011
	20.3-30.5		27.1-35.1	29.0-39.8		
Had 4 or more sexual partners in life						
	26.4		27.3	31.4	-	No change, 2009-2011
	20.3-33.5		23.9-30.9	24.9-38.8		
First intercourse before 13						
	9.7		8.5	11.1	-	No change, 2009-2011
	6.3-14.8		6.0-11.9	8.1-15.0		
Ever forced to have intercourse						
	13.8		12.3	15.1	-	No change, 2009-2011
	11.0-17.0		10.0-15.0	11.5-19.6		
Lifetime illegal injection drug use						
	0.9		3.0	3.1	-	No change, 2009-2011
	0.3-2.5		2.0-4.6	2.0-4.7		

AR	2005	2007	2009	2011	Linear change	t-test
Ever been tested for HIV						
			27.4	25.3	-	No change, 2009-2011
			23.8-31.4	20.8-30.4		
Did not use condom at last sexual intercourse						
			37.6	42.1	-	No change, 2009-2011
			32.2-43.4	36.5-47.9		
Had 4 or more sexual partners in life						
			34.0	38.2	-	No change, 2009-2011
			29.2-39.1	34.0-42.6		
First intercourse before 13						
			18.9	16.4	-	No change, 2009-2011
			14.7-23.9	11.6-22.8		
Ever forced to have intercourse						
			17.9	14.7	-	No change, 2009-2011
			14.2-22.3	11.7-18.3		
Lifetime illegal injection drug use						
			4.7	2.0	-	Decreased, 2009-2011
			3.2-7.0	1.0-3.9		

NC	2005	2007	2009	2011	Linear change	t-test
Ever been tested for HIV						
			23.3	22.5	-	No change, 2009-2011
			20.7-26.2	19.2-26.3		
Did not use condom at last sexual intercourse						
			35.0	42.9	-	Increased, 2009-2011
			32.6-37.5	39.7-46.1		
Had 4 or more sexual partners in life						
			30.9	34.4	-	No change, 2009-2011
			27.6-34.6	30.2-38.9		
First intercourse before 13						
			14.4	16.4	-	No change, 2009-2011
			12.3-16.9	13.7-19.6		
Ever forced to have intercourse						
			13.7	14.5	-	No change, 2009-2011
			11.5-16.2	11.7-18.0		
Lifetime illegal injection drug use						
			2.1		-	-
			1.4-3.4			

ND	2005	2007	2009	2011	Linear change	t-test
Ever been tested for HIV						
			23.4	23.8	-	No change, 2009-2011
			20.3-26.8	20.1-27.9		
Did not use condom at last sexual intercourse						
			40.4	36.6	-	No change, 2009-2011
			36.1-44.8	32.8-40.5		
Had 4 or more sexual partners in life						
			26.4	29.6	-	No change, 2009-2011
			21.9-31.5	25.2-34.4		
First intercourse before 13						
			7.5	8.1	-	No change, 2009-2011
			5.6-10.1	6.0-10.9		
Ever forced to have intercourse						
			11.1	12.4	-	No change, 2009-2011
			8.5-14.4	9.9-15.4		
Lifetime illegal injection drug use						
			3.5	2.8	-	No change, 2009-2011
			2.2-5.7	1.6-5.0		

Figure 4. Cross State Comparison of HIV Testing Among High School Students who Had Sexual Intercourse, National and State YRBS, 2011



APPENDICES

Appendix A: YRBS questionnaire

National Youth Risk Behavior Survey (YRBS) data are available at:
http://www.cdc.gov/HealthyYouth/yrbs/data/

National questionnaire for 2011:
ftp://ftp.cdc.gov/pub/data/yrbs/2011/YRBS_2011_National_User_Guide.pdf

Appendix B: Collinearity macro

```
 OPTIONS MPRINT SYMBOLGEN mlogic;

* COLLINEARITY DIAGNOSTICS USING THE INFORMATION MATRIX;
* Original MACRO FROM SAS-L BY MATHEW ZACK;
* Modified 26 April 2005 by Jim Singleton to handle covariates included in class statement;
* Modified November 2010 by Kristin Wall and Kevin Delaney to
          INCLUDE CODE FOR GENMOD, MIXED and GLIMMIX
           and
          Explicitly name the output dataset containing Collinearity Diagnostics;

*The MACRO contains four named parameters:

          COVDSN=DATASETNAME is the input dataset containing the Covariance Matrix output from
          the LOGISTIC, MIXED, GLIMMIX, PHREG or GENMOD procedures

          OUTPUT=DATASETNAME is the name you choose to contain the output Collinearity Diagnostics

          PROCDR=SAS procedure that produced the collinearity Matrix output.
          Currently, Valid values include:
          Logistic, Phreg, Genmod, Glimmix and Mixed
          To maintain consistency with previous versions of the MACRO, this parameter is not required for
Logistic or Phreg

          PARMINFO=Dataset generated with the Statement:
                    ods output genmod.parminfo=parms
          That contains the names of the variables included in the model;

*IN PROC LOGISTIC OR PROC PHREG SPECIFY THE COVOUT AND THE OUTEST=DATASETNAME ;
*  OPTIONS IN THE PROC STATEMENT.
* IF USING LOGISTIC OR PHREG Only COVDSN and OUTPUT are required;
*    %COLLIN(COVDSN=DATASETNAME, OUTPUT=DATASETNAME2);


*IF USING PROC LOGISTIC for CONDITIONAL LOGISTIC REGRESSION ie. for Matched data you need to tell
SAS to
```

```
            Drop the intercept column from the COVOUT DATASET: ;
            /*          proc logistic data=Data1 covout outest=test(DROP=intercept);
                        strata ID;
                        model outcome(event='1')=Gall hyper;
                        run;
                        */
*Doing so will allow you to use the COVOUT DATASET for conditional Logistic regression like any other
Logistic
            output:
             %COLLIN(COVDSN=DATASETNAME, OUTPUT=DATASETNAME2); *In the example above
COVDSN=TEST;

*If using PROC SURVEYLOGISTIC add the follwing ODS OUTPUT STATEMENT to your code;
 *ods output surveylogistic.covb=DATASETNAME;
*Also, add the /covb option to the MODEL statement, e.g. model outcome=exp covars/covb;
*Then call the Macro as:
            %Collin(COVDSN=DATASETNAME,PROCDR=SURVEYLOGISTIC,OUTPUT=DATASETNAME2);

* When using this Macro with GLIMMIX:
* Use the /covb option after the model statement and include the line of code:  ods output
glimmix.CovB=DATASETNAME;
* Call macro as: %COLLIN(COVDSN=DATASETNAME, PROCDR=GLIMMIX, OUTPUT=);

*When using the Macro with GENMOD:  ;

* IF REPEATED IS NOT USED (UNCLUSTERED DATA -> NO GEE) THEN ;
* ADD COVB TO THE MODEL STATEMENT (MODEL / COVB) and include the following two statements
immediately
            Before  ;
*ods output genmod.parminfo=parms;
*ods output genmod.covb=covdsn;

* IF REPEATED IS USED FOR CLUSTERED DATA THEN ;
* ADD COVB TO THE REPEATED STATEMENT (REPEATED / COVB);
 *ods output genmod.parminfo=parms;
 *ods output genmod.geercov=covdsn;

* When using GENMOD Call the MACRO with PROCDR=GENMOD  and PARMINFO=parms
                        (from the ODS OUTPUT STATEMENT)
*    %COLLIN(COVDSN=COVDSN, PROCDR=GENMOD, PARMINFO=Parms, OUTPUT=DATASETNAME);

* When using this Macro with MIXED:
* Use the /covb option after the model statement and include the line of code:  ods output
mixed.CovB=DATASETNAME;
* Call macro as: %COLLIN(COVDSN=DATASETNAME, PROCDR=MIXED, OUTPUT=);


%MACRO COLLIN(COVDSN=, PROCDR=, PARMINFO=,OUTPUT=);

%* MACRO TO CALCULATE COLLINEARITY DIAGNOSTICS FROM ;
%*  VARIANCE-COVARIANCE MATRIX IN NONLINEAR REGRESSION;

%* REF: DAVIS CE, HYDE JE, BANGDIWALA SI, NELSON JJ.;
```

```
%*      AN EXAMPLE OF DEPENDENCIES AMONG VARIABLES IN A;
%*      CONDITIONAL LOGISTIC REGRESSION.  IN: MOOLGAVKAR SH,;
%*      PRENTICE RL, EDS.  MODERN STATISTICAL METHODS IN;
%*      CHRONIC DISEASE EPIDEMIOLOGY.  NEW YORK:;
%*      JOHN WILEY & SONS, INC., 1986:140-7.;



%let Drop=%str();

%* MAKE GENMOD COVARIANCE OUTPUT SIMILAR ENOUGH TO LOGISTIC AND PHREG THAT THIS MACRO
WILL
%* WORK.;

%IF %Upcase(&PROCDR)=GENMOD %THEN %DO;

%* FOR SOME INEXPLICABLE REASON, SAS DOES NOT RECORD THE VARIABLE NAMES IN THE OUTPUT;
%* VARIANCE-COVARIANCE DATA SET. THIS NEXT SECTION OF CODE REPLACES THE PARM VARIABLE;
%* WITH THE NAMES OF THE VARIABLES AND RENAMES PARM TO _NAME_ TO CONFORM TO THE
OUTPUT;
%* DATA SETS GENERATED BY LOGISTIC AND GENMOD.;

%* IF THERE ARE MORE THAN 9 VARIABLES IN THE MODEL STATEMENT, SAS WILL STOP PROCESSING;
%* ON THE DATA NEXT_2 STEP DECLARING THE BY VARIABLE (PARM) IS NOT IN THE CORRECT SORTED;
%* ORDER. THIS DOESNT HAPPEN FOR LESS THAN NINE VARIABLES. WHEN YOU SORT THE DATA SET;
%* ON PARM, THE SORT DOES NOT TAKE PLACE AS EXPECTED, MESSING UP THE VARIANCE-COVARIANCE;
%* MATRIX. THE PROBLEM IS THAT THE VALUES OF PARM PROGRESS AS PARM1, PARM2, PARM3, ...;
%* PARM9, PARM10, ETC. WHEN YOU SORT ON PARM, PARM10, PARM11 THROUGH PARM19 SORT
AFTER;
%* PARM1 AND BEFORE PARM2, DUE TO THE WAY SORTING WORKS ON CHARACTER VARIABLES. THE
ONLY;
%* WAY TO FIX THIS IS TO RENAME THE VARIABLES TO PARM01, PARM02, ETC. SO THE SORTING WORKS;
%* CORRECTLY.;


DATA NEXT_1; SET &PARMINFO;
ATTRIB PARNUM FORMAT=$12.;
PARNUM=PARAMETER;
IF PARNUM = 'Prm1' THEN PARNUM = 'Prm01';
IF PARNUM = 'Prm2' THEN PARNUM = 'Prm02';
IF PARNUM = 'Prm3' THEN PARNUM = 'Prm03';
IF PARNUM = 'Prm4' THEN PARNUM = 'Prm04';
IF PARNUM = 'Prm5' THEN PARNUM = 'Prm05';
IF PARNUM = 'Prm6' THEN PARNUM = 'Prm06';
IF PARNUM = 'Prm7' THEN PARNUM = 'Prm07';
IF PARNUM = 'Prm8' THEN PARNUM = 'Prm08';
IF PARNUM = 'Prm9' THEN PARNUM = 'Prm09';

RENAME PARNUM=PARM;

RUN;
PROC SORT;
 BY PARM;
```

```
RUN;

DATA NEXT_1A; SET &COVDSN;
ATTRIB PARM FORMAT=$12.;
PARM=ROWNAME;
IF PARM = 'Prm1' THEN PARM = 'Prm01';
IF PARM = 'Prm2' THEN PARM = 'Prm02';
IF PARM = 'Prm3' THEN PARM = 'Prm03';
IF PARM = 'Prm4' THEN PARM = 'Prm04';
IF PARM = 'Prm5' THEN PARM = 'Prm05';
IF PARM = 'Prm6' THEN PARM = 'Prm06';
IF PARM = 'Prm7' THEN PARM = 'Prm07';
IF PARM = 'Prm8' THEN PARM = 'Prm08';
IF PARM = 'Prm9' THEN PARM = 'Prm09';

RUN;
PROC SORT;
 BY PARM;
RUN;

DATA NEXT_2(DROP=EFFECT); MERGE NEXT_1A(IN=IN1A) NEXT_1(IN=IN1); BY PARM; IF IN1A;
PARM=EFFECT;
RENAME PARM=_NAME_;
RUN;

  %* IN SOME OUTPUT VARIANCE-COVARIANCE MATRICES, THERE WILL BE A RECORD FOR;
  %* SCALE. DELETE THIS RECORD.;
  DATA NEXT_3; SET NEXT_2;
  IF _NAME_='SCALE' THEN DELETE;
  RUN;
  %* INSERT A DUMMY RECORD FOR ESTIMATE TO SIMULATE COVARIANCE OUTPUT FROM LOGISTIC
  %*  AND PHREG.;
  DATA NEXT_4;
  _NAME_= 'ESTIMATE';
  OUTPUT;
  RUN;
  DATA NEXT_5; SET NEXT_4 NEXT_3;

  RUN;
proc print; run;

 %END;

%* MAKE MIXED COVARIANCE OUTPUT SIMILAR ENOUGH TO LOGISTIC AND PHREG THAT THIS MACRO
WILL WORK.;
%* Use the /covb option after the model statement and inlcude the line of code:  ods output
CovB=dataset1;
%* Call macro as: %COLLIN(COVDSN=, PROCDR=MIXED, PARMINFO=dataset1);

%IF %upcase(&PROCDR)=MIXED %THEN %DO;
DATA NEXT_1 (Keep=_NAME_ col:); SET &COVDSN;
          RENAME EFFECT=_NAME_;
          if Row = 1          then  RowName = 'Prm01';
```

```
            if Row = 2          then  RowName = 'Prm02';
            if Row = 3          then  RowName = 'Prm03';
            if Row = 4          then  RowName = 'Prm04';
            if Row = 5          then  RowName = 'Prm05';
            if Row = 6          then  RowName = 'Prm06';
            if Row = 7          then  RowName = 'Prm07';
            if Row = 8          then  RowName = 'Prm08';
            if Row = 9          then  RowName = 'Prm09';
RUN;

data next_2(Drop=covars);
set next_1;

array cols col:;
covars=dim(cols);
call symput("Numcols",left(covars));
run;


data next_2a;
attrib dummy length=$1000;
retain dummy ;
set next_2;
if sum(of Col1-COL&Numcols)=0 then do;
if dummy=" " then dummy="Drop= COL"||trim(left(put(_N_,8.)));
else if dummy ne " " then do;

dummy=trim(left(dummy))||" COL"||trim(left(put(_N_,8.)));

end;
Call symput("Drop",dummy);
delete;
end;
run;



DATA NEXT_3; SET NEXT_2a(&drop);
  IF _NAME_='SCALE' THEN DELETE;
  RUN;

  %* INSERT A DUMMY RECORD FOR ESTIMATE TO SIMULATE COVARIANCE OUTPUT FROM LOGISTIC
  %*  AND PHREG.;
  DATA NEXT_4;                    ATTRIB _NAME_ FORMAT=$12.;
  _NAME_= 'ESTIMATE';
  OUTPUT;
  RUN;

  DATA NEXT_5; SET NEXT_4 NEXT_3; ATTRIB _NAME_ FORMAT=$12.;
  RUN;

proc print; run;
```

```
 %END;


%* MAKE GLIMMIX COVARIANCE OUTPUT SIMILAR ENOUGH TO LOGISTIC AND PHREG THAT THIS MACRO
WILL WORK.;

%IF %Upcase(&PROCDR)=GLIMMIX %THEN %DO;
DATA NEXT_1 (Keep=_NAME_ COL:); SET &COVDSN;
        RENAME EFFECT=_NAME_;
        if Row = 1      then  RowName = 'Prm01';
        if Row = 2      then  RowName = 'Prm02';
        if Row = 3      then  RowName = 'Prm03';
        if Row = 4      then  RowName = 'Prm04';
        if Row = 5      then  RowName = 'Prm05';
        if Row = 6      then  RowName = 'Prm06';
        if Row = 7      then  RowName = 'Prm07';
        if Row = 8      then  RowName = 'Prm08';
        if Row = 9      then  RowName = 'Prm09';
RUN;

data next_2(Drop=covars);
set next_1;

array cols col:;
covars=dim(cols);
call symput("Numcols",left(covars));
run;


data next_2a;
attrib dummy length=$1000;
retain dummy ;
set next_2;
if sum(of Col1-COL&Numcols)=0 then do;
if dummy=" " then dummy="Drop= COL"||trim(left(put(_N_,8.)));
else if dummy ne " " then do;

dummy=trim(left(dummy))||" COL"||trim(left(put(_N_,8.)));
end;
Call symput("Drop",dummy);
delete;
end;
run;


*Depending on the reference coding used in GENMOD and MIXED the Covariance MATRIX output by the
procedure may have
        Columns and corresponding Rows with all Zeros.
The MACRO Variable DROP (created in NEXT_2A) Isolates and removes these extraneous columns before
we get to
Manipulating the matrix in the IML code below;

DATA NEXT_3; SET NEXT_2a(&drop);
```

```
  IF _NAME_='SCALE' THEN DELETE;
  RUN;

  %* INSERT A DUMMY RECORD FOR ESTIMATE TO SIMULATE COVARIANCE OUTPUT FROM LOGISTIC
  %*  AND PHREG.;
  DATA NEXT_4;                          ATTRIB _NAME_ FORMAT=$12.;
  _NAME_= 'ESTIMATE';
  OUTPUT;
  RUN;

  DATA NEXT_5; SET NEXT_4 NEXT_3; ATTRIB _NAME_ FORMAT=$12.;
  RUN;

proc print; run;

%END;

%* MAKE SURVEYLOGISTIC COVARIANCE OUTPUT SIMILAR ENOUGH TO LOGISTIC AND PHREG THAT THIS
MACRO WILL WORK.;

%IF %Upcase(&PROCDR)=SURVEYLOGISTIC %THEN %DO;
DATA NEXT_1 ; SET &covdsn;
          RENAME Parameter=_NAME_;

RUN;


DATA NEXT_3; SET NEXT_1;
  IF _NAME_='SCALE' THEN DELETE;
  RUN;

  %* INSERT A DUMMY RECORD FOR ESTIMATE TO SIMULATE COVARIANCE OUTPUT FROM LOGISTIC
  %*  AND PHREG.;
  DATA NEXT_4;   ATTRIB _NAME_ FORMAT=$32.;
  _NAME_= 'ESTIMATE';
  OUTPUT;
  RUN;

  DATA NEXT_5; SET NEXT_4 NEXT_3; ATTRIB _NAME_ FORMAT=$32.;
  RUN;

proc print; run;

%END;


%IF &PROCDR=%str()
          or %upcase(&PROCDR)=LOGISTIC
          or %upcase(&PROCDR)=PHREG
%THEN %DO;
  DATA NEXT_5; SET &COVDSN;
  RUN;
```

```
%END;


proc print data=next_5; run;

%IF (NEXT_5 NE ) %THEN %DO;

OPTION MPRINT;

%LET __STOP=0;

PROC IML;
 USE NEXT_5;
 READ ALL VAR {_NAME_} INTO _VARNAME;

 _NRVNAME=NROW(_VARNAME);


 IF (_NRVNAME>1) THEN DO;
   _VARNAM2=_VARNAME(|2:_NRVNAME, |);
   NMISSING=J(NROW(_VARNAM2),1,.);
   LABELS={"EIGENVAL","CONDINDX","        "};
   _VARNAM2=LABELS//_VARNAM2;
   FREE _VARNAME LABELS;
   READ ALL VAR _NUM_ INTO VARCOV(|COLNAME=_NVNAME|);
   _NRCVC=NCOL(VARCOV);
   LASTVNAM=_NVNAME(|1,_NRCVC|);
   IF (LASTVNAM="_LNLIKE_") THEN VARCOV2=VARCOV(|2:_NRVNAME,1:_NRCVC-1|);
   IF (LASTVNAM^="_LNLIKE_") THEN VARCOV2=VARCOV(|2:_NRVNAME,|);

%* IF COVARIANCE MATRIX IS FROM PROC GENMOD USING THE REPEATED MEASURES DESIGN;
%* THEN THE LOWER DIAGONAL WILL HAVE THE CORRELATIONS AND THE UPPER DIAGONAL WILL HAVE;
%* THE COVARIANCES. THIS NEXT SECTION OF CODE REPLACES THE LOWER DIAGONAL WITH THE UPPER;
%* DIAGONAL TO MAKE A SYMMETRIC COVARIANCE MATRIX. IF THE MATRIX IS SYMMETRICAL ALREADY;
%* THEN THE NEXT SECTION OF CODE WILL NOT AFFECT ANYTHING.;


     VC2_C = NCOL(VARCOV2);
     VC2_R = NROW(VARCOV2);
     DO CL=1 TO VC2_C;
       DO RW=1 TO VC2_R;
         VARCOV2(|RW,CL|) = VARCOV2(|CL,RW|);
       END;
     END;

%* PRINT THE VARIANCE-COVARIANCE MATRIX FOR DIAGNOSTIC PURPOSES;
    PRINT VARCOV2;

   FREE VARCOV _NRCVC LASTVNAM VC2_C VC2_R CL;
   COVBINV=INV(VARCOV2);
   SCALE=INV(SQRT(DIAG(COVBINV)));
   R=SCALE*COVBINV*SCALE;
```

```
    FREE COVBINV SCALE;
    CALL EIGEN(MUSQR,V,R);
    FREE R;
    SROOTMUS=SQRT(MUSQR);
    CI=1/(SROOTMUS/MAX(SROOTMUS));
    PHI=(V##2)*DIAG(MUSQR##(-1));
    SUMPHI=PHI(|,+|);
    PI=PHI#(SUMPHI##(-1));
    FREE PHI SUMPHI SROOTMUS V;
    FINAL=(MUSQR||CI||NMISSING||PI`)`;
    FREE PI MUSQR CI NMISSING;
    _NCFINAL=NCOL(FINAL);
    _NRFINAL=NROW(FINAL);
    FINAL2=J(_NRFINAL,_NCFINAL,0);
    _NCFP1=_NCFINAL+1;
    __VDP="VDP";
    DO I=1 TO _NCFINAL;
      FINAL2(|,_NCFP1-I|)=FINAL(|,I|);
      X=CHAR(I,3);
      Y=COMPRESS(CONCAT(__VDP,X));
      IF I=1 THEN _VDPNAME=Y;
        ELSE _VDPNAME=_VDPNAME||Y;
    END;
    FREE FINAL _NRFINAL _NCFINAL I X Y;
    CREATE &output FROM FINAL2(|ROWNAME=_VARNAM2 COLNAME=_VDPNAME|);
    APPEND FROM FINAL2(|ROWNAME=_VARNAM2|);
    FREE _VARNAM2 _VDPNAME FINAL2;
  END;
  IF (_NRVNAME=1) THEN DO;
    X="1";
    CALL SYMPUT("__STOP",LEFT(X));
    PRINT " ";
    PRINT "*******************************************************";
    PRINT "YOU NEED TO SPECIFY THE  COVOUT  OPTION";
    PRINT " IN EITHER PROC LOGISTIC OR PROC PHREG.";
    PRINT " THIS PROGRAM WILL NOT CALCULATE COLLINEARITY DIAGNOSTICS.";
    PRINT "*******************************************************";
    PRINT " ";
  END;
  QUIT;
RUN;

%IF (&__STOP EQ 0) %THEN %DO;
   PROC PRINT DATA=&output LABEL NOOBS;
    ID _VARNAM2;
            Title7 "Input DATASET &COVDSN, Submitted &sysdate9";
    TITLE8 "COLLINEARITY DIAGNOSTICS FOR NONLINEAR MODELS USING";
    TITLE9 "THE INFORMATION MATRIX:  EIGENVALUES, CONDITION INDEXES,";
    TITLE10 "AND VARIANCE DECOMPOSITION PROPORTIONS (VDP'S)";

    LABEL _VARNAM2="VARIABLE";
   RUN;
%END;
```

```
%END;
%ELSE %DO;
  %PUT;
  %PUT "********************************************************";
  %PUT "WHEN YOU INVOKE THIS MACRO, YOU HAVE TO SPECIFY THE NAME";
  %PUT " OF A SAS DATA SET THAT CONTAINS THE VARIANCE-COVARIANCE";
  %PUT " MATRIX FROM EITHER PROC LOGISTIC OR PROC PHREG.";
  %PUT;
  %PUT "YOU CAN CREATE THIS MATRIX BY INCLUDING THE FOLLOWING OPTIONS";
  %PUT " ON THE PROC STATEMENT:  COVOUT  AND  OUTEST=SASDSN,";
  %PUT " WHERE SASDSN IS THE NAME OF THE SAS DATA SET CONTAINING";
  %PUT " THE VARIANCE-COVARIANCE MATRIX.";
  %PUT "********************************************************";
  %PUT;
%END;

PROC DATASETS;
DELETE NEXT_1 NEXT_1A NEXT_2 Next_2a NEXT_3 NEXT_4 NEXT_5;
RUN;
QUIT;

title;

%MEND COLLIN;
```

Appendix C: SAS Code

```
options obs= 13917;
                                               ***/CHANGE/***;
libname dataset 'H:\0-NATIONAL\National2005\yrbs2005-formatted.xls';
proc print data=dataset.'YRBS_2005 formatted$'n (obs=2); run;




*******************************************************************
                    *          DATASET PREPARATION  - YRBS 2005, 2007, 2009, 2011 *
*******************************************************************/;
options obs= 59793;
libname dataset 'H:\0-NATIONAL\National2005_2011\yrbs2005_2011-formatted.xls';

/* Check files for Analysis */;
proc print data=dataset.'YRBS_2005_2011 formatted$'n (obs=2); run;

*/ Create the needed dataset /*;
DATA ST_YEAR; set dataset.'YRBS_2005_2011 formatted$'n; run; */ST_YEAR dataset /*;
proc print data=ST_YEAR (obs=2); run;

/* Recode 1-2 and 0-1 for future analysis, var coded 0 or 2 is the referent */;
data ST_YEAR_01;
set ST_YEAR;
if nTST=1 then do; nTST01=1; nTST21=1; end; /*event: tested*/
if nTST=2 then do; nTST01=0; nTST21=2; end; /*referent: not tested*/
if nTST not in (1 2) then do; nTST01=.; nTST21=.; end;

if nTCH=1 then do; nTCH01=1; nTCH21=1; end; /*event: taught*/
if nTCH=2 then do; nTCH01=0; nTCH21=2; end; /*referent: not taught*/
if nTCH not in (1 2) then do; nTCH01=.; nTCH21=.; end;

if SEX=1 then do; SEX01=1; SEX21=1; end; /*event: female*/
if SEX=2 then do; SEX01=0; SEX21=2; end; /*referent: male*/
if SEX not in (1 2) then do; SEX01=.; SEX21=.; end;

if GRD=1 then do; grade=4; end; /*referent: 9th grade*/
if GRD=2 then do; grade=3; end;
if GRD=3 then do; grade=2; end;
if GRD=4 then do; grade=1; end;
if GRD not in (1 2 3 4) then do; grade=.; end;

if HS=1 then do; HS01=1; HS21=1; end; /*event: Had sex*/
if HS=2 then do; HS01=0; HS21=2; end; /*referent: not sexually active*/
if HS not in (1 2) then do; HS01=.; HS21=.; end;

if nnRACETH=1 then do; race_eth=4; end; /*referent: white*/
if nnRACETH=2 then do; race_eth=3; end;
if nnRACETH=3 then do; race_eth=2; end;
if nnRACETH=4 then do; race_eth=1; end;
```

```
if nnRACETH not in (1 2 3 4) then do; race_eth=.; end;

if IF=1 then do; IF01=1; IF21=1; end; /*event: had a forced intercourse*/
if IF=2 then do; IF01=0; IF21=2; end; /*referent: intercourse not forced*/
if IF not in (1 2) then do; IF01=.; IF21=.; end;

if nAAI=1 then do; nAAI01=1; nAAI21=1; end; /*event: had intercourse before 13*/
if nAAI=2 then do; nAAI01=0; nAAI21=2; end; /*referent: no intercourse before 13*/
if nAAI not in (1 2) then do; nAAI01=.; nAAI21=.; end;

if nNPL=1 then do; nNPL01=1; nNPL21=1; end; /*event: had 4+ partner in life*/
if nNPL=2 then do; nNPL01=0; nNPL21=2; end; /*referent: less than 4 partners in life*/
if nNPL not in (1 2) then do; nNPL01=.; nNPL21=.; end;

if nUZC=2 then do; nUZC01=0; nUZC21=2; end; /*referent: use of condom*/
if nUZC=1 then do; nUZC01=1; nUZC21=1; end; /*event: did not use a condom*/
if nUZC not in (1 2) then do; nUZC01=.; nUZC21=.; end;

if nIUD=1 then do; nIUD01=1; nIUD21=1; end; /*event: used injected drug*/
if nIUD=2 then do; nIUD01=0; nIUD21=2; end; /*referent: no use of injected drug in life*/
if nIUD not in (1 2) then do; nIUD01=.; nIUD21=.; end;

run;

proc print data=ST_YEAR_01 (obs=3); run;

/*Verify data */
proc contents data=ST_YEAR_01 varnum; run;
proc freq data=ST_YEAR_01; Tables YR nYEAR; run;
proc freq data=ST_YEAR_01; Tables STR; run;
proc freq data=ST_YEAR_01; Tables PSU; run;
proc freq data=ST_YEAR_01; Tables WT; run;
proc freq data=ST_YEAR_01; Tables nTST nTST01 nTST21; run;
proc freq data=ST_YEAR_01; Tables nTCH nTCH01 nTCH21; run;
proc freq data=ST_YEAR_01; Tables SEX SEX01 SEX21; run;
proc freq data=ST_YEAR_01; Tables GRD grade; run;
proc freq data=ST_YEAR_01; Tables nnRACETH race_eth; run;
proc freq data=ST_YEAR_01; Tables HS HS01 HS21; run;
proc freq data=ST_YEAR_01; Tables IF IF01 IF21; run;
proc freq data=ST_YEAR_01; Tables nAAI nAAI01 nAAI21; run;
proc freq data=ST_YEAR_01; Tables nNPL nNPL01 nNPL21; run;
proc freq data=ST_YEAR_01; Tables nUZC nUZC01 nUZC21; run;
proc freq data=ST_YEAR_01; Tables nIUD nIUD01 nIUD21; run;

/* store in EXCEL file */;
proc export data= ST_YEAR_01
outfile = 'H:\0-NATIONAL\National2005_2011\dataset_recod.XLS'
dbms = excel replace;
sheet = "STYEARrecod";
run;
```

```
************************************************
/*           DATA OVERVIEW - NATIONAL YRBS                    */;
************************************************;

/* OVERALL - 2005_2011  */;
options obs= 59793;
libname RecodDat 'H:\0-NATIONAL\National2005_2011\dataset_recod.XLS';

proc surveyfreq data=RecodDat.'STYEARrecod$'n nomcar;
stratum STR;
cluster PSU;
weight WT;
Tables SEX21 / row(deff) CL(type=logit) ;
Tables grade / row(deff) CL(type=logit) ;
Tables race_eth  / row(deff) CL(type=logit) ;
Tables nTST21 / row(deff) CL(type=logit) ;
Tables nTCH21 / row(deff) CL(type=logit) ;
Tables nUZC21 / row(deff) CL(type=logit) ;
Tables nNPL21 / row(deff) CL(type=logit) ;
Tables nAAI21/ row(deff) CL(type=logit) ;
Tables IF21 / row(deff) CL(type=logit) ;
Tables nIUD21 / row(deff) CL(type=logit) ;
run;

/* POPULATION OF INTEREST (HAD SEX) - 2005_2011  */;
options obs= 59793;
libname RecodDat 'H:\0-NATIONAL\National2005_2011\dataset_recod.XLS';

proc surveyfreq data=RecodDat.'STYEARrecod$'n nomcar;
stratum STR;
cluster PSU;
weight WT;
Tables HS21 / row(deff) CL(type=logit) ;
run;

proc surveyfreq data=RecodDat.'STYEARrecod$'n nomcar;
stratum STR;
cluster PSU;
weight WT;
Tables HS21*SEX21 / row(deff) CL(type=logit) ;
Tables HS21*grade / row(deff) CL(type=logit) ;
Tables HS21*race_eth  / row(deff) CL(type=logit) ;
Tables HS21*nTST21 / row(deff) CL(type=logit) ;
Tables HS21*nTCH21 / row(deff) CL(type=logit) ;
Tables HS21*nUZC21 / row(deff) CL(type=logit) ;
Tables HS21*nNPL21 / row(deff) CL(type=logit) ;
Tables HS21*nAAI21/ row(deff) CL(type=logit) ;
Tables HS21*IF21 / row(deff) CL(type=logit) ;
Tables HS21*nIUD21 / row(deff) CL(type=logit) ;
run;

/* POPULATION OF INTEREST (HAD SEX) - 2005  */;
```

```
options obs= 13917;

libname RecodDat 'H:\0-NATIONAL\National2005\dataset_recod.XLS';

proc surveyfreq data=RecodDat.'STYEARrecod$'n nomcar;
stratum STR;
cluster PSU;
weight WT;
Tables HS21 / row(deff) CL(type=logit) ;
run;

options obs= 13917;

libname RecodDat 'H:\0-NATIONAL\National2005\dataset_recod.XLS';

proc surveyfreq data=RecodDat.'STYEARrecod$'n nomcar;
stratum STR;
cluster PSU;
weight WT;
Tables HS21*SEX21 / row(deff) CL(type=logit) ;
Tables HS21*grade / row(deff) CL(type=logit) ;
Tables HS21*race_eth  / row(deff) CL(type=logit) ;
Tables HS21*nTST21 / row(deff) CL(type=logit) ;
Tables HS21*nTCH21 / row(deff) CL(type=logit) ;
Tables HS21*nUZC21 / row(deff) CL(type=logit) ;
Tables HS21*nNPL21 / row(deff) CL(type=logit) ;
Tables HS21*nAAI21/ row(deff) CL(type=logit) ;
Tables HS21*IF21 / row(deff) CL(type=logit) ;
Tables HS21*nIUD21 / row(deff) CL(type=logit) ;
run;

/* POPULATION OF INTEREST (HAD SEX) - 2007  */;
options obs= 14041;

libname RecodDat 'H:\0-NATIONAL\National2007\dataset_recod.XLS';

proc surveyfreq data=RecodDat.'STYEARrecod$'n nomcar;
stratum STR;
cluster PSU;
weight WT;
Tables HS21;
run;

options obs= 14041;

libname RecodDat 'H:\0-NATIONAL\National2007\dataset_recod.XLS';

proc surveyfreq data=RecodDat.'STYEARrecod$'n nomcar;
stratum STR;
cluster PSU;
weight WT;
Tables HS21*SEX21 / row(deff) CL(type=logit) ;
Tables HS21*grade / row(deff) CL(type=logit) ;
```

```
Tables HS21*race_eth  / row(deff) CL(type=logit) ;
Tables HS21*nTST21 / row(deff) CL(type=logit) ;
Tables HS21*nTCH21 / row(deff) CL(type=logit) ;
Tables HS21*nUZC21 / row(deff) CL(type=logit) ;
Tables HS21*nNPL21 / row(deff) CL(type=logit) ;
Tables HS21*nAAI21/ row(deff) CL(type=logit) ;
Tables HS21*IF21 / row(deff) CL(type=logit) ;
Tables HS21*nIUD21 / row(deff) CL(type=logit) ;
run;

/* POPULATION OF INTEREST (HAD SEX) - 2009  */;
options obs= 16410;

libname RecodDat 'H:\0-NATIONAL\National2009\dataset_recod.XLS';

proc surveyfreq data=RecodDat.'STYEARrecod$'n nomcar;
stratum STR;
cluster PSU;
weight WT;
Tables HS21 ;
run;

options obs= 16410;

libname RecodDat 'H:\0-NATIONAL\National2009\dataset_recod.XLS';

proc surveyfreq data=RecodDat.'STYEARrecod$'n nomcar;
stratum STR;
cluster PSU;
weight WT;
Tables HS21*SEX21 / row(deff) CL(type=logit) ;
Tables HS21*grade / row(deff) CL(type=logit) ;
Tables HS21*race_eth  / row(deff) CL(type=logit) ;
Tables HS21*nTST21 / row(deff) CL(type=logit) ;
Tables HS21*nTCH21 / row(deff) CL(type=logit) ;
Tables HS21*nUZC21 / row(deff) CL(type=logit) ;
Tables HS21*nNPL21 / row(deff) CL(type=logit) ;
Tables HS21*nAAI21/ row(deff) CL(type=logit) ;
Tables HS21*IF21 / row(deff) CL(type=logit) ;
Tables HS21*nIUD21 / row(deff) CL(type=logit) ;
run;

/* POPULATION OF INTEREST (HAD SEX) - 2011  */;
options obs= 15425;
libname RecodDat 'H:\0-NATIONAL\National2011\dataset_recod.XLS';

proc surveyfreq data=RecodDat.'STYEARrecod$'n nomcar;
stratum STR;
cluster PSU;
weight WT;
Tables HS21;
run;
```

```
options obs= 15425;
libname RecodDat 'H:\0-NATIONAL\National2011\dataset_recod.XLS';

proc surveyfreq data=RecodDat.'STYEARrecod$'n nomcar;
stratum STR;
cluster PSU;
weight WT;
Tables HS21*SEX21 / row(deff) CL(type=logit) ;
Tables HS21*grade / row(deff) CL(type=logit) ;
Tables HS21*race_eth  / row(deff) CL(type=logit) ;
Tables HS21*nTST21 / row(deff) CL(type=logit) ;
Tables HS21*nTCH21 / row(deff) CL(type=logit) ;
Tables HS21*nUZC21 / row(deff) CL(type=logit) ;
Tables HS21*nNPL21 / row(deff) CL(type=logit) ;
Tables HS21*nAAI21/ row(deff) CL(type=logit) ;
Tables HS21*IF21 / row(deff) CL(type=logit) ;
Tables HS21*nIUD21 / row(deff) CL(type=logit) ;
run;



*****************************************************************
/*        NATIONAL TRENDS AMONG HIGH SCHOOL STUDENTS WHO HAD INTERCOURSE - 2005-2011   */
*****************************************************************;

/*National 2005-2011 */
options obs= 59793;
libname RecodDat 'H:\0-NATIONAL\National2005_2011';

/*Transfert from EXCEL to temporary dataset */
proc import datafile='H:\0-NATIONAL\National2005_2011\dataset_recod.XLS'
dbms = EXCEL
replace
out = TRENDWK;
sheet = "STYEARrecod";
run;
proc print data=TRENDWK (obs=10); run;

/*Sort for future use */
proc sort data=TRENDWK;
by YR STR PSU;
run;
proc print data=TRENDWK (obs=10); run;

/*Put temporary dataset into SAS format */
data RecodDat.TRENDWK;
set TRENDWK;
run;
proc print data=RecodDat.TRENDWK (obs=3); run;

/*Calculate orthogoanal coefficients: linear and quadratic */
PROC IML; X={2005 2007 2009 2011}; XP=ORPOL(X,3); PRINT XP; QUIT;
```

```
/*Plug orthogonal coefficients */
DATA VARSET; SET RecodDat.TRENDWK;
   IF (YR=2005 OR YR=2007 OR YR=2009 OR YR=2011);
   IF YR=2005 THEN DO; T4L=-0.67082; T4Q=0.5; end;
ELSE IF YR=2007 THEN DO; T4L=-0.223607; T4Q=-0.5; end;
ELSE IF YR=2009 THEN DO; T4L=0.2236068; T4Q=-0.5; end;
ELSE IF YR=2011 THEN DO; T4L=0.6708204; T4Q=0.5; end;

/*Test if change over time: for outcome and risk behaviors */;

/* nTST01 */;
PROC RLOGIST DATA=VARSET DESIGN=WR FILETYPE=SAS;
NEST YR STR PSU/PSULEV=3 MISSUNIT;
WEIGHT WT;
SUBPOPN HS21=1;
CLASS SEX01 RACE_ETH GRADE;
REFLEVEL SEX01=0 RACE_ETH=4 GRADE=4;
MODEL nTST01 = SEX01 RACE_ETH GRADE T4L ;
test waldf satadjf satadjchi;
PRINT/BETAFMT=F8.5 SEBETAFMT=F8.5 P_BETAFMT=F8.5;
RUN;

PROC RLOGIST DATA=VARSET DESIGN=WR FILETYPE=SAS;
NEST YR STR PSU/PSULEV=3 MISSUNIT;
WEIGHT WT;
SUBPOPN HS21=1;
CLASS SEX01 RACE_ETH GRADE;
REFLEVEL SEX01=0 RACE_ETH=4 GRADE=4;
MODEL nTST01 = SEX01 RACE_ETH GRADE T4L T4Q;
test waldf satadjf satadjchi;
PRINT/BETAFMT=F8.5 SEBETAFMT=F8.5 P_BETAFMT=F8.5;
RUN;

/* nTCH01 */;
PROC RLOGIST DATA=VARSET DESIGN=WR FILETYPE=SAS;
NEST YR STR PSU/PSULEV=3 MISSUNIT;
WEIGHT WT;
SUBPOPN HS21=1;
CLASS SEX01 RACE_ETH GRADE;
REFLEVEL SEX01=0 RACE_ETH=4 GRADE=4;
MODEL nTCH01 = SEX01 RACE_ETH GRADE T4L ;
test waldf satadjf satadjchi;
PRINT/BETAFMT=F8.5 SEBETAFMT=F8.5 P_BETAFMT=F8.5;
RUN;

PROC RLOGIST DATA=VARSET DESIGN=WR FILETYPE=SAS;
NEST YR STR PSU/PSULEV=3 MISSUNIT;
WEIGHT WT;
SUBPOPN HS21=1;
CLASS SEX01 RACE_ETH GRADE;
REFLEVEL SEX01=0 RACE_ETH=4 GRADE=4;
MODEL nTCH01 = SEX01 RACE_ETH GRADE T4L T4Q ;
```

```
test waldf satadjf satadjchi;
PRINT/BETAFMT=F8.5 SEBETAFMT=F8.5 P_BETAFMT=F8.5;
RUN;

/* nUZC01 */;
PROC RLOGIST DATA=VARSET DESIGN=WR FILETYPE=SAS;
NEST YR STR PSU/PSULEV=3 MISSUNIT;
WEIGHT WT;
SUBPOPN HS21=1;
CLASS SEX01 RACE_ETH GRADE;
REFLEVEL SEX01=0 RACE_ETH=4 GRADE=4;
MODEL nUZC01 = SEX01 RACE_ETH GRADE T4L ;
test waldf satadjf satadjchi;
PRINT/BETAFMT=F8.5 SEBETAFMT=F8.5 P_BETAFMT=F8.5;
RUN;

PROC RLOGIST DATA=VARSET DESIGN=WR FILETYPE=SAS;
NEST YR STR PSU/PSULEV=3 MISSUNIT;
WEIGHT WT;
SUBPOPN HS21=1;
CLASS SEX01 RACE_ETH GRADE;
REFLEVEL SEX01=0 RACE_ETH=4 GRADE=4;
MODEL nUZC01 = SEX01 RACE_ETH GRADE T4L T4Q;
test waldf satadjf satadjchi;
PRINT/BETAFMT=F8.5 SEBETAFMT=F8.5 P_BETAFMT=F8.5;
RUN;

/* nNPL01 */;
PROC RLOGIST DATA=VARSET DESIGN=WR FILETYPE=SAS;
NEST YR STR PSU/PSULEV=3 MISSUNIT;
WEIGHT WT;
SUBPOPN HS21=1;
CLASS SEX01 RACE_ETH GRADE;
REFLEVEL SEX01=0 RACE_ETH=4 GRADE=4;
MODEL nNPL01 = SEX01 RACE_ETH GRADE T4L ;
test waldf satadjf satadjchi;
PRINT/BETAFMT=F8.5 SEBETAFMT=F8.5 P_BETAFMT=F8.5;
RUN;

PROC RLOGIST DATA=VARSET DESIGN=WR FILETYPE=SAS;
NEST YR STR PSU/PSULEV=3 MISSUNIT;
WEIGHT WT;
SUBPOPN HS21=1;
CLASS SEX01 RACE_ETH GRADE;
REFLEVEL SEX01=0 RACE_ETH=4 GRADE=4;
MODEL nNPL01 = SEX01 RACE_ETH GRADE T4L T4Q;
test waldf satadjf satadjchi;
PRINT/BETAFMT=F8.5 SEBETAFMT=F8.5 P_BETAFMT=F8.5;
RUN;

/* nAAI01 */;
PROC RLOGIST DATA=VARSET DESIGN=WR FILETYPE=SAS;
NEST YR STR PSU/PSULEV=3 MISSUNIT;
```

```
WEIGHT WT;
SUBPOPN HS21=1;
CLASS SEX01 RACE_ETH GRADE;
REFLEVEL SEX01=0 RACE_ETH=4 GRADE=4;
MODEL nAAI01 = SEX01 RACE_ETH GRADE T4L ;
test waldf satadjf satadjchi;
PRINT/BETAFMT=F8.5 SEBETAFMT=F8.5 P_BETAFMT=F8.5;
RUN;

PROC RLOGIST DATA=VARSET DESIGN=WR FILETYPE=SAS;
NEST YR STR PSU/PSULEV=3 MISSUNIT;
WEIGHT WT;
SUBPOPN HS21=1;
CLASS SEX01 RACE_ETH GRADE;
REFLEVEL SEX01=0 RACE_ETH=4 GRADE=4;
MODEL nAAI01 = SEX01 RACE_ETH GRADE T4L T4Q;
test waldf satadjf satadjchi;
PRINT/BETAFMT=F8.5 SEBETAFMT=F8.5 P_BETAFMT=F8.5;
RUN;

/* IF01 */;
PROC RLOGIST DATA=VARSET DESIGN=WR FILETYPE=SAS;
NEST YR STR PSU/PSULEV=3 MISSUNIT;
WEIGHT WT;
SUBPOPN HS21=1;
CLASS SEX01 RACE_ETH GRADE;
REFLEVEL SEX01=0 RACE_ETH=4 GRADE=4;
MODEL IF01 = SEX01 RACE_ETH GRADE T4L ;
test waldf satadjf satadjchi;
PRINT/BETAFMT=F8.5 SEBETAFMT=F8.5 P_BETAFMT=F8.5;
RUN;

PROC RLOGIST DATA=VARSET DESIGN=WR FILETYPE=SAS;
NEST YR STR PSU/PSULEV=3 MISSUNIT;
WEIGHT WT;
SUBPOPN HS21=1;
CLASS SEX01 RACE_ETH GRADE;
REFLEVEL SEX01=0 RACE_ETH=4 GRADE=4;
MODEL IF01 = SEX01 RACE_ETH GRADE T4L T4Q;
test waldf satadjf satadjchi;
PRINT/BETAFMT=F8.5 SEBETAFMT=F8.5 P_BETAFMT=F8.5;
RUN;

/* nIUD01 */;
PROC RLOGIST DATA=VARSET DESIGN=WR FILETYPE=SAS;
NEST YR STR PSU/PSULEV=3 MISSUNIT;
WEIGHT WT;
SUBPOPN HS21=1;
CLASS SEX01 RACE_ETH GRADE;
REFLEVEL SEX01=0 RACE_ETH=4 GRADE=4;
MODEL   nIUD01 = SEX01 RACE_ETH GRADE T4L ;
test waldf satadjf satadjchi;
PRINT/BETAFMT=F8.5 SEBETAFMT=F8.5 P_BETAFMT=F8.5;
```

```
RUN;

PROC RLOGIST DATA=VARSET DESIGN=WR FILETYPE=SAS;
NEST YR STR PSU/PSULEV=3 MISSUNIT;
WEIGHT WT;
SUBPOPN HS21=1;
CLASS SEX01 RACE_ETH GRADE;
REFLEVEL SEX01=0 RACE_ETH=4 GRADE=4;
MODEL   nIUD01 = SEX01 RACE_ETH GRADE T4L T4Q;
test waldf satadjf satadjchi;
PRINT/BETAFMT=F8.5 SEBETAFMT=F8.5 P_BETAFMT=F8.5;
RUN;


*******************************************************
/*                    CRUDE ODD RATIOS - NATIONAL LEVEL - 2005-2011   */;
*******************************************************;
/*National 2005-2011 */
options obs= 59793;
libname RecodDat 'H:\0-NATIONAL\National2005_2011\dataset_recod.XLS';

proc surveylogistic DATA=RecodDat.'STYEARrecod$'n ;
domain HS21;
stratum STR;
cluster PSU;
weight WT;
model nTST01 (EVENT='1')= nNPL01;
run;

proc surveylogistic DATA=RecodDat.'STYEARrecod$'n ;
domain HS21;
stratum STR;
cluster PSU;
weight WT;
model nTST01 (EVENT='1')= IF01;
run;

proc surveylogistic DATA=RecodDat.'STYEARrecod$'n ;
domain HS21;
stratum STR;
cluster PSU;
weight WT;
model nTST01 (EVENT='1')= nAAI01;
run;

proc surveylogistic DATA=RecodDat.'STYEARrecod$'n ;
domain HS21;
stratum STR;
cluster PSU;
weight WT;
model nTST01 (EVENT='1')= nUZC01;
run;
```

```
proc surveylogistic DATA=RecodDat.'STYEARrecod$'n ;
domain HS21;
stratum STR;
cluster PSU;
weight WT;
model nTST01 (EVENT='1')= nIUD01;
run;

proc surveylogistic DATA=RecodDat.'STYEARrecod$'n ;
domain HS21;
stratum STR;
cluster PSU;
weight WT;
model nTST01 (EVENT='1')= SEX01;
run;

proc surveylogistic DATA=RecodDat.'STYEARrecod$'n ;
domain HS21;
stratum STR;
cluster PSU;
weight WT;
class grade (REF='4')/param=ref;
model nTST01 (EVENT='1')= grade;
run;

proc surveylogistic DATA=RecodDat.'STYEARrecod$'n ;
domain HS21;
stratum STR;
cluster PSU;
weight WT;
class race_eth (REF='4')/param=ref;
model nTST01 (EVENT='1')= race_eth;
run;


*************************************************************************
/* CRUDE ODD RATIOS BY STRATA YEAR - NATIONAL LEVEL - 2005, 2007, 2009, 2011 */;
*************************************************************************;
proc surveylogistic DATA=RecodDat.'STYEARrecod$'n ;
domain HS21;
stratum STR;
cluster PSU;
weight WT;
by nyear;
model nTST01 (EVENT='1')= nNPL01;
run;

proc surveylogistic DATA=RecodDat.'STYEARrecod$'n ;
domain HS21;
stratum STR;
cluster PSU;
weight WT;
```

```
by nyear;
model nTST01 (EVENT='1')= IF01;
run;

proc surveylogistic DATA=RecodDat.'STYEARrecod$'n ;
domain HS21;
stratum STR;
cluster PSU;
weight WT;
by nyear;
model nTST01 (EVENT='1')= nAAI01;
run;

proc surveylogistic DATA=RecodDat.'STYEARrecod$'n ;
domain HS21;
stratum STR;
cluster PSU;
weight WT;
by nyear;
model nTST01 (EVENT='1')= nUZC01;
run;

proc surveylogistic DATA=RecodDat.'STYEARrecod$'n ;
domain HS21;
stratum STR;
cluster PSU;
weight WT;
by nyear;
model nTST01 (EVENT='1')= nIUD01;
run;

proc surveylogistic DATA=RecodDat.'STYEARrecod$'n ;
domain HS21;
stratum STR;
cluster PSU;
weight WT;
by nyear;
model nTST01 (EVENT='1')= SEX01;
run;

proc surveylogistic DATA=RecodDat.'STYEARrecod$'n ;
domain HS21;
stratum STR;
cluster PSU;
weight WT;
by nyear;
class grade (REF='4')/param=ref;
model nTST01 (EVENT='1')= grade;
run;

proc surveylogistic DATA=RecodDat.'STYEARrecod$'n ;
domain HS21;
stratum STR;
```

```
cluster PSU;
weight WT;
by nyear;
class race_eth (REF='4')/param=ref;
model nTST01 (EVENT='1')= race_eth;
run;


*************************************************************
/*                    ADJUSTED ODD RATIOS - NATIONAL LEVEL - 2005-2011          */;
*************************************************************;
/*National 2005-2011 */
options obs= 59793;
libname RecodDat 'H:\0-NATIONAL\National2005_2011\dataset_recod.XLS';

/* For the pool - All Years Together */;
proc surveylogistic DATA=RecodDat.'STYEARrecod$'n ;
domain HS21;
stratum STR;
cluster PSU;
weight WT;
class grade (REF='4')/param=ref;
class race_eth (REF='4')/param=ref;
class nYEAR (REF='4')/param=ref;
model nTST01 (EVENT='1')= nNPL01 IF01 nAAI01 nUZC01 nIUD01 SEX01 grade race_eth nYEAR  ;
run;

/* By strata year */;
proc surveylogistic DATA=RecodDat.'STYEARrecod$'n ;
domain HS21;
stratum STR;
cluster PSU;
weight WT;

by nyear;

class grade (REF='4')/param=ref;
class race_eth (REF='4')/param=ref;
model nTST01 (EVENT='1')= nNPL01 IF01 nAAI01 nUZC01 nIUD01 SEX01 grade race_eth;
run;


********************************************************
/* INTERACTION TEST - NATIONAL DATASET 2005-2011 */
********************************************************;


/*National 2005-2011 */
options obs= 59793;
libname RecodDat 'H:\0-NATIONAL\National2005_2011\dataset_recod.XLS';

/* Note: When combining multiple national YRBS datasets:      /*
/* no need to adjust the weights Note: national YRBS data are  /*
```

```
/* weighted to the sample size; that is, the weighted count of /*
/* respondents is equal to the unweighted count of respondents /*


/*                    CHUNK TEST FOR THE POOL            */;

proc surveylogistic DATA=RecodDat.'STYEARrecod$'n ;
domain HS21;
stratum STR;
cluster PSU;
weight WT;
class grade (REF='4')/param=ref;
class race_eth (REF='4')/param=ref;
class nYEAR (REF='4')/param=ref;
model nTST01 (EVENT='1')= nNPL01 IF01 nAAI01 nUZC01 nIUD01 SEX01 grade race_eth nYEAR
                                        nNPL01*SEX01 IF01*SEX01 nAAI01*SEX01
nUZC01*SEX01 nIUD01*SEX01
                                        nNPL01*race_eth IF01*race_eth nAAI01*race_eth
nUZC01*race_eth nIUD01*race_eth
                                        nNPL01*grade IF01*grade nAAI01*grade
nUZC01*grade nIUD01*grade
                                        nNPL01*nYEAR IF01*nYEAR nAAI01*nYEAR
nUZC01*nYEAR nIUD01*nYEAR;
run;
/* (WT) -2log = 21493.611 */;

proc surveylogistic DATA=RecodDat.'STYEARrecod$'n ;
domain HS21;
stratum STR;
cluster PSU;
weight WT;
class grade (REF='4')/param=ref;
class race_eth (REF='4')/param=ref;
class nYEAR (REF='4')/param=ref;
model nTST01 (EVENT='1')= nNPL01 IF01 nAAI01 nUZC01 nIUD01 SEX01 grade race_eth nYEAR ;
run;
/* (WT) -2log = 21619.136 */;

data pvalue;
df = 50; chisq = 125.5;
pvalue = 1 - probchi (chisq, df);
run;
proc print data=pvalue;
run;


/*        BACKWARD ELIMINATION FROM THE COMPLETE MODEL  */

proc surveylogistic DATA=RecodDat.'STYEARrecod$'n ;
domain HS21;
stratum STR;
cluster PSU;
```

```
weight WT;
class grade (REF='4')/param=ref;
class race_eth (REF='4')/param=ref;
class nYEAR (REF='4')/param=ref;
MODEL nTST01 = nNPL01 IF01 nAAI01 nUZC01 nIUD01 SEX01 grade race_eth nYEAR
                        nNPL01*SEX01 IF01*SEX01 nAAI01*SEX01 nUZC01*SEX01 nIUD01*SEX01
                        nNPL01*race_eth IF01*race_eth nAAI01*race_eth nUZC01*race_eth
nIUD01*race_eth
                        nNPL01*grade IF01*grade nAAI01*grade nUZC01*grade nIUD01*grade
                        nNPL01*nYEAR IF01*nYEAR nAAI01*nYEAR nUZC01*nYEAR nIUD01*nYEAR;
run;

proc surveylogistic DATA=RecodDat.'STYEARrecod$'n ;
domain HS21;
stratum STR;
cluster PSU;
weight WT;
class grade (REF='4')/param=ref;
class race_eth (REF='4')/param=ref;
class nYEAR (REF='4')/param=ref;
MODEL nTST01 = nNPL01 IF01 nAAI01 nUZC01 nIUD01 SEX01 grade race_eth nYEAR
                        IF01*SEX01 nAAI01*SEX01 nUZC01*SEX01
                        nNPL01*race_eth
                        IF01*grade ;
run;


/*EVALUATE IF INTERACTION TERM LEFT ARE MEANINGFUL: TEST BY STRATA */;

/*      NPL-RACE */;
proc surveylogistic DATA=RecodDat.'STYEARrecod$'n ;
domain HS21;
stratum STR;
cluster PSU;
weight WT;
class grade (REF='4')/param=ref;
class race_eth (REF='4')/param=ref;
class nYEAR (REF='4')/param=ref;
model nTST01 (EVENT='1')= nNPL01 IF01 nAAI01 nUZC01 nIUD01 SEX01 grade race_eth nYEAR
                                    IF01*SEX01 nAAI01*SEX01 nUZC01*SEX01
nNPL01*race_eth IF01*grade ;
contrast 'Strata 4: White ; Drug Use 1 vs 0' nNPL01 1 /estimate=both; *** strata 4: White ;
contrast 'Strata 3: Black ; Drug Use 1 vs 0' nNPL01 1 race_eth*nNPL01 0 0 1/estimate=both ; *** strata 3:
Black ;
contrast 'Strata 2: Hisp ; Drug Use 1 vs 0' nNPL01 1 race_eth*nNPL01 0 1 0/estimate=both ; *** strata 2:
Hisp ;
contrast 'Strata 1: Other ; Drug Use 1 vs 0' nNPL01 1 race_eth*nNPL01 1 0 0/estimate=both ; *** strata 1:
Other ;
run;

/*      IF - GRADE */;
proc surveylogistic DATA=RecodDat.'STYEARrecod$'n ;
domain HS21;
```

```
stratum STR;
cluster PSU;
weight WT;
class grade (REF='4')/param=ref;
class race_eth (REF='4')/param=ref;
class nYEAR (REF='4')/param=ref;
model nTST01 (EVENT='1')= nNPL01 IF01 nAAI01 nUZC01 nIUD01 SEX01 grade race_eth nYEAR
                                                         IF01*SEX01 nAAI01*SEX01 nUZC01*SEX01
nNPL01*race_eth IF01*grade ;
contrast 'Strata 4: 9th grade ; Intercourse Forced 1 vs 0' grade 1 /estimate=both; *** strata 4: 9th grade ;
contrast 'Strata 3: 10th grade ; Intercourse Forced 1 vs 0' grade 1 IF01*grade 0 0 1/estimate=both ; ***
strata 3: 10th grade ;
contrast 'Strata 2: 11th grade ; Intercourse Forced 1 vs 0' grade 1 IF01*grade 0 1 0/estimate=both ; ***
strata 2: 11th grade ;
contrast 'Strata 1: 12th grade ; Intercourse Forced 1 vs 0' grade 1 IF01*grade 1 0 0/estimate=both ; ***
strata 1: 12th grade ;
run;

/*        SEX - AAI IF UZC  */;
proc surveylogistic DATA=RecodDat.'STYEARrecod$'n ;
domain HS21;
stratum STR;
cluster PSU;
weight WT;
class grade (REF='4')/param=ref;
class race_eth (REF='4')/param=ref;
class nYEAR (REF='4')/param=ref;
model nTST01 (EVENT='1')= nNPL01 IF01 nAAI01 nUZC01 nIUD01 SEX01 grade race_eth nYEAR
                                                         IF01*SEX01 nAAI01*SEX01 nUZC01*SEX01
nNPL01*race_eth IF01*grade ;
contrast 'Strata Sex=1, Female ; Age at Intercourse 1 vs 0' nAAI01 1 SEX01*nAAI01 1 /estimate=both; ***
strata 1: Sex 1 Female ;
contrast 'Strata Sex=0, Male ; Age at Intercourse 1 vs 0' nAAI01 1 /estimate=both ; *** strata 2: Sex 0
Male ;
contrast 'Strata Sex=1, Female ; Intercourse Forced 1 vs 0' IF01 1 SEX01*IF01 1 /estimate=both; *** strata
1: Sex 1 Female ;
contrast 'Strata Sex=0, Male ; Intercourse Forced 1 vs 0' IF01 1 /estimate=both ; *** strata 2: Sex 0 Male ;
contrast 'Strata Sex=1, Female ; Use Condom 1 vs 0' nUZC01 1 SEX01*nUZC01 1 /estimate=both; ***
strata 1: Sex 1 Female ;
contrast 'Strata Sex=0, Male ; Use Condom 1 vs 0' nUZC01 1 /estimate=both ; *** strata 2: Sex 0 Male ;
run;

********************************************************************
 * MULTICOLLINEARITY ASSESSMENT)   *
********************************************************************/;

/*National 2005-2011 */
options obs= 59793;
libname RecodDat 'H:\0-NATIONAL\National2005_2011\dataset_recod.XLS';

/* Initial model*/;
FILENAME collin "H:\0-STATES-FOCUS-working\collin_2011.sas";
```

```
%INCLUDE collin;
ODS OUTPUT SURVEYLOGISTIC.COVB=collin_info;
proc surveylogistic DATA=RecodDat.'STYEARrecod$'n ;
domain HS21;
stratum STR;
cluster PSU;
weight WT;
class grade (REF='4')/param=ref;
class race_eth (REF='4')/param=ref;
class nYEAR (REF='4')/param=ref;
MODEL nTST01 = nNPL01 IF01 nAAI01 nUZC01 nIUD01 SEX01 grade race_eth nYEAR
                                    IF01*SEX01 nAAI01*SEX01 nUZC01*SEX01
                                    nNPL01*race_eth
                                    IF01*grade /COVB ;
run;
%Collin(COVDSN=collin_info,PROCDR=SURVEYLOGISTIC,OUTPUT=collin_info2)

/* */;
FILENAME collin "H:\0-STATES-FOCUS-working\collin_2011.sas";
%INCLUDE collin;
ODS OUTPUT SURVEYLOGISTIC.COVB=collin_info;
proc surveylogistic DATA=RecodDat.'STYEARrecod$'n ;
domain HS21;
stratum STR;
cluster PSU;
weight WT;
class grade (REF='4')/param=ref;
class race_eth (REF='4')/param=ref;
class nYEAR (REF='4')/param=ref;
model nTST01 (EVENT='1')= nNPL01 IF01 nAAI01 nUZC01 nIUD01 SEX01 grade race_eth nYEAR /COVB;
run;
%Collin(COVDSN=collin_info,PROCDR=SURVEYLOGISTIC,OUTPUT=collin_info2)

/* Final model*/;
FILENAME collin "H:\0-STATES-FOCUS-working\collin_2011.sas";
%INCLUDE collin;
ODS OUTPUT SURVEYLOGISTIC.COVB=collin_info;
proc surveylogistic DATA=RecodDat.'STYEARrecod$'n ;
domain HS21;
stratum STR;
cluster PSU;
weight WT;
class grade (REF='4')/param=ref;
class race_eth (REF='4')/param=ref;
class nYEAR (REF='4')/param=ref;
model nTST01 (EVENT='1')= nNPL01 IF01 nAAI01 nUZC01 nIUD01 SEX01 grade race_eth /COVB;
run;
%Collin(COVDSN=collin_info,PROCDR=SURVEYLOGISTIC,OUTPUT=collin_info2)
```

```
********************************************************
                    *           DATASET DESCRIPTION  - MA  *
********************************************************/;
libname RecodDat 'H:\0-STATES\dataset_recod.XLS';

/* DESCRIPTION: Demographics and Behavior Characteristic */;

proc surveyfreq data=RecodDat.'MA_recod$'n total=stratum_pop nomcar; /* nTST21 */;
stratum STR;
cluster PSU;
weight WT;
Tables YR*SEX21 / row(deff) CL(type=logit) ;
Tables YR*grade / row(deff) CL(type=logit) ;
Tables YR*race_eth / row(deff) CL(type=logit) ;
Tables YR*nTST21 / row(deff) CL(type=logit) ;
Tables YR*nTCH21 / row(deff) CL(type=logit) ;
Tables YR*nUZC21 / row(deff) CL(type=logit) ;
Tables YR*nNPL21 / row(deff) CL(type=logit) ;
Tables YR*nAAI21 / row(deff) CL(type=logit) ;
Tables YR*IF21 / row(deff) CL(type=logit) ;
Tables YR*nIUD21 / row(deff) CL(type=logit) ;
run;

proc surveyfreq data=RecodDat.'MA_recod$'n total=stratum_pop nomcar;
stratum STR;
cluster PSU;
weight WT;
Tables HS21*SEX21 / row(deff) CL(type=logit) ;
Tables HS21*grade / row(deff) CL(type=logit) ;
Tables HS21*race_eth / row(deff) CL(type=logit) ;
Tables HS21*nTST21 / row(deff) CL(type=logit) ;
Tables HS21*nTCH21 / row(deff) CL(type=logit) ;
Tables HS21*nUZC21 / row(deff) CL(type=logit) ;
Tables HS21*nNPL21 / row(deff) CL(type=logit) ;
Tables HS21*nAAI21 / row(deff) CL(type=logit) ;
Tables HS21*IF21 / row(deff) CL(type=logit) ;
Tables HS21*nIUD21 / row(deff) CL(type=logit) ;
run;

proc surveyfreq data=RecodDat.'MA_recod$'n total=stratum_pop nomcar;
stratum STR;
cluster PSU;
weight WT;
Tables HS21 / row(deff) CL(type=logit) ;
run;


********************************************************
                    *           MA - TREND OVERTIME  *
********************************************************/;
options obs= 12089;
libname dataset 'H:\0-STATES\YRBS_TBA_MA.xls';
```

```
DATA stratum_pop; set dataset.'Stratum_pop_MA-STR$'n; run;

libname RecodDat 'H:\0-STATES\dataset_recod.XLS';
proc surveyfreq data=RecodDat.'MA_recod$'n total=stratum_pop nomcar;
stratum STR;
cluster PSU;
weight WT;
Tables HS21*YR*nTST21 / row(deff) CL(type=logit) ;
Tables HS21*YR*nUZC21 / row(deff) CL(type=logit) ;
Tables HS21*YR*nNPL21 / row(deff) CL(type=logit) ;
Tables HS21*YR*nAAI21 / row(deff) CL(type=logit) ;
Tables HS21*YR*IF21 / row(deff) CL(type=logit) ;
Tables HS21*YR*nIUD21 / row(deff) CL(type=logit) ;
Tables YR*HS21 / row(deff) CL(type=logit) ;
run;

/*Calculate orthogoanal coefficients: linear and quadratic */
PROC IML; X={2005 2007 2009 2011}; XP=ORPOL(X,3); PRINT XP; RUN; QUIT;

/*Plug orthogoanal coefficients */
DATA VARSET; SET RecodDat.TRENDWK;
   IF (YR=2005 OR YR=2007 OR YR=2009 OR YR=2011);
   IF YR=2005 THEN DO; T4L=-0.67082; T4Q=0.5; end;
ELSE IF YR=2007 THEN DO; T4L=-0.223607; T4Q=-0.5; end;
ELSE IF YR=2009 THEN DO; T4L=0.2236068; T4Q=-0.5; end;
ELSE IF YR=2011 THEN DO; T4L=0.6708204; T4Q=0.5; end;

/*Test if change over time: for outcome and risk behaviors */

/*TST01 */
PROC RLOGIST DATA=VARSET DESIGN=WR FILETYPE=SAS;
NEST YR nSTR PSU/PSULEV=3 MISSUNIT;
WEIGHT WT;
SUBPOPN HS21=1;
CLASS SEX01 RACE_ETH GRADE;
REFLEVEL SEX01=0 RACE_ETH=4 GRADE=4;
MODEL nTST01 = SEX01 RACE_ETH GRADE T4L ;
/*test waldf satadjf satadjchi; */
PRINT/BETAFMT=F8.5 SEBETAFMT=F8.5 P_BETAFMT=F8.5;
RUN;
PROC RLOGIST DATA=VARSET DESIGN=WR FILETYPE=SAS;
NEST YR nSTR PSU/PSULEV=3 MISSUNIT;
WEIGHT WT;
SUBPOPN HS21=1;
CLASS SEX01 RACE_ETH GRADE;
REFLEVEL SEX01=0 RACE_ETH=4 GRADE=4;
MODEL nTST01 = SEX01 RACE_ETH GRADE T4L T4Q;
/*test waldf satadjf satadjchi; */
PRINT/BETAFMT=F8.5 SEBETAFMT=F8.5 P_BETAFMT=F8.5;
RUN;

/*nUZC01 */
```

```
PROC RLOGIST DATA=VARSET DESIGN=WR FILETYPE=SAS;
NEST YR nSTR PSU/PSULEV=3 MISSUNIT;
WEIGHT WT;
SUBPOPN HS21=1;
CLASS SEX01 RACE_ETH GRADE;
REFLEVEL SEX01=0 RACE_ETH=4 GRADE=4;
MODEL nUZC01 = SEX01 RACE_ETH GRADE T4L ;
/* test waldf satadjf satadjchi; */
PRINT/BETAFMT=F8.5 SEBETAFMT=F8.5 P_BETAFMT=F8.5;
RUN;
PROC RLOGIST DATA=VARSET DESIGN=WR FILETYPE=SAS;
NEST YR nSTR PSU/PSULEV=3 MISSUNIT;
WEIGHT WT;
SUBPOPN HS21=1;
CLASS SEX01 RACE_ETH GRADE;
REFLEVEL SEX01=0 RACE_ETH=4 GRADE=4;
MODEL nUZC01 = SEX01 RACE_ETH GRADE T4L T4Q;
/* test waldf satadjf satadjchi; */
PRINT/BETAFMT=F8.5 SEBETAFMT=F8.5 P_BETAFMT=F8.5;
RUN;

/*nNPL01 */
PROC RLOGIST DATA=VARSET DESIGN=WR FILETYPE=SAS;
NEST YR nSTR PSU/PSULEV=3 MISSUNIT;
WEIGHT WT;
SUBPOPN HS21=1;
CLASS SEX01 RACE_ETH GRADE;
REFLEVEL SEX01=0 RACE_ETH=4 GRADE=4;
MODEL nNPL01 = SEX01 RACE_ETH GRADE T4L ;
/* test waldf satadjf satadjchi; */
PRINT/BETAFMT=F8.5 SEBETAFMT=F8.5 P_BETAFMT=F8.5;
RUN;
PROC RLOGIST DATA=VARSET DESIGN=WR FILETYPE=SAS;
NEST YR nSTR PSU/PSULEV=3 MISSUNIT;
WEIGHT WT;
SUBPOPN HS21=1;
CLASS SEX01 RACE_ETH GRADE;
REFLEVEL SEX01=0 RACE_ETH=4 GRADE=4;
MODEL nNPL01 = SEX01 RACE_ETH GRADE T4L T4Q;
/* test waldf satadjf satadjchi; */
PRINT/BETAFMT=F8.5 SEBETAFMT=F8.5 P_BETAFMT=F8.5;
RUN;

/*nAAI01 */
PROC RLOGIST DATA=VARSET DESIGN=WR FILETYPE=SAS;
NEST YR nSTR PSU/PSULEV=3 MISSUNIT;
WEIGHT WT;
SUBPOPN HS21=1;
CLASS SEX01 RACE_ETH GRADE;
REFLEVEL SEX01=0 RACE_ETH=4 GRADE=4;
MODEL nAAI01 = SEX01 RACE_ETH GRADE T4L;
test waldf satadjf satadjchi;
/* test waldf satadjf satadjchi; */
```

```
PRINT/BETAFMT=F8.5 SEBETAFMT=F8.5 P_BETAFMT=F8.5;
RUN;
PROC RLOGIST DATA=VARSET DESIGN=WR FILETYPE=SAS;
NEST YR nSTR PSU/PSULEV=3 MISSUNIT;
WEIGHT WT;
SUBPOPN HS21=1;
CLASS SEX01 RACE_ETH GRADE;
REFLEVEL SEX01=0 RACE_ETH=4 GRADE=4;
MODEL nAAI01 = SEX01 RACE_ETH GRADE T4L T4Q;
test waldf satadjf satadjchi;
/* test waldf satadjf satadjchi; */
PRINT/BETAFMT=F8.5 SEBETAFMT=F8.5 P_BETAFMT=F8.5;
RUN;

/*IF01 */
PROC RLOGIST DATA=VARSET DESIGN=WR FILETYPE=SAS;
NEST YR nSTR PSU/PSULEV=3 MISSUNIT;
WEIGHT WT;
SUBPOPN HS21=1;
CLASS SEX01 RACE_ETH GRADE;
REFLEVEL SEX01=0 RACE_ETH=4 GRADE=4;
MODEL IF01 = SEX01 RACE_ETH GRADE T4L ;
test waldf satadjf satadjchi;
/* test waldf satadjf satadjchi; */
PRINT/BETAFMT=F8.5 SEBETAFMT=F8.5 P_BETAFMT=F8.5;
RUN;
PROC RLOGIST DATA=VARSET DESIGN=WR FILETYPE=SAS;
NEST YR nSTR PSU/PSULEV=3 MISSUNIT;
WEIGHT WT;
SUBPOPN HS21=1;
CLASS SEX01 RACE_ETH GRADE;
REFLEVEL SEX01=0 RACE_ETH=4 GRADE=4;
MODEL IF01 = SEX01 RACE_ETH GRADE T4L T4Q;
test waldf satadjf satadjchi;
/* test waldf satadjf satadjchi; */
PRINT/BETAFMT=F8.5 SEBETAFMT=F8.5 P_BETAFMT=F8.5;
RUN;

/*nIUD01 */
PROC RLOGIST DATA=VARSET DESIGN=WR FILETYPE=SAS;
NEST YR nSTR PSU/PSULEV=3 MISSUNIT;
WEIGHT WT;
SUBPOPN HS21=1;
CLASS SEX01 RACE_ETH GRADE;
REFLEVEL SEX01=0 RACE_ETH=4 GRADE=4;
MODEL   nIUD01 = SEX01 RACE_ETH GRADE T4L ;
test waldf satadjf satadjchi;
/* test waldf satadjf satadjchi; */
PRINT/BETAFMT=F8.5 SEBETAFMT=F8.5 P_BETAFMT=F8.5;
RUN;
PROC RLOGIST DATA=VARSET DESIGN=WR FILETYPE=SAS;
NEST YR nSTR PSU/PSULEV=3 MISSUNIT;
WEIGHT WT;
```

```
SUBPOPN HS21=1;
CLASS SEX01 RACE_ETH GRADE;
REFLEVEL SEX01=0 RACE_ETH=4 GRADE=4;
MODEL   nIUD01 = SEX01 RACE_ETH GRADE T4L T4Q;
test waldf satadjf satadjchi;
/* test waldf satadjf satadjchi; */
PRINT/BETAFMT=F8.5 SEBETAFMT=F8.5 P_BETAFMT=F8.5;
RUN;



*************************************************************************
 *          MA - TESTING FOR LINEARITY TO THE RIGHT OF THE JOINPOINT 2007  *
*************************************************************************/;

/*Calculate orthogonal coefficients: linear and quadratic */
PROC IML; X={2007 2009 2011}; XP=ORPOL(X,3); PRINT XP; QUIT;

DATA VARSET; SET RecodDat.TRENDWK;
   IF (YR=2007 OR YR=2009 OR YR=2011);
   IF YR=2007 THEN DO; T4L=-0.707107; end;
ELSE IF YR=2009 THEN DO; T4L=0; end;
ELSE IF YR=2011 THEN DO; T4L=0.7071068; end;

/*TST01 */
PROC RLOGIST DATA=VARSET DESIGN=WR FILETYPE=SAS;
NEST YR STR PSU/PSULEV=3 MISSUNIT;
WEIGHT WT;
SUBPOPN HS21=1;
CLASS SEX01 RACE_ETH GRADE;
REFLEVEL SEX01=0 RACE_ETH=4 GRADE=4;
MODEL nTST01 = SEX01 RACE_ETH GRADE T4L ;
PRINT/BETAFMT=F8.5 SEBETAFMT=F8.5 P_BETAFMT=F8.5;
RUN;

/*TST01 */
PROC RLOGIST DATA=VARSET DESIGN=WR FILETYPE=SAS;
NEST YR STR PSU/PSULEV=3 MISSUNIT;
WEIGHT WT;
SUBPOPN HS21=1;
CLASS SEX01 RACE_ETH GRADE;
REFLEVEL SEX01=0 RACE_ETH=4 GRADE=4;
MODEL nTST01 = SEX01 RACE_ETH GRADE T4L ;
PRINT/BETAFMT=F8.5 SEBETAFMT=F8.5 P_BETAFMT=F8.5;
RUN;

/*nAAI01 */
PROC RLOGIST DATA=VARSET DESIGN=WR FILETYPE=SAS;
NEST YR nSTR PSU/PSULEV=3 MISSUNIT;
WEIGHT WT;
SUBPOPN HS21=1;
CLASS SEX01 RACE_ETH GRADE;
```

```
REFLEVEL SEX01=0 RACE_ETH=4 GRADE=4;
MODEL nAAI01 = SEX01 RACE_ETH GRADE T4L;
PRINT/BETAFMT=F8.5 SEBETAFMT=F8.5 P_BETAFMT=F8.5;
RUN;

/*IF01 */
PROC RLOGIST DATA=VARSET DESIGN=WR FILETYPE=SAS;
NEST YR nSTR PSU/PSULEV=3 MISSUNIT;
WEIGHT WT;
SUBPOPN HS21=1;
CLASS SEX01 RACE_ETH GRADE;
REFLEVEL SEX01=0 RACE_ETH=4 GRADE=4;
MODEL IF01 = SEX01 RACE_ETH GRADE T4L ;
PRINT/BETAFMT=F8.5 SEBETAFMT=F8.5 P_BETAFMT=F8.5;
RUN;

/*nIUD01 */
PROC RLOGIST DATA=VARSET DESIGN=WR FILETYPE=SAS;
NEST YR nSTR PSU/PSULEV=3 MISSUNIT;
WEIGHT WT;
SUBPOPN HS21=1;
CLASS SEX01 RACE_ETH GRADE;
REFLEVEL SEX01=0 RACE_ETH=4 GRADE=4;
MODEL   nIUD01 = SEX01 RACE_ETH GRADE T4L ;
test waldf satadjf satadjchi;
PRINT/BETAFMT=F8.5 SEBETAFMT=F8.5 P_BETAFMT=F8.5;
RUN;


*************************************************************************
 *          MA - TESTING FOR CHANGE TO THE LEFT OF THE JOINPOINT 2007  *
*************************************************************************/;
data VARSET2005_2007;
set RecodDat.TRENDWK;
where YR=2005 or YR=2007;
run;
proc sort data=VARSET2005_2007; by STR PSU; run;

/* t-test for TST01 */
proc multilog data=VARSET2005_2007 FILETYPE=SAS SEMETHOD=ZEGER R=INDEPENDENT;
    nest STR PSU ;
    weight WT;
    subpopn HS21=1 ;
          class YR nTST01 SEX01 RACE_ETH GRADE ;
          model nTST01 = YR SEX01 RACE_ETH GRADE / CUMLOGIT;
          CONDMARG YR / adjrr;
          SETENV LABWIDTH=28 COLWIDTH=7 DECWIDTH=4 COLSPCE=2 TOPMGN=0;
          PRINT beta sebeta deft="Design Effect" t_beta p_beta;
run ;

/* t-test for nAAI01 */
proc multilog data=VARSET2005_2007 FILETYPE=SAS SEMETHOD=ZEGER R=INDEPENDENT;
    nest STR PSU ;
```

```
    weight WT;
    subpopn HS21=1 ;
        class YR nAAI01 SEX01 RACE_ETH GRADE ;
        model nAAI01 = YR SEX01 RACE_ETH GRADE / CUMLOGIT;
        CONDMARG YR / adjrr;
        SETENV LABWIDTH=28 COLWIDTH=7 DECWIDTH=4 COLSPCE=2 TOPMGN=0;
        PRINT beta sebeta deft="Design Effect" t_beta p_beta;
run ;

/* t-test for IF01 */
proc multilog data=VARSET2005_2007 FILETYPE=SAS SEMETHOD=ZEGER R=INDEPENDENT;
    nest STR PSU ;
    weight WT;
    subpopn HS21=1 ;
        class YR IF01 SEX01 RACE_ETH GRADE ;
        model IF01 = YR SEX01 RACE_ETH GRADE / CUMLOGIT;
        CONDMARG YR / adjrr;
        SETENV LABWIDTH=28 COLWIDTH=7 DECWIDTH=4 COLSPCE=2 TOPMGN=0;
        PRINT beta sebeta deft="Design Effect" t_beta p_beta;
run ;

/* t-test for nIUD01 */
proc multilog data=VARSET2005_2007 FILETYPE=SAS SEMETHOD=ZEGER R=INDEPENDENT;
    nest STR PSU ;
    weight WT;
    subpopn HS21=1 ;
        class YR nIUD01 SEX01 RACE_ETH GRADE ;
        model nIUD01 = YR SEX01 RACE_ETH GRADE / CUMLOGIT;
        CONDMARG YR / adjrr;
        SETENV LABWIDTH=28 COLWIDTH=7 DECWIDTH=4 COLSPCE=2 TOPMGN=0;
        PRINT beta sebeta deft="Design Effect" t_beta p_beta;
run ;
```

REFERENCES

1. Center for Disease Control and Prevention. HIV Surveillance Report. *http://wwwcdcgov/hiv/topics/surveillance/resources/reports/* 2011.

2. Center for Disease Control and Prevention. HIV Surveillance Supplemental Report. *http://wwwcdcgov/hiv/topics/surveillance/resources/reports/* 2012;17(3):Part A.

3. Branson BM, Viall A,Marum E. Expanding HIV testing: back to the future. *J Acquir Immune Defic Syndr* 2013;63 (Suppl 2):S117-21.

4. Branson BM, Handsfield HH, Lampe M, et al. Revised recommendations for HIV testing of adults, adolescents, and pregnant women in health-care settings. *MMWR Recomm Rep* 2006;55(RR-14):1-17; quiz CE1-4.

5. Moyer VA,U.S. Preventive Services Task Force. Screening for HIV: U.S. Preventive Services Task Force Recommendation Statement. *Ann Intern Med* 2013 159(1):51-60.

6. Balaji AB, Eaton DK, Voetsch AC, et al. Association between HIV-related risk behaviors and HIV testing among high school students in the United States, 2009. *Arch Pediatr Adolesc Med* 2012;166(4):331-6.

7. Eaton DK, Kann L, Kinchen S, et al. Youth risk behavior surveillance - United States, 2011. *MMWR Surveill Summ* 2012 61(4):1-162.

8. Anglemyer A, Rutherford GW, Horvath T, et al. Antiretroviral therapy for prevention of HIV transmission in HIV-discordant couples. *Cochrane Database Syst Rev* 2013.

9. Fowler MG, Gable AR, Lampe MA, et al. Perinatal HIV and its prevention: progress toward an HIV-free generation. *Clin Perinatol* 2010;37(4):699-719.

10. Peterman TA, Todd KA,Mupanduki I. Opportunities for targeting publicly funded human immunodeficiency virus counseling and testing. *J Acquir Immune Defic Syndr Hum Retrovirol* 1996;12(1):69-74.

11. Long EF, Brandeau ML,Owens DK. The cost-effectiveness and population outcomes of expanded HIV screening and antiretroviral treatment in the United States. *Ann Intern Med* 2010;21(153):12.

12. Paltiel AD, Walensky RP, Schackman BR, et al. Expanded HIV screening in the United States: effect on clinical outcomes, HIV transmission, and costs. *Ann Intern Med* 2006;145(11):797-806.

13. Qaseem A, Snow V, Shekelle P, et al. Screening for HIV in health care settings: a guidance statement from the American College of Physicians and HIV Medicine Association. *Ann Intern Med* 2009;150(2):125-31.

14. Lubinski C, Aberg J, Bardeguez AD, et al. HIV policy: the path forward--a joint position paper of the HIV Medicine Association of the Infectious Diseases Society of America and the American College of Physicians. *Clin Infect Dis* 2009;48(10):1335-44.

15. American College of Obstetricians and Gynecologists. ACOG Committee Opinion. Routine human immunodeficiency virus screening. *Obstet Gynecol* 2008;112(2 Pt 1):401-3.

16. Committee on Pediatric AIDS, Emmanuel PJ,Martinez J. Adolescents and HIV infection: the pediatrician's role in promoting routine testing. *Pediatrics* 2011;128(5):1023-9.

17. American Academy of Family Physicians. Screening for HIV and treatment of acquired immunodeficiency policy statement. *http://wwwaafporg/patient-care/clinical-recommendations/all/hivhtml* 2007.

18. Centers for Disease Control and Prevention. Results of the Expanded HIV Testing Initiative--25 jurisdictions, United States, 2007-2010. *MMWR Morb Mortal Wkly Rep*, 2011:805-10.

19. Hutchinson AB, Farnham PG, Duffy N, et al. Return on public health investment: CDC's Expanded HIV Testing Initiative. *J Acquir Immune Defic Syndr* 2012.

20. Peralta L, Deeds BG, Hipszer S, et al. Barriers and facilitators to adolescent HIV testing. *AIDS Patient Care STDS* 2007;21(6):400-8.

21. Lehrer JA, Pantell R, Tebb K, et al. Forgone health care among U.S. adolescents: Associations between risk characteristics and confidentiality concern. *Journal of Adolescent Health* 2007;40:218-26.

22. Lim SW, Chhabra R, Rosen A, et al. Adolescents' views on barriers to health care. *Journal of Primary Care & Community Health* 2012;3:99-103.

23. Grusky O, Duan N, Swanson AN, et al. Evaluating the accessibility of HIV testing organizations. . *Evaluation & the Health Professions* 2004;27:189-205.

24. Akinbami LJ, Gandhi H,Cheng TL. Availability of adolescent health services and confidentiality in primary care practices. *Pediatrics* 2003;111:394-401.

25. Hyden C, Allegrante JP,Cohall AT. HIV Testing Sites' Communication About Adolescent Confidentiality: Potential Barriers and Facilitators to Testing. *Health Promot Pract http://hppsagepubcom/content/early/2013/08/20/1524839913499347* 2013

26. Kirby D, Coyle K, Alton F, et al. Reducing Adolescent Sexual Risk A Theoretical Guide for Developing and Adapting Curriculum-Based Programs. *http://pubetrorg/upfiles/reducing_adolescent_sexual_riskpdf* 2011.

27. Center for Control Disease and Prevention. HIV Testing Among Adolescents: What Schools and Education Agencies Can Do. http://www.cdc.gov/healthyyouth/sexualbehaviors/pdf/hivtesting_adolescents.pdf. 2012.

28. Branson BM, Handsfield HH, Lampe MA, et al. Revised recommendations for HIV testing of adults, adolescents, and pregnant women in health-care settings. *Morbidity and Mortality Weekly Report* 2006;55:1-17.

29. Center for Disease Control and Prevention. State HIV laws. *http://wwwcdcgov/hiv/policies/law/states/indexhtml*, 2013.

30. Hahn EK. Incorporating the CDC Recommendations for Adolescent HIV Screening into Practice. *Journal for Nurse Practitioners* 2009;5(4):265-73.

31. Brener ND, Kann L, McManus TL, et al. Reliability of the 1999 Youth Risk Behavior Survey Questionnaire. *Journal of Adolescent Health* 2002;31:336-42.

32. Brener ND, Collins JL,Kann L. Reliability of the Youth Risk Behavior Survey questionnaire. *American Journal of Epidemiology* 1995;141:575-80.

33. National Center for Education Statistics. Percentage of high school dropouts among persons 16 through 24 years old (status dropout rate), by sex and race/ethnicity: Selected years, 1960 through 2011. *http://ncesedgov/programs/digest/d12/tables/dt12_128asp* 2012.