BIOSTATISTICS 600

Global Activity Two

Male Circumcision and HIV Infection in African populations: Revisited

INTRODUCTION

Many studies have investigated the association between *Male Circumcision* and *HIV Prevalence*. Global Activity One explored this relationship in 37 African countries using a continuous measure for *Male Circumcision* and for *HIV Prevalence*.

A more recent article (Drain 2006) explores the relationship between *Male Circumcision* and *HIV Prevalence* in a somewhat different way. *HIV Prevalence* was estimated from the Joint United Nations Programme on HIV/AIDS. *Male circumcision* rates were categorized as either "Low" (<20%), "Intermediate" (20-80%), or "High" (>80%). Investigators found a significantly higher prevalence of HIV in African countries with "Low" rate of male circumcision compared to African countries with a "High" rate of male circumcision from a two- sample t-test. Investigators also explored the relationship between the two factors using univariate linear regression and correlation. Using the data provided (primarily updated values from the Joint United Nations: Report on the Global HIV/AIDS epidemic 2008), students will investigate the relationship between *Male Circumcision* and *HIV Prevalence* in Africa using methods based on the (Drain 2006) article. Students will reproduce some results from the original article as well as develop some new results.

SOURCES

Drain PK, Halperin DT, Hughes JP, Klausner JD, Bailey RC. 2006. Male circumcision, religion and infectious diseases: an ecologic analysis of 118 developing countries. BMC Infectious Diseases 6:172.
Open Access: www.biomedcentral.com/1471-2334/6/172

Joint United Nations Programme on HIV/AIDS: Report on the Global HIV/AIDS Epidemic 2004 Geneva: UNAIDS; 2004.

Joint United Nations Programme on HIV/AIDS: Report of the Global HIV/AIDS Epidemic 2008 Mexico City: UNAIDS; 2008

Note: Data for questions based on the (Drain 2006) article were obtained from http://www.unaids.org/en/KnowledgeCentre/HIVData/GlobalReport/Archive.asp (accessed June 2, 2009)

QUESTIONS

For the following problems, refer to the data set "GA_Two_Drain.xls". Questions are based on the methods and findings from the article Drain (2006). Students' answers may be different in some cases from the original article due to updated data sources and other factors. Forty-four African nations were grouped into three categories based on the percentage of males circumcised: *Low (<20%), Intermediate (20-80%)* or *High (>80%).* Unless otherwise stated, use the column labeled *2007 HIV prevalence percent Adult (15–49)* as the outcome variable.

1. Calculate the mean, standard deviation, median, and IQR for the *2007 HIV Prevalence Percent* for each *Male Circumcision Category*. Report your results in a well-organized table.

2. Construct side-by-side boxplots for *2007 HIV Prevalence Percent* for the three *Male Circumcision Categories* (by hand or software). Describe the relationship.

3. Conduct a statistical test to compare the average *2007 HIV Prevalence* in the two *Male Circumcision Categories: Low*. Vs. *High*. Include these steps: State the null and alternative hypothesis. Check the assumptions. Calculate an appropriate test statistic and corresponding *p*-value. Interpret the *p*-value for a nonstatistician.

4. Calculate a 95% Confidence interval for the difference in average 2007 *HIV prevalence* in the two groups, *Low* vs. *High Male Circumcision Categories*. Carefully interpret the confidence interval for someone who has no statistical training.

5. i) The three categories for *Male Circumcision* were coded as *1=Low, 2=Intermediate, 3= High* in the original journal article (Drain 2006). Those values are provided in the given dataset. Produce a scatterplot with *x= Numerical Category of Male Circumcision* and *y= 2007 Prevalence of HIV in Adults.* Include the linear regression equation fit to the data.

ii) Conduct a linear regression analysis for *x=Numerical Category of Male Circumcision* and *y= 2007 Prevalence of HIV in Adults*. Include a plot of the residuals. Check the assumptions. What assumption(s) are violated?

iii) There are more countries with *Numerical Category of Male Circumcision =3* than the other categories. Is this problematic?

iv) Consider the coding of the categories as *1=Low, 2=Intermediate, 3= High*. Do you think it is valid to assume equal spacing for these three categories?

6. i) Produce a scatterplot with *x= Numerical Category of Male Circumcision* and *y= ln(2007 Prevalence of HIV per 100,000 Adults).* These *y* values are calculated for you in the given dataset. Include the linear regression equation fit to the data.

ii) Conduct a linear regression analysis for *x=Numerical Category of Male Circumcision* and *y= ln(2007 Prevalence of HIV per 100,000 Adults).* Include a plot of the residuals. Check the

assumptions. Do the log transformed data in this problem meet the linear regression assumptions better than the original data above?

7. Calculate the correlation coefficient and the coefficient of variation between *Numeric Category of Male Circumcision* and *log(2007 HIV Prevalence Percent in Adults)*. (Use the transformed data, as in Question 6). Conduct a statistical test for the correlation coefficient equal to zero and carefully interpret the *p*-value.

8. Investigators are interested in the change in average *HIV Prevalence* between the years 2001 and 2007. Conduct a statistical test to address this question.

9. Suppose investigators wished to compare the three mean *HIV Prevalences* for all the *Male Circumcision Categories* (rather than just *Low* vs. *High* in previous problem and original article). Conduct a statistical test to compare the three averages (*Low*, *Intermediate* and *High*).

Additional Sources regarding the relationship between *HIV* and *Male Circumcision* in Africa:

Auvert B, Taljaard P, Lagarde E, Sobngwi-Tarbekou J, Sitta R, Puren A. 2005. Randomized, Controlled Intervention Trial of Male Circumcision for Reduction of HIV Infection Risk: The ANRS 1265 Trial. PLoS Med 2(11):e298, 1112-22. www.plosmedicine.org

Bailey RC, Moses S, Parker CB, Agot K, Maclean I, Krieger JN, Williams CFM, Campbell RT, Ndinya-Achola JO. 2007. Male circumcision for HIV prevention in young men in Kisumu, Kenya: a randomized, controlled trial. The Lancet 369: 643-56.

Bailey RC, Plummer FA, Moses S. 2001. Male circumcision and HIV prevention: current knowledge and future research directions. The Lancet Infectious Disease 1:223-31.

Bongaarts J, Reining P, Way P, Conant F. 1989. The relationship between male circumcision and HIV infection in African populations. AIDS 3:373-7.

BIOSTATISTICS TOPICS:

DESCRIPTIVE STATISTICS, TWO-SAMPLE T-TEST, CONFIDENCE INTERVAL FOR DIFFERENCE IN MEANS, LINEAR REGRESSION, DATA TRANSFORMATION, CORRELATION, MATCHED PAIRS T-TEST, ANOVA.