1 Male circumcision & HIV prevention: Kenya

The Figure below is from the article: "Male circumcision for HIV prevention in young men in Kisumu, Kenya: a randomised controlled trial" (Lancet 2007; 369: 643-656). The full article is included in the zip file, and also available under "resources for rates" in course `EPI634`.

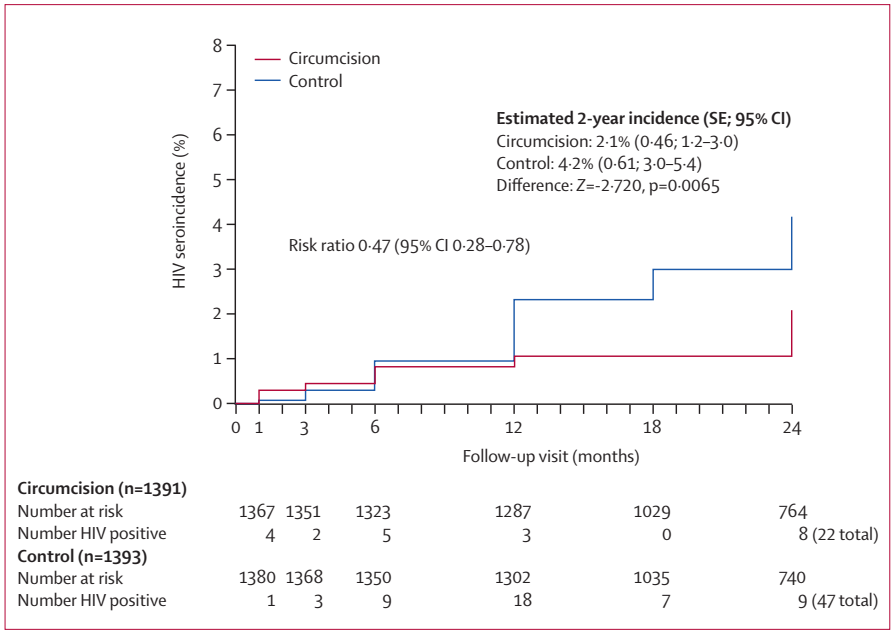


Figure 2: **Cumulative HIV seroincidence across follow-up visits by treatment**
Time to HIV-positive status is taken as the first visit when a positive HIV test result is noted. Time is credited as the follow-up visit month. Participants without HIV-positive status are censored at the last regular follow-up visit completed where HIV testing was done, credited specifically as months 1, 3, 6, 12, 18, and 24.

For those of you who prefer (and JH encourages you) to use `R`, SAS, Stata or Excel rather than your portable $10 calculator, the raw data in the figure are also available in an enclosed data-file, and can be processed using some provided R code (with some steps left for you to fill in); these files are also available under the "HIV" link in course `EPI634`.

Exercise

a Replicate each of the 3 lines of statistics in the insert beginning with the text "Estimated 2-year incidence" in the top right portion of the Figure.[1]

b Compare the reported SE of 0.61% with the (naive) SE one would obtain by the formula $100 \times \{0.042 \times 0.958/1380\}^{1/2}$. Explain the difference.[2]

c The reported *risk* ratio of 0.47 is either a misnomer for the estimate of the *hazard* (i.e., *rate*) ratio, from Cox regression, or the ratio of the two reported 2-year cumulative incidence estimates; the 95% limits are compatible with the latter). One could also calculate a single crude *rate ratio* [i.e., *incidence ratio* in Rothman's terminology, *incidence density ratio* in Miettinen's]. Calculate this crude *rate ratio* and its associated 95%CI (*hints*: see Rothman chapter 7; for person-time calculations, assume seroconversions occurred midway through the intervals).

d The assumption of a 'constant rate ratio across time windows' is implied by the calculation in (c), and implicit in the "reduction in the risk of acquiring an HIV infection of 53%" reported in the abstract. Also, the distribution of follow-up time is weighted slightly more to the first year than the second; thus if the true rate ratio id different in different follow-up windows, a single ratio will reflect this.

   If you calculated a separate rate ratio and 95% confidence interval for each of the six follow-up windows, the confidence intervals would be so wide that you could 'drive a very wide truck' *horizontally* through virtually all six of them; i.e., there aren't sufficient data in each interval to check for trends in the rate ratios, i.e., to check tell whether (hazards) are or are not proportional. Thus, as a next-best approach, aggregate the data into 1-year windows, and determine whether the 95% confidence intervals for the two resulting rate ratios are distinct, or overlap.[3]

[1] "The Kaplan-Meier method was used to estimate the HIV event distribution over time by treatment, accounting for staggered enrolment and incomplete, discrete follow-up. Estimates of 2-year HIV seroincidences and corresponding standard errors obtained by Greenwoods formula were used to test for differences between the treatments on the primary outcome (HIV seroconversion)."

[2] Comment: some authors refer to the n' that satisfies the equality $100 \times \{0.042 \times 0.958/n'\}^{1/2} = 0.61\%$ as the *'effective'* sample size.

[3] Checking whether two independent 95% CI's don't or do overlap is *not quite the same* as testing whether the two point estimates are significantly different at the 0.05 level. If they don't overlap, then yes $P < 0.05$. But if they do overlap a little, it can be still be that $P < 0.05$. The more accurate way is to calculate a single 95% CI for the difference, using the SE of this difference. If interested, see the enclosed article by Wolfe and Hanley: "If we're so different, why do we keep overlapping? When 1 plus 1 doesn't make 2." *Canadian Medical Association Journal.* 2002 Jan 8;166(1):65-66 – it is permanently available under `r e p r i n t s` on JH's homepage.

2 HIV prevention trial: Uganda

The Table and Figure opposite are from "Male circumcision for HIV prevention in men in Rakai, Uganda: a randomised trial," Lancet 2007; 369: 657-666, published back to back with the report on the trial in Kenya.

a Comment on the appropriateness/accuracy of the terms (i) "*Cumulative number of participants*" and (ii) "*Cumulative incidence per 100 person-years*" in the bottom panel of Table 3.

b Try to replicate the point estimate of 0.35, the 95% CI of 0.10-1.04, and the p = 0.0389[4] in the last row of the '6-12 months follow-up interval' panel.

c "*An overall risk difference and risk ratios were calculated at the end of follow-up, with CI based on standard Greenwood formula variance estimates.*" [Statistical analysis, Methods section.]

"*Figure 2 shows the Kaplan-Meier survival curves for time-to-detection of HIV infection for the modified intention-to-treat analysis. The difference between the cumulative probabilities of HIV detection was significant (p = 0.003)*" [Results].

JH presumes this p-value was calculated in the same way as in the Kenya study. Instead, test the equality of the two curves using the log-rank test.

d Estimate the risk (i.e., cumulative incidence) difference at the end of follow-up using Nelson-Aalen rather than Kaplan-Meier estimators.[5]

==========

The results in these next two quotes are based on Poisson and Cox regression, and are included here merely so we begin thinking about how to represent these concepts in regression models – the topic of a future lecture.

"Table 3 shows HIV incidence by study arm and follow-up visit intervals [...] The intention-to-treat analysis showed *a progressive decrease in incidence in the intervention group over the entire follow-up period* (p for trend 0.014). Incidence fell in the control group between the time of first follow-up and the time of second follow-up, and remained stable thereafter; however, the trend was not significant (p=0.6)."

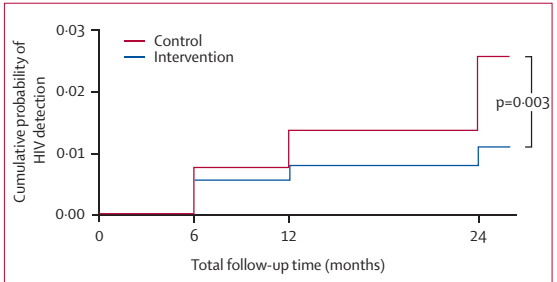
"The IRR [Incidence Rate Ratio] of HIV acquisition associated with circumcision also fell over time; this increase in efficacy was of borderline significance (p=0.054 for the time-by-study arm interaction)."

[4] PS: do you see a conflict between this CI and this p value?

[5] R code specific for these data not provided, but if you wish, cf. R code for JUPITER data.

	Intervention group	Control group	Incidence rate ratio (95% CI)	p value
0–6 months follow-up interval				
Number of participants	2263	2319		
Incident events	14	19		
Person-years	1172·1	1206·7		
Incidence per 100 person-years	1·19	1·58	0·76 (0·35–1·60)	0·439
6–12 months follow-up interval				
Number of participants	2235	2229		
Incident events	5	14		
Person-years	1190·7	1176·3		
Incidence per 100 person-years	0·42	1·19	0·35 (0·10–1·04)	0·0389
12–24 months follow-up interval				
Number of participants	964	980		
Incident events	3	12		
Person-years	989·7	1008·7		
Incidence per 100 person-years	0·30	1·19	0·25 (0·05–0·94)	0·0233
Total 0–24 months follow-up				
Cumulative number of participants	2387	2430		
Cumulative incident events	22	45		
Cumulative person-years	3352·4	3391·8		
Cumulative incidence per 100 person-years	0·66	1·33	0·49 (0·28–0·84)	0·0057

Table 3: **HIV incidence by study group and follow-up interval, and cumulative HIV incidence over 2 years**



Cases of HIV/total participants

Intervention	0/2474	14/2387	5/2274	3/964
Control	0/2522	19/2430	14/2279	12/980

Figure 2: **Kaplan-Meier cumulative probabilities of HIV detection by study group**
Actual visits grouped by the three scheduled visits at 6 months, 12 months, and 24 months after enrolment. The cumulative probabilities of HIV infection were 1·1% in the intervention group and 2·6% in the control group over 24 months.

The raw data in the Table are available in a data-file, and can be processed using the provided R code (with some steps left for you to fill in).