BIOSTATISTICS 600
Global Topics Activity Four
Insecticide Treated Nets for Prevention of Anemia due to Malaria
ANSWER KEY

INTRODUCTION

In sub-saharan African, malaria is a major cause of anemia during pregnancy. Several studies have explored whether Insecticide Treated Nets (ITNs) could reduce malaria prevalence and anemia prevalence among pregnant women.

In a randomized, placebo-controlled double blind study (Njagi, 2003), investigators collected data from 752 pregnant women who were randomized to 4 groups: ITN+SP (Insecticide Treated Nets + Sulfadoxine Pyrimethamine, an oral malaria preventative medication), ITN (Insecticide Treated Net plus Oral Malaria Placebo), SP (Oral Malaria Preventative Medication), Placebo (Oral Malaria Placebo). Outcomes of interest were percentage of women who were anemic at delivery and average hemoglobin levels at delivery. The relationship between malaria exposure/anemia is hypothesized to be different depending on whether the pregnancy is the first (primagravida) or the second (secundigravida); therefore, most analyses were conducted separately for the two groups. The study reports a significant difference in the proportion of anemic women at delivery across the intervention groups for primagravida women (p=0.02). For example, 48% of primagravida women in the placebo group were anemic at delivery compared to 35% of primagravida women in the ITN group.

An observational study (Marchant, 2002) also investigated the association between ITN and malaria. In this study, pregnant women were asked about ITN use and other factors and blood samples were taken. Women who use ITN were less likely to be positive for malaria than women who did not use ITN (25% vs. 33%, p=0.06). ITN users were also less likely to be anemic than non-ITN users (72% vs. 82%, p= 0.01).

In this activity, students will reproduce many of the results from the original articles (Njagi, 2003, Marchant 2002) as well as explore new results available from the data provided. Attention to assumptions and interpretation of results will be emphasized.

SOURCES

A.

Njagi JK, Magnussen P, Estambale B, Ouma J, Mugo B. 2003. Prevention of anaemia in pregnancy using insecticide-treated bednets and sulfadoxine-pyrimethamine in a highly malarious area of Kenya: a randomized controlled trial. Transactions of the Royal Society of Tropical Medicine and Hygiene. 97:277-82.

B.

Marchant T, Schellenberg JA, Edgar T, Nathan R, Abdulla S, Mukasa O, Mponda H, Lengeler C. 2002. Socially marketed insecticide-treated nets improve malaria and anaemia in pregnancy in Southern Tanzania. 7(2):149-58.

QUESTIONS

1. The Njagi article reports for the following for women in the Insecticide Treated Net (ITN) primagravidas group. At recruitment, $n=104$ women had the average *hemoglobin*, $\bar{x}_{rec}=103.4$ g/L with standard deviation $s_{rec}=18.2$ g/L. At delivery, their average *hemoglobin* for these women was $\bar{x}_{del}=108.4$ g/L with standard deviation $s_{del}=19.3$ g/L. The reported difference in these averages was 5.0 g/L with *p*-value = 0.059.

a) Explain why we can **not** use a two sample *t*-test to test whether the average *hemoglobin* levels are equal. What statistical test would be appropriate? Are we able to reproduce this statistical test with the given information?

b) Consider the following values given in the paper. In the ITN group at delivery, the sample *hemoglobin* statistics are $\bar{x}_{itn} = 108.4$ g/L, $s_{itn}=19.3$ g/L, $n_{itn}=104$. In the Placebo group at delivery, the sample *hemoglobin* statistics are $\bar{x}_{pl} = 102.2$ g/L, $s_{pl}=22.4$ g/L, $n_{pl}=94$. Is there evidence that the population average *hemoglobin* levels are different at delivery in the two populations?

c) Which question is more interesting, "Is the average *hemoglobin* in the *ITN group* different at recruitment compared to average *hemoglobin* in the *ITN group* at delivery? (Q1a)" OR "Is the average *hemoglobin* in the *ITN group* equal to the average *hemoglobin* in the *Placebo group* at delivery? (Q1b)" . Why?

d) What statistical method would you use to test whether averages for the four groups (*ITN+SP, SP, ITN, Placebo*) were equal at delivery for primagravidas in the population?

> **1. a)** Because the two sample averages are for the same patients (at two different time points, recruitment and delivery), the assumption of independence is violated. We would need to conduct a matched pairs *t*-test to compare these averages. To conduct a matched pairs *t*-test, we would need the recruitment and delivery *hemoglobin* for each woman, and calculate the difference for each women ("after *hemoglobin* level" minus "before *hemoglobin* value", for example). Then we could calculate average of those differences and test whether the average in the population is equal to zero. Without the individual values for <u>each</u> woman, we can't reproduce the matched pairs *t*-test.
>
> **b)** We <u>can</u> conduct a two-sample *t*-test to compare the average *hemoglobin*

level in the *ITN group* with the average *hemoglobin* level in the *Placebo group* at delivery. This comparison is not made in the actual journal article, but we are given enough information to conduct the statistical test.

H_0: The average *hemoglobin* level in the *ITN group* is the same as the average *hemoglobin* level in the *Placebo group* at delivery in the population.
H_a: The average *hemoglobin* level in the *ITN group* is not the same as the average *hemoglobin* level in the *Placebo group* at delivery in the population.

$$H_0: \mu_{ITN} = \mu_P \qquad H_a: \mu_{ITN} \neq \mu_P$$

Check Assumptions:
- The observations need to be independent. This is likely true, for example we would like the participants to be unrelated – for example not sisters or not from the same household.
- The sample should be a simple random sample. Our sample is not a SRS – however, this assumption is important in this example for generalizability rather than validity.
- *Hemoglobin* should be normally distributed in the two groups. We would need to view the original data and have some knowledge about *hemoglobin* levels in general to check this assumption. However, recall that the t-test is relatively robust. So even if the data were skewed, we can likely proceed with the analysis, because of the relatively large sample size. [One common "rule of thumb", is $(n_1 + n_2) > 39$.] Also see Note 2 below.

Calculate test statistic and p-value:

$$t = \frac{\bar{x}_1 - \bar{x}_2 - (\mu_1 - \mu_2)}{\sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}}} = \frac{108.4 - 102.2 - 0}{\sqrt{\frac{19.3^2}{104} + \frac{22.4^2}{94}}} = 2.076$$

t-distribution
df = 93

$$2P(t > 2.076) = 0.04$$

-2.08   2.08

$$\text{degrees freedom} = \min(n_1 - 1, n_2 - 1) = 93$$
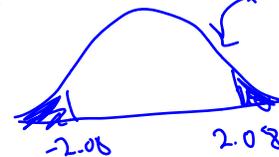
Conclusion:
If the average *hemoglobin* at delivery were the same in the two populations (*ITN* vs. *Placebo*), the probability that we'd observe a sample difference in averages as extreme as 6.2 g/L is 0.04. Our sample averages are somewhat unusual if the populations averages are the same. We have some evidence that the average *hemoglobin* at delivery may be different in the *ITN group* compared to the *Placebo group* in the population.

Note1:
Randomization helps to ensure that the two groups are similar with respect to baseline measures. In this study, the baseline *hemoglobin* levels for the two

groups were, $\bar{x}_{itn} = 103.4$ g/L, $s_{itn}=18.2$ g/L and $\bar{x}_{pl} = 99.3$ g/L, $s_{pl}=18.0$ g/L. Do you think the baseline *hemoglobin* levels could explain the difference the *hemoglobin* levels at delivery?

Note 2:
The article mentions that skewed data were log transformed before analysis. Depending on the extent of skewness and outliers, this transformation may be helpful, but not necessary since t-tests are relatively "robust". This means that even if the normality assumption is not met, the tests give valid results when the sample sizes are relatively large.

c) To determine if the treatment is effective, we are interested in "Is the average *hemoglobin* in the treatment group(s) at delivery equal to the average *hemoglobin* in the *placebo group* at delivery in the population?" This is a much more interesting question than the statistical test reported in the article and above in question 1a). We are more interested in "Is the intervention(s) more effective than placebo?" rather than "In the intervention group, are average *hemoglobin* values different comparing value at recruitment and at delivery?". The participants in the intervention group(s) could have improved *hemoglobin* levels, but not as much improved as the *placebo group*. In this case, the intervention would less or equally effective as the placebo.

d) To compare the mean *hemoglobin* levels at delivery for the four groups, investigators could have used ANOVA model (assuming the assumptions were met). Then, if a difference were found, proceed with pairwise comparisons such as in 1c.

2. Consider the following values given for *Gestational Age* at baseline for the two groups *Primigravida* and *Secondigravida*: $\bar{x}_{pr} = 20.8$ weeks, $s_{pr}=3.5$ weeks, $n_{pr}=400$ and $\bar{x}_{sec} = 20.5$ weeks, $s_{sec}=3.8$ weeks, $n_{sec}=352$.

a) Conduct a statistical test to compare the average *Gestational Age* in the *Primigravida* and *Secondigravida* groups in the population. State null and alternative hypothesis, test statistic, and interpret the *p*-value. (Because of rounding, your answer will be slightly different from the value reported in the paper (*p*=0.2).)

b) The table of baseline characteristics in this article [Table 1, p. 279] compares the baseline values for *Primigravida* vs. *Secondigravida* groups, just as we did in 2a). At baseline, what comparisons would be <u>more</u> helpful? Why?

2a) H$_0$: The average *Gestational Age* in the *Primigravida group* is the same as the average *Gestational Age* in the *Secondigravida* group in the population at baseline.
   H$_a$: These two population averages are not the same.

$$H_0: \mu_p = \mu_s \qquad H_a: \mu_p \ne \mu_s$$

Check Assumptions: (See discussion in Q1 above concerning assumptions)
- The observations need to be independent. (Likely are.)
- The sample should be a simple random sample. (It is not.)
- *Gestational age* should be normally distributed in the two groups. (Unknown, however because of sufficient sample size, this assumption

can be relaxed.)

Conduct a two-sample t-test:

$$t = \frac{\bar{x}_p - \bar{x}_s - (\mu_p - \mu_s)}{\sqrt{\frac{s_p^2}{n_p} + \frac{s_s^2}{n_s}}} = \frac{20.8 - 20.5 - 0}{\sqrt{\frac{3.5^2}{400} + \frac{3.5^2}{352}}} = \frac{.3}{.2677} = 1.121$$

$$2 \, Pr(t > 1.121) = 0.26$$

$$df = \min(399, 351) = 351$$
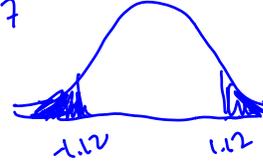
Conclusion:
If the average *Gestational Age* at baseline were different in these two populations (*Primigravida* vs. *Secundigravida*), then the probability we would obtain two sample averages this extreme (20.8 weeks vs. 20.5 weeks) is about 0.26. Our result is not unusual if the population averages are the same. We have insufficient evidence to conclude that *Gestational Ages* are different in the groups.

b) In a table comparing baseline characteristics in a randomized study, we are primarily interested in "Did the randomization work?" In other words, we are interested in whether the four intervention groups are similar at baseline (for factors such as gestational age, % anemic, maternal age, weight, etc). So, it would be helpful to see all these baseline characteristics for each intervention group by parity. We are less interested in the whether the two groups, *Primigravida* and *Secundigravida,* are different (they likely are different at baseline! Else why would we stratify the analysis?) rather we would like to know whether the intervention groups are different at baseline. [The baseline *hemoglobin* levels are given for the different intervention groups by parity at baseline in Table 2 – however, it would be helpful to have other baseline characteristics by intervention group.]

It is important, when viewing baseline characteristics in a randomized study to consider, "Are the differences in the intervention groups important?" (clinically significant) rather than "Are the differences in the intervention groups statistically significant?"

3. At baseline in the *Primigravida group* (*n=400*), 43.5% of women were anemic compared with 28.4% anemic in the *Secondigravida group* (*n=352*). Is there evidence these percents are different are in the populations? State the null and alternative hypothesis, check the assumptions, calculate and interpret the *p*-value. [The journal article reports *p*-value <0.001.]

3.

We will conduct a test for the difference in proportions equal to zero.

Group	P-hat: Anemic	X: Number Anemic	n
Primigravida	0.435	174	400
Secondigravida	0.284	100	352

$$H_0: P_p = P_s \qquad H_a: P_p \neq P_s$$

Pooled phat: $\dfrac{X_p + X_s}{n_p + n_s} = \dfrac{174 + 100}{400 + 352} = \dfrac{274}{752} = 0.364$

$$Z = \dfrac{.435 - .284 - 0}{\sqrt{(.364)(.636)\left(\frac{1}{352} + \frac{1}{400}\right)}} = 4.29$$

p-value: 2*P(*z*>4.29) = 0.0000179

Conclusion:
Assuming the proportions anemic were the same in the two groups (*Primi* vs. *Secondi*) in the populations, the probability that we'd have sample proportions this different (0.435 vs. 0.284) is practically zero. We have strong evidence that the proportions anemic are not the same in the groups.

4. Consider the following results from the (Njagi, 2003) study for *Primigravida* women at delivery:

Intervention Group	Number Anemic (%)	n
ITN+SP	30 (28.6%)	105
ITN	36 (34.6%)	104
SP	30 (30.9%)	97
Placebo	45 (47.9%)	94

a) Construct a two-way table with *Intervention Group* as the row variable and *Anemic/Not Anemic* as the column variable. Conduct a statistical test for any association between *Anemia* and *Intervention Group*. Interpret the *p*-value and include a table of expected values assuming no association.

b) Is there evidence that the percent anemic is different in the *Placebo group* compared to the *Insecticide Treated Net (ITN) group*? (Results are not given in the article.)

c) A colleague suggests the following for testing whether the percent anemic is the same in the *Placebo group* and *ITN group*: H_0: $p_{itn}=0.479$ H_a: $p_{itn} \neq 0.479$,

$$z = \frac{p - p_0}{\sqrt{\frac{p_0(1-p_0)}{n}}} = \frac{0.346 - 0.479}{\sqrt{\frac{0.479(1-0.479)}{104}}}$$. Explain why this test statistic is not appropriate for this case.

d) Is there evidence that the percent anemic is different in the *ITN+SP group* compared to the *SP only group*? In other words, does the addition of the Insecticide Treated Nets to sulfadoxine pyrimethamine give additional benefit? (Results are not given in the article.)

4a) Two-way table of observed values:

Group	Anemic	Not Anemic	Total
ITN+SP	30	75	105
ITN	36	68	104
SP	30	67	97
Placebo	45	49	94
Total	141	259	400

Expected Values:

Group	Anemic	Not Anemic	Total
ITN+SP	37.01	67.99	105
ITN	36.66	67.34	104
SP	34.19	62.81	97
Placebo	33.14	60.87	94
Total	141	259	400

There is sufficient sample size to conduct chi-square test.

H_0: No association between intervention and anemia
H_a: Some association between intervention and anemia

Calculate the test statistic and pvalue:
(By hand:)
 Degrees of freedom = (rows -1) * (columns -1) = (4-1)*(2-1) =3

$$\chi^2 = \frac{(30-37.01)^2}{37.01} + \frac{(75-67.99)^2}{67.99} + \cdots = 9.43$$
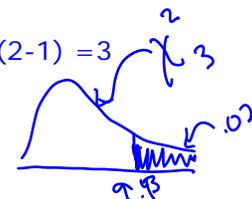
Pr(X^2>9.426) = 0.024

(Or by software:)

Frequency Expected Row Pct	Table of itn by anemic			
		anemic		
itn		an_yes	an_no	Total
itn_sp		30 37.013 28.57	75 67.988 71.43	105
itn_only		36 36.66 34.62	68 67.34 65.38	104
sp_only		30 34.193 30.93	67 62.808 69.07	97
placebo		45 33.135 47.87	49 60.865 52.13	94
Total		141	259	400

(handwritten annotations: "OBSERVED", "EXPECTED", and $\frac{259 \times 105}{400} = 67.988$)

Statistics for Table of itn by anemic

Statistic	DF	Value	Prob
Chi-Square	3	9.4258	0.0241

Conclusion:
If there were no association between the *Intervention Group* and *Anemia*, the probability of observing values as extreme as our sample is pretty small, about 0.02. Therefore we have some evidence that the assumption of "no association" may not be correct.

b) Once you have evidence for "some association" between the variables, we proceed to investigate WHICH groups are different. One interesting question is which Treatment group(s) are different from the *Placebo group*. We will conduct a test for the difference in proportions anemic equal to zero in the *ITN* vs. the *Placebo group*.

Group	Phat (anemic)	X	n
ITN	0.346	36	104
Placebo	0.479	45	94

$$H_0: P_{ITN} = P_P \qquad H_a: P_{ITN} \neq P_P$$

Pooled phat: $\dfrac{X_{ITN} + X_P}{n_{ITN} + n_P} = \dfrac{36 + 45}{104 + 94} = \dfrac{81}{198} = .409$

$$Z = \frac{.346 - .479 - 0}{\sqrt{(.409)(.591)\left(\frac{1}{104} + \frac{1}{94}\right)}} = 1.89$$

(handwritten: N(0,1) DISTRIBUTION, -1.89, 1.89)

p-value: 2*Pr(z<-1.89) = 0.058

Conclusion:
Assuming the proportions anemic were the same in the two groups at delivery

(ITN vs. Placebo) in the populations, the probability that we'd have sample proportions as different as 0.346 and 0.479 is about 0.06. Our result is 'borderline significant'. If the population proportions were the same, our results are somewhat unusual.

c) The suggested null hypothesis and test statistic are used when comparing a single sample proportion, $\hat{p}$, to a <u>constant,</u> p_0. In our example, we are comparing two sample statistics (two $\hat{p}$ values). The estimate for the Placebo group (0.479) is a statistic (it varies depending on the sample), not a parameter (a fixed population value). The Placebo estimate (0.479) is a $\hat{p}$, not a constant, p_0. Therefore, we must use the methods above in 2b) for comparing to two sample proportions (in other words, two $\hat{p}$ values).

d) We will conduct a test for the difference in proportions anemic equal to zero in the *ITN+SP group* vs. the *SP Only group*.

Group	P-hat: anemic	X: Number Anemic	n
ITN +SP	0.286	30	105
SP	0.309	30	97

$$H_0: P_{i6} = P_s \qquad H_a: P_{i6} \neq P_s$$

Pooled phat: $\dfrac{X_{i6} + X_s}{n_{i6} + n_s} = \dfrac{30 + 30}{105 + 97} = \dfrac{60}{202} = 0.297$

$$z = \dfrac{.286 - .309 - \phi}{\sqrt{(.297)(.703)(\frac{1}{97} + \frac{1}{105})}} = 0.366$$
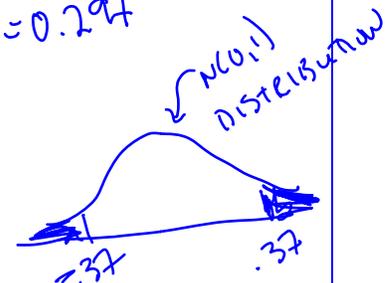
p-value: 2*Pr(*z*<-0.366) = 0.71

<u>Conclusion:</u>
If proportion anemic in the *ITN+SP group* were the same as proportion anemic in the *SP only group* at delivery in the population, we'd expect to observe sample values as different as ours (0.309 vs. 0.286) about 71% of the time just by chance. Our results are not at all unusual if the true proportions anemic are the same in these two populations. We have no reason to suspect the proportions anemic are different in the two groups. The addition of ITN to SP did decrease the proportion anemic (compared to the *SP-Only group*), however, that decrease was not statistically significant.
[Incorrect conclusion: "Since p>0.05, we have evidence the proportions are the same." Or "The probability the proportions are the same in the two groups is about 0.71"]

5. Other studies have investigated the relationship between *Insecticide Treated Nets* and *Anemia* (as well as other outcomes related to malaria during pregnancy) by randomizing villages to either ITN or no net. Discuss the advantages and disadvantages for randomizing villages rather than individuals to the intervention.

- Browne EN, Maude GH, Binka FN. The impact of insecticide treated bednets on malaria and anemia in pregnancy in Kassena-Nankana districts, Ghana: a randomized controlled trial. Tropical Medicine and International Health 2001;6(9) 667-76.
- Shulman CE, Dorman EK, Talisuna AO, Lowe BS, Nevill C et al. A community randomized controlled trial of insecticide treated benets for the prevention of malaria and anemia among prmigravid women on the Kenyan coast. Tropical Medicine and International Health 1998; 3(3)197-204.
- terKuile FO, Terlouw DJ, Phillips-Howard PA, Hawley WA, Friedman JF, Kariuki SK et al Reduction of malaria during pregnancy by permethrin-treated bednets in an area of intense perennial malaria transmission in Kenya. American Journal of Tropical Medicine and Hygiene 2003, 68 Suppl4: 50-60.

> **5.** Logistically, randomizing villages is easier than randomizing individuals. There are fewer randomizations to do. Also it is much easier to distribute the same type of nets to each village rather than different types for the individuals.
>
> Randomizing villages rather than individuals will complicate the analysis however. The methods we have used above (and most? All? Methods in Bios 600) assume that <u>the observations are independent</u>. When villages are the unit of randomization, that assumption no longer is valid. Methods for "clustered" data or "multivariate methods" will be needed. Also a greater sample size will likely be needed to detect a difference in the groups, if villages are the unit of randomization.

6. Consider the following result given in (Marchant 2002). Sixty eight out of 266 *ITN* users tested positive for malaria. Seventy nine out of 239 *non-ITN* users tested positive for malaria (chisquare statistic = 3.4, *p*-value = 0.06, *RR* = 0.77(0.59-1.02))

a) Conduct a statistical test for any association between *ITN use* and *Malaria* using a chi-square test. Include the null and alternative hypothesis. Display the expected values if there were no association. Interpret the *p*-value for a non-statistician.

b) Show how the relative risk and CI were calculated. Interpret the CI for a non-statistician.

c) Conduct a statistical test to compare the proportions of Malaria positive women in the two groups.

d) Explain the relationship between the results in Q6a), b) c)

e) Calculate a CI for the proportion of Malaria positive women in the *ITN user group*. Interpret that CI for a non-statistician. [Results are not given in the paper.]

> **6a)**
> H_0: No association between ITN use and Malaria.
> H_a: Some association between ITN use and Malaria

Frequency Expected Row Pct	Table of itn by malaria			
	malaria			
itn	mal_pos	mal_neg	Total	
itn	68 / 77.43 / 25.56	198 / 188.57 / 74.44	266	
no_itn	79 / 69.57 / 33.05	160 / 169.43 / 66.95	239	
Total	147	358	505	

OBSERVED
EXPECTED

$\frac{266 \times 147}{505}$

Statistics for Table of itn by malaria

Statistic	DF	Value	Prob
Chi-Square	1	3.4229	0.0643

$$\chi^2 = \sum \frac{(OBS - EXP)^2}{EXP} = \frac{(68 - 77.43)^2}{77.43} + \frac{(198 - 188.57)^2}{188.57} + \cdots = 3.42$$

Degrees of freedom = 1
$P(X^2 > 3.42) = 0.064$

Conclusion: If there were no association between *ITN* and *Malaria*, then we'd expect to see observed values as extreme as ours by chance about 6% percent of the time. Our results are borderline unusual if there is no association.

b) $\hat{RR} = \dfrac{68/266}{79/239} = 0.77$

$$95\% \ CI: \ RR \ e^{\pm z \ast \sqrt{\frac{b/a}{a+b} + \frac{d/c}{c+d}}}$$

$$= .77 e^{\pm 1.96 \sqrt{\frac{198/68}{266} + \frac{160/79}{239}}}$$

$$= .77 e^{\pm .2731} = (0.59, 1.02)$$

Suppose we could conduct this study many times - taking many different samples of 505 women and computing a RR for each sample and computing a CI in this way for each of those RRs. About 95% of the CI we compute would contain the true RR parameter- the relative risk in the population between *ITN use* and *Malaria*.

Since 1 is contained in the 95% CI (just barely), we know that a statistical test of the relative risk equal to 1 would have p value >0.05.

Statistics for Table of itn by malaria

Estimates of the Relative Risk (Row1/Row2)			
Type of Study	Value	95% Confidence Limits	
Case-Control (Odds Ratio)	0.6956	0.4731	1.0226
Cohort (Coll Risk)	0.7734	0.5885	1.0163

c) We want to compare the proportions Malaria positive in the two groups (*ITN* vs. *No ITN*).

$$H_0: P_{ITN} = P_{NOTN} \qquad H_a: P_{ITN} \neq P_{NOTN}$$

$$Z = \frac{P_1 - P_2 - (P_{01} - P_{02})}{\sqrt{(P_{pooled})(1 - P_{pooled})\left(\frac{1}{n_1} + \frac{1}{n_2}\right)}} = \frac{.3305 - .2556 - 0}{\sqrt{(.2911)(.7089)\left(\frac{1}{266} + \frac{1}{239}\right)}} = 1.85$$

$2Pr(z>1.85) = 0.064$

Conclusion: Assuming that the proportions of Malaria positive women in the two groups (*ITN* vs . *no ITN*) are the same in the populations, then the probability that we'd see proportions this extreme in our sample (0.33 vs. 0.26) is about 0.06. Our sample values are borderline unusual under the assumption of no difference.

d) The results in questions 6 a)b)&c) all address the same hypothesis of "no association" between the two variables – they just use different statistics. Notice the *p*-values are the same.

e)

$$\hat{p} \pm z^* \sqrt{\frac{\hat{p}(1-\hat{p})}{n}} = .2556 \pm 1.96 \sqrt{\frac{(.2556)(.7444)}{266}} = .2556 \pm .0524$$

$$= (.203, .308)$$

	Column 1 Risk Estimates			
	Risk	ASE	(Asymptotic) 95% Confidence Limits	
Row 1	0.2556	0.0267	0.2032	0.3081
Row 2	0.3305	0.0304	0.2709	0.3902
Total	0.2911	0.0202	0.2515	0.3307
Difference	-0.0749	0.0405	-0.1543	0.0045

If we were to take many samples of *n*=266 women from the population of *ITN* users from this population, and compute a CI for each of these samples, then we'd capture the true proportion of Malaria positive about 95% of the time. We can be reasonably sure that the true % Malaria positive among *ITN* users in this population is between 20% and 31%.

7. Consider the following result from (Marchant 2002).

	Parasite Density		
Group	None	Low	High
ITN User	199	37	30
ITN Non User	161	35	43

   a) If there were no association between *Parasite Density* and *ITN use*, how many women would be expect to be "*ITN User*" and "*Parasite Density: None*"?

   b) Conduct a statistical test for any association between *ITN Use* and *Parasite Density* (3 levels). Include the null and alternative hypothesis. Check the sample size requirements. Display the expected values if there were no association and degrees of freedom. Interpret the *p*-value for a non-statistician. (Rather than the chi-squared test for trend reported in the article, you may use the chi-squared test for any association which is taught in Bios 600 and is more conservative.)

   c) Collapse the categories, *Parasite Density* "*Low*" and "*High*", into one category "*Parasite Positive*". Relabel the "*None*" category to "*Parasite Negative*". Conduct a statistical test to test an association between *ITN Use* (Yes vs. No) and *Parasite* (Negative vs. Positive) [Results are not given in the article.]

---

7. a) Expected value:
Row total * Column total / Overall count = (360 * 266)/505 =189.62 women

b) H_0: No association H_a: Some association
   For sample size requirements, we need the average of the expected counts is 5 or more and the smallest expected count is 1 or more. These sample size requirements are met.
   Degrees of freedom = (rows-1) * (columns -1) = 1*2 =2.

   By hand:

$$\chi^2 = \frac{(199-189.62)^2}{189.62} + \frac{(37-37.93)^2}{37.93} + \cdots \doteq 4.92$$
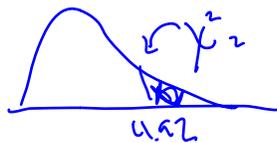
$$Pr(\chi^2 > 4.92) = 0.084$$

Or by software:

Global Activity Four
Malaria and Anemia

The FREQ Procedure

Frequency Expected Row Pct	Table of itn by parasite			
		parasite		
itn	para_no	para_lo	para_hi	Total
itn	199 189.62 74.81	37 37.925 13.91	30 38.451 11.28	266
no_itn	161 170.38 67.36	35 34.075 14.64	43 34.549 17.99	239
Total	360	72	73	505

Statistics for Table of itn by parasite

Statistic	DF	Value	Prob
Chi-Square	2	4.9523	0.0841

If there were no association between *ITN use* (Yes vs. No) and *Parasite* (None, Low, High) then the probability that we'd observe a table of values as extreme as in this study is about 0.08. Our results are not very unusual if there is truly no association in the population.

c)
Sample size requirements are met to conduct a chi-square test for a 2x2 table, with all four the expected counts greater than or equal to 5.

Frequency Expected Row Pct	Table of itn by parasite		
		parasite	
itn	para_neg	para_pos	Total
itn	199 189.62 74.81	67 76.376 25.19	266
no_itn	161 170.38 67.36	78 68.624 32.64	239
Total	360	145	505

OBS
EXP

BY HAND

$$x^2 = \frac{(199-189.62)^2}{189.62} + \frac{(67-76.376)^2}{76.376}$$

$$\cdots = 3.4118$$

$$P_R = (x^2 > 3.4118) = 0.06$$

Statistics for Table of itn by parasite

Statistic	DF	Value	Prob
Chi-Square	1	3.4118	0.0647

If there were no association between *ITN use* and *Parasite Status* (Positive vs. Negative) then the probability that we would have a table of observed values as extreme as in this study is about 0.06. Our results are somewhat unusual if there is no association.

8. Discuss the relative strength of the two studies (Njagi, 2003, Marchant 2002) used in this activity. Summarize the advantages and disadvantages.

---

8. a)

In general, randomized controlled trials such as (Njagi, 2003) are considered stronger than observational studies (Marchant, 2002). In a randomized study, the patients will be very similar at baseline across the intervention groups with respect to any potentially confounding factors. So if randomization has "worked", any differences in the groups at the end of the study will be due to the intervention itself. In an observational study, the *ITN group* may be different than the *non-ITN group* in ways other than the exposure. While it is possible to adjust for factors that differ between the *ITN* and *non-ITN groups* (such as age, education, parity, etc.) in an observational study, we can only do so if those factors are known, can be measured, and can be measured accurately.   Randomization has the advantage of being able to equally distribute factors both known and unknown. Randomized studies have the disadvantage of being more expensive (in general) and taking longer to conduct.

Observational studies have the advantage of being less expensive.

Observational studies may sometimes be the only option, for example when randomization is considered unethical because of potential risk associated with the exposure/intervention.