From the BBC website (9/25/09) comes the following report:

An experimental HIV vaccine has for the first time cut the risk of infection, researchers say. The vaccine - a combination of two earlier experimental vaccines - was given to 16,000 people in Thailand, in the largest ever such vaccine trial. Researchers found that it reduced by nearly a third the risk of contracting HIV, the virus that leads to Aids. It has been hailed as a significant, scientific breakthrough, but a global vaccine is still some way off. The study was carried out by the US army and the Thai government over seven years on volunteers - all HIV-negative men and women aged between 18 and 30 - in parts of Thailand.

$\approx$

The vaccine was a combination of two older vaccines that on their own had not cut infection rates. The vaccine is based on B and E strains of HIV that most commonly circulate in Thailand not the C strain which predominates in Africa.

Half of the volunteers were given the vaccine, while the other half were given a placebo - and all were given counselling on HIV/Aids prevention. Participants were tested for HIV infection every six months for three years.

$\approx$

Question 3: How would you analyze these data using the hypothesis test techniques you've learned in MA 2611?

$\approx$

Let's use the 5-step method:

1. **The Scientific Hypothesis:** The vaccine is effective.
2. **The Statistical Model:** Two population binomial: $b(8198, p_p)$ and $b(8197, p_v)$, where p_p is the population proportion that would get HIV infection if given a placebo, and p_v, the population proportion that would get HIV infection if given the vaccine.
3. **The Statistical Hypotheses:**

$$\begin{array}{lccc} H_0: & p_p - p_v & = & 0 \\ H_a: & p_p - p_v & > & 0 \\ \text{or} & & & \\ H_a: & p_p - p_v & \neq & 0? \end{array}$$

We will assume the latter.

$\approx$

4. **The Test Statistic:**

The observed value of the standardized test statistic is

$$z_0^* = \frac{\hat{p}_p - \hat{p}_v}{\sqrt{\hat{p}(1-\hat{p})\,(1/n_1 + 1/n_2)}}, \text{ where}$$

$\hat{p} = (74 + 51)/(8198 + 8197) = 0.0076$, so

$$z_0^* = \frac{74/8198 - 51/8197}{\sqrt{(0.0076)(0.9924)(1/8198 + 1/8197)}} = 2.0644.$$

Then we have $p^+ = Pr(N(0,1) > 2.0644) = 0.0195$ and $p^+ = Pr(N(0,1) < 2.0644) = 0.9805$ The p-value is then $p_\pm = 2\min(p_-, p^+) = 0.0390$.

Note that we could use the large sample interval, since the numbers of HIV cases and non-cases are large (> 10) in both groups.

$\approx$

5. **Results and Interpretation:**
The p-value, 0.0390, is interpreted as follows: If H_0 is true (ie, if the vaccine has no effect), then only 3.9% of similar experiments would result in as large (or larger) a difference in the observed proportions of infected individuals.

Statistically, we can say the observed difference is significant at the 0.05 level (and, in fact, at any significance level greater than 0.039).

$\approx$