# BIOSTATISTICS 600

## Global Activity Three

## Acupuncture and Correction of Breech Presentation

## ANSWER KEY

## INTRODUCTION

An inexpensive, low-risk procedure to correct breech presentation would have large public health impact, particularly in populations with limited access to skilled health care for complicated deliveries.

Toward the end of gestation, most babies turn to cephalic position (head down). Breech position (head up and bottom or feet down) is the most common type of non-cephalic presentation. Delivery of a baby in the breech position increases the risks of complications for both the mother and the baby. For the baby, delivery in breech position increases the risk of physical injury. For the mother, breech position delivery increases the risk of a cesarean section delivery which in turn increases the risk of surgical complications. Other downsides for C-section deliveries include greater recovery time compared to vaginal deliveries, and greater cost. Also medical care and facilities may not be available for a C-section delivery in some areas.

The method of "moxibustion" has been practiced in traditional Chinese medicine "since ancient times" to correct breech presentation. Moxibustion is the application of heat (by burning herbs) next to an acupuncture point to stimulate fetal movement. Investigators hypothesize that one way to get a baby to move from breech position to cephalic position, is to get the baby just to move more. Increasing spontaneous fetal movement is one of the goals of this acupuncture intervention – the increased movement then could result in the baby moving into the preferable cephalic position.

In a study conducted in Jiangxi Province, China (Cardini 1998) 260 pregnant women whose fetus was in breech position were randomized to either moxibustion or usual care. Fetal activity and breech presentation at birth were among the outcome measures of interest. The intervention group had significantly more fetal movement (48.5 fetal movements in the invention group compared to 35.4 fetal movements in the control group, $p<0.001$). The moxibustion group also had significantly higher proportion of cephalic presentation at birth (98/130 in the intervention group vs. 81/130 in the control group, $p=0.02$). An ongoing study in Spain (Vas, 2008) will investigate the relationship between this acupuncture method and cephalic presentation by randomizing pregnant women with a fetus in breech presentation to one of three groups: real moxibustion, sham moxibustion, or usual care.

In this activity, students will reproduce many of the results from the original article (Cardini 1998) including discussing the assumptions and several issues involved in the study being conducted outside the US. In addition, students will be asked to calculate and interpret results from a hypothetical third arm of a trial (such as in the design by Vas, 2008).

## SOURCES

A.
Cardini F, Weixin H. 1998. Moxibustion for Correction of Breech Presentation: A Randomized Controlled Trial. JAMA 280(18):1580-4.

B.
Vas J, Arnanda JM, Baron M, Perea-Milla E, Mendez C, Ramirez C, Aguilar I, Modesto M, Lara AM, Martos F, Garcia-Ruiz AJ. 2008. Correcting non-cephalic presentation with moxibustion: study protocol for a multi-centre randomized controlled trial in general practice. BMC Complementary and Alternative Medicine 8:22.
http://www.biomedcentral.com/1472-6882/8/22

C.
Glickman SW, McHutchinson JG, Peterson ED, Cairns CB, Harrington RA, Califf RM, Schulman KA, Ethical and Scientific Implications of the Globalization of Clinical Research. N Engl J Med 2009;360:816-23.

*Terminology*:
ECV, external cephalic version, is a procedure performed by a health care provider to attempt to reposition the baby into cephalic position. The mother is given muscle relaxants, and the health care provider applies pressure to try to turn the baby around.
Primigravidas are women pregnant for the first time.

## QUESTIONS

**1.** The first trial (Cardini, 1998) was conducted was China and the second trial in (Vas, 2008) is being conducted in Spain. Acupuncture has been practiced in China "since ancient times", and is less common in Western countries. What impact may attitudes and familiarity towards alternative medicine play in the results of these studies?

**1.**
Participants in the trial conducted in China may be more familiar with the true moxibustion. So a design such as (Vas, 2008) would likely be more difficult in China (rather than Spain or the US) because participants randomized to the "sham moxibustion" could be aware that the treatment was not at the correct acupuncture point.
Recruitment may be more difficult in areas less trustful of alternative medicine.
Also, differences in attitudes toward acupuncture could influence recruitment in the two studies, which could affect the generalizability of the study. Since the mechanism by which this acupuncture method may work is unclear, the results may differ due to the psychological as well as physiologic effects of the intervention.

**2**. Investigate rates of cesarean section births in China, Spain, and the US (using internet resources as well as the articles). Then investigate rates of non-cephalic presentation (including breech presentation) in the three countries. If moxibustion is associated with decreased rates of cesarean section, compare how the three countries may be impacted by the use of moxibustion to increase cephalic presentation at term? Include aspects such as economics, and health outcomes for mother and child.

**2.** China:

Cesarean sections rates in China are rising particularly in urban areas. For example, the cesarean section rate in Shantou, China in 1997 was 29.9%. This was a large increase compared to the 1990 rate of 11.1% in a period in which the number of births declined. Breech presentation was the indication in approximately 18% of the c-section births.

Investigators hypothesize that the "one child rule" has been a contributing factor to the rise of c-section births as some women request c-section deliveries for convenience or in low-risk cases.

- Wu WL. Cesarean delivery in Shantou, 2000. China:a retrospective analysis on 1922 women. Birth, 27(2):86-90.
- Sufang G, Padmadas SS, Fenmin A, Brown JJ, Stones RW. 2007. Delivery settings and caesarean section rates in China. Bulletin of the World Health Organization 85(10):755-762.

Spain:

The incidence of breech presentation in Spain was estimated at 3.8% and ECV is relatively uncommon.

- Breech presentation in Spain, 1992: a collaborative study European Journal of Obstetrics & Gynecology and Reproductive Biology, 62(1):19-24.

US:

The C-Section rate in the US is approximately 20%. The rates of c-sections for breech presentation vary by state –for example, in California one study (Gilbert 2003) reports that 96.1% of breech presentations births were C-sections.

General:

Breech presentation occurs in 3-4% of births. Some common known risk factors for breech presentation are previous breech pregnancy, lack of prenatal care or insufficient prenatal care, prematurity, and low birth weight. Rates of vaginal birth for breech presentation varies widely by country from Japan 56% to Canada <5.0%. In most countries, rates of c-sections births have risen in the last 20 years. Despite efforts made to reduce the proportion of c-section births, experts disagree about recommendations for indications.

- Hickok DE, Gordon D, Milber, JA, Williams, MA, Daling JR. 1992. The Frequency of Breech Presentation by Gestational Age at Birth: A large Population-Based Study. American Journal of Obstetrics and Gynecology 166:851-52.
- Koike T, Minakami, H, Sasaki M, Tamada T, Sato I. 1995. The Problem of Relating Fetal Outcome with Breech Presentation to Mode of Delivery. Archives of Gynecology and Obstetrics 258:119-23.
- Gilbert WM, Hicks SM, Boe NM, Danielson B. 2003. Vaginal Versus Cesarean Delivery for Breech Presentation in California: A Population-Based Study. Obstetrics and Gynecology 102:911-17.
- ican-online.org/pregnancy/breech-presentation-fact-sheet.

Discussion:
The issue of the risks and benefits of c-section birth for breech presentation is hotly debated. The rates of c-section birth for breech presentation vary widely in the countries where these studies were conducted as well as developing other countries that could benefit from the research.
An effective intervention for breech presentation would have enormous economic and public health impact in preventing complications from breech delivery or from c-section deliveries. Consider developing countries where c-section deliveries are less common but riskier, and mortality and morbidity from breech presentation are more common - this intervention could reduce injury of the neonate by increasing the percentage of cephalic births. In more developed countries, the impact would be both economic (fewer c-sections) and health related (fewer breech vaginal births and fewer breech c-sections).

**3.** If you were part of an Institutional Review Board reviewing these trials, what questions would you have for the investigators in the two studies? Discuss the ethical issues with conducting these clinical trials outside the US – in particular with the patients randomized to placebo.

**3.**

- Is this study necessary? (or has the hypothesis been sufficiently answered by previous research)
- Is there representation by the host country on the IRB? Among the investigators?
- What are the risks of the intervention? Of the placebo?
- Explain the informed consent process – how will investigators ensure the participants understand the risks and benefits of participation and for not participating? Are country specific practices taken into consideration? What are the risks (for participants and for non-participants) and how will the risks be explained?
- Will this research benefit the host country?
- What is the "standard of care" for pregnant women in this country with breech presentation? How will "standard of care" be ensured for women in the placebo/control group?
- What will patients be given to participate in the study and is this "coercive" *in this population*?
- Do the investigators have any conflicts of interests? Disclose funding sources and sponsors of the study.
- Will the results be reported, regardless of the outcome? How will the results be reported?

**4.** Consider the following statements from the results section of the abstract in article (A. Cardini 1998):

> "The intervention group experienced a mean of 48.45 fetal movements vs. 35.35 in the control group ($p$<.001; 95% CI for difference, 10.56- 15.60)".

This result is also discussed on page 1583. Discuss this result by addressing these questions:

**a)** State the null hypothesis and alternative hypothesis. State the interpretation of this *p*-value. In other words, *p* is a probability… what is it the probability of? Use words a medical professional with no statistical training will understand. Notice that zero is not in the confidence interval – explain how is this CI related to the *p*-value.

**4a)** Given: $\bar{x}_{trt}$ = 48.45 fetal movements

$\bar{x}_{cntl}$ = 35.35 fetal movements

Difference in sample averages = 13.1

$p$<0.001 for the difference; CI for the difference = (10.56 – 15.60)

$H_0: \mu_T = \mu_C$ or $H_0: \mu_T - \mu_C = 0$

$H_a: \mu_T \neq \mu_C$ EQUIVALENTLY $H_a: \mu_T - \mu_C \neq 0$

Null hypothesis: the average number of *Fetal Movements* in the *Treatment* population is the same as the average number of *Fetal Movements* in the *Control* population

Alternative hypothesis: the average number of *Fetal Movements* in the *Treatment* population is NOT the same as the average number of *Fetal Movements* in the *Control* population

The article doesn't explicitly mention whether the alternative is one-sided or two-sided, but we know the default is two-sided. We'll assume they used a two-sided test. If they used a one-sided test, they should have mentioned it and would need to defend that choice.

INTERPRETATION: If the difference in the mean number of *Fetal Movements* in the *Treatment* and *Control* populations is zero, we would expect the difference in the sample averages also to be small. But the difference in the sample values of mean number of *Fetal Movements* is 48 (sample mean for the treatment group) minus 35 (sample mean for the control group) which is 13 *Fetal Movements*. Assuming that the true difference in the population is zero, getting a sample difference of 13 fetal movements is very unusual. We'd expect to get a value this extreme (or more extreme) by chance less than 0.1% of the time. We have strong evidence that the difference in the averages in the populations is NOT zero. The averages in the treatment population and control population are likely different.

Since zero is not in the confidence interval (in fact very far outside the CI) we know that zero is not a likely value for the true difference in average number of fetal movements. Since zero is not in the 95% confidence interval, we also know the *p* value for testing the difference in averages equal to zero will be less than 0.05.

**b)**

- The article states that this result was from a '*t* test'. Which would be appropriate for this analysis: a two-sample *t* test, a one sample *t* test, a *z* test or a matched pairs *t* test. Why?
- Can you reproduce these values (the *p*-value and the CI) from the information given in the paper? If so, do it. If not, why not?
- What are the assumptions to conduct this *t* test, and do you think those assumptions are met? Explain.

**4b)** The investigators used a two sample *t*- test. Since they are interested in whether two averages are equal (in the two populations) and had sample standard deviations (*s* values not σ values), a two sample - test is appropriate (assuming assumptions are met).

We can't reproduce this result because we don't have the *s* values, the sample standard deviation for each group (and we don't have the original data). If we had the original data or the sample standard deviations for each group, we could reproduce the *t*-test.

For a two sample *t* –test, we assume that the *Treatment* sample is a simple random sample from a normal population with mean $\mu_1$ and that the control sample is an independent simple random sample from a normal population with mean $\mu_2$. The assumption of independent observations is likely true (for example, there are not likely two sisters in the samples). We don't know if *Fetal Movements* are normally distributed, but the *t*-test is relatively robust, meaning that even if this assumption is not met we can still use the procedure if the sample size is sufficiently large. Our sample size in this problem is sufficiently large.

**c)** Consider a similar hypothetical study of *Factor X* (similar to moxibustion). You can think of it as third arm of the trial- another type of intervention, such as "sham moxibustion". Address this question: "Are the mean number of *Fetal Movements* the same in the *Factor X* group and the *Control* group in the population?". Assume the mean number of *Fetal Movements* for the *Factor X* group is 36.79 movements with s.d. 7.42 and $n_T=130$. The mean number of *Fetal Movements* for the *Control* group is still 35.35 and $n_C=130$ (and assume sample s.d.=7.85). Compute the *p*-value for the difference in active *Fetal Movements* in the two groups. As usual, interpret the conclusion precisely in language understandable to a nonstatistician. (In other words, you should be able to explain the *p*-value without using terminology such as 'reject/ accept the null hypothesis'…).

**4c)**

$\bar{x}_T = 36.79$ $\quad \bar{x}_C = 35.35$ $\quad H_0: \mu_T - \mu_C = 0$

$s_T = 7.42$ $\quad s_C = 7.85$ $\quad H_a: \mu_T - \mu_C \neq 0$

$$t = \frac{\bar{x}_T - \bar{x}_C - (\mu_T - \mu_C)}{\sqrt{\frac{s_T^2}{n_T} + \frac{s_C^2}{n_C}}} = \frac{36.79 - 35.35 - 0}{\sqrt{\frac{7.42^2}{130} + \frac{7.85^2}{130}}} = \frac{1.44}{0.947} = 1.52$$

Conducting a two sample t-test for the difference in means gives us a test statistic, $t = 1.52$. Using t tables, the two-tailed $p$ value is between 0.1 and 0.2. Using statistical software or calculator, we get a more exact p-value, p = 0.13.

Assuming that the mean number of *Fetal Movements* is the same for the *Treatment* group (*Factor X*) and the *Control* group, we'd expect sample values to be this far apart (or more far apart) about 10% of the time. So if the population means are the same in the populations, our result is not very unusual. We don't have much evidence to reject the assumption that the average number of *Fetal Movements* in two groups is the same.

**d)** Compute a 95 % confidence interval for the difference in active *Fetal Movements* in the *Factor X Group* and the *Control Group* using the information provided in part c. Interpret the confidence interval in language a nonstatistician can understand. Do you think *Factor X* is effective in producing a different number of *Fetal Movements*, compared to the *Control group*? Why?

**4d)**
CI: 1.44 +/- 1.98 SEd = 1.44 +/- 1.98 (0.987) = ( -0.44, 3.32)
(Get the $t^*$ value = 1.98 from the tables ( with $df$ = 129 or conservatively row = smaller df = 100) or from software. )

The confidence interval is ( -0.44, 3.32). This is a CI for the difference *in average number of Fetal Movements between the Treatment and the Control* group in the populations. We know 95% of all CI computed in this way contain the true difference in averages. So the true difference in averages is likely between Negative 0.44 and 3.32. In particular, from the CI, we can see the "zero" is one of the many likely values for the difference in average fetal movements.
[You can't just say, "Because 0 is in the CI, we conclude there is no difference in the average number of fetal movements."!]

Notice the 95% confidence interval contains 0 and the $p$ value is > 0.05. We have insufficient evidence to suspect that the average number of fetal movements is not the same.

Our $p$-value being greater than 0.05 could mean two things: there is a difference in average *Fetal Movements* but we just didn't detect it (either because the difference is small or because the sample size was too small or because we happened to get a sample by chance where the averages were close together) or there is no difference in the average number *of Fetal Movements* in the population.
[You can't just say, "There is no difference in average number *of Fetal Movements*".]

**5.** Consider the following statement from the results section of the abstract:

"….98 (75.4%) of 130 fetuses in the intervention group were cephalic at birth vs. 81 (62.3%) of the 130 fetuses in the control group ($p$=.02; RR, 1.21; 95% CI 1.02-1.43)".

This information is also in Table 2 in the article. Discuss this result by addressing these questions:

**a)** Construct a 2x2 table for this data. Test the hypothesis that there is no association between the *Treatment* and the baby being *Cephalic* at birth. Explain/interpret your answer so that a nonstatistician could understand it.

**5a)**

OBSERVED

| | CEPH | NO | |
|---|---|---|---|
| T | 98 | 32 | 130 |
| C | 81 | 49 | 130 |
| | 179 | 81 | |

EXPECTED

| | CEPH | NO | |
|---|---|---|---|
| T | 89.5 | 40.5 | 130 |
| C | 89.5 | 40.5 | 130 |
| | 179 | 81.0 | |

$$\chi^2 = \frac{(98-89.5)^2}{89.5} + \frac{(32-40.5)^2}{40.5} + \frac{(81-89.5)^2}{89.5} + \frac{(49-40.5)^2}{40.5} = 5.18$$

DF=1 Look up $\chi^2_1 \Rightarrow .02 < p < .025$

$H_o$: No association between *Treatment* status and *Cephalic* birth status
$H_a$: Some association between *Treatment* status and *Cephalic* birth status

If there were no association between *Treatment* status (*treatment* vs. *control*) and *Cephalic* position, then we'd expect results this extreme less than 2.5% of the time just by chance. We have some evidence of an association.
In other words, if there is no association then our results are somewhat unusual which leads us to believe that there may be an association.

[More explanation: Two things could have happened here. The first possibility is that there is no association between the *Cephalic* birth and *Treatment* in the population, and we just happened to select a sample that was extreme. Our sample values, by chance, are somewhat different than we would expect if there is no association. The second possibility is that there is an association. We don't know which of these two possibilities is true.
An old fashioned interpretation of this result is "reject the null hypotheses and conclude that there no association"… however this interpretation is **not** preferred. You should understand why people unfortunately sometimes still use the terminology (because it is easy to remember and concrete/black and white) and why it is not preferred. ]

**b)** Investigators are interested in whether the <u>proportion</u> of *Cephalic* presentations at birth is the same for the two groups.

- Compute the 'risk of cephalic birth' for the *Treatment* group, the 'risk of cephalic birth' for the *Control* group
- Compute the 'risk difference'
- Test whether the risk difference is equal to zero in the population . (State the null and alternative hypotheses, the test statistic, precisely interpret the *p*value in words a nonstatistician could understand.)
- Discuss the use a one-sided alternative vs. a two-sided alternative. (For example, could the use of a one-sided alternative be defended for this scenario, in your opinion?)

**5b)**

$\hat{p}_T = \frac{98}{130} = 0.754 = 75.4\%$ $\quad H_0: p_T = p_C$

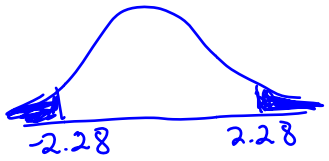
$p_C = \frac{81}{130} = 0.623 = 62.3\%$ $\quad H_a: p_T \neq p_C$

$SE_{\hat{p}_T - \hat{p}_C} = \sqrt{\hat{p}(1-\hat{p})\left(\frac{1}{130} + \frac{1}{130}\right)} = 0.0575$

where $\hat{p} = \frac{179}{260} = 0.688$

$Z = \frac{\hat{p}_T - \hat{p}_C}{SE_{\hat{p}_T - \hat{p}_C}} = \frac{0.754 - 0.623}{0.0575} = 2.28$

$p = 2PR(z > 2.28) = 0.0223$




If the two proportions of *Cephalic* deliveries were the same in the two populations (*Treatment* and *Control*), we'd expect sample values this extreme (or more extreme) about 2% of the time. We have some evidence that the proportion of 'head down' births is not the same in the two populations.

A one-sided alternative would <u>not</u> be recommended for this example. This decision is made before the data are collected, and it was possible to have the *Treatment* with a lower proportion of cephalic births. The two-sided alternative is conservative and always recommended and unless a very strong case can be made for the one-sided alternative.

**c)** Compute a 95% CI for the difference in the proportions of *Cephalic* presentations ( the risk difference computed in b)) . Interpret the CI for a nonstatistician.

**5c)**

$$SE_{\hat{p}_1-\hat{p}_2} = \sqrt{\frac{(.75)(.25)+(.62)(.38)}{130}} = 0.0575$$

$$(\hat{p}_1-\hat{p}_2) \pm z^* SE_{\hat{p}_1-\hat{p}_2} = 0.131 \pm 1.96(0.0575) = (0.02, 0.24)$$

We know that if repeatedly took samples and computed a confidence interval associated with each sample as in the problem, about 95% of the CI would contain the true difference in proportions in the population. So, the true difference in the proportion of cephalic deliveries is likely between (0.02,0.24).

There are two possibilities: We have calculated one of the approximately 95% of CIs that contain the true difference in proportions. OR we happened to have calculated one of the approximately 5% of CI that don't contain the true difference in proportions.

Since the confidence interval doesn't contain zero, the difference in proportions is likely not zero in the population. So the proportions of cephalic births in the *Treatment* group and the *Control* group are likely different in the populations.

You can't say, "The true difference in proportions is between (0.02 and 0.24) about 95% of the time". The true difference in proportions in the population is fixed. The true population parameter doesn't bounce around and land between (0.02 and 0.24) 95% of the time. It is the confidence intervals that vary from sample to sample, but the parameter is a constant, unknown value.

Can you say "We are 95% confident that the true difference in proportions is between 0.02 and 0.24"? Many people use this phrasing, but it is not preferred in my opinion. The problem is, what does it really mean to be "95% confident" -? I don't know! This interpretation is certainly not as clear as some interpretations. It suggests (to me) that the CI is fixed and the population mean is variable (which is not correct).
Being "95% confident" has an ambiguous meaning to me. For example:

* If this phrasing means (to you) "95% of the time the true difference in proportions is between 0.02 and 0.24" then this interpretation is NOT correct.
* If this phrasing means (to you) " CI's computed in this way capture the true difference in proportions about 95% of the time, so we are 95% confident in the procedure to produce a CI that contains the difference in proportions in the population.", then the interpretation is correct. So, saying "We are 95% confident that ….." is not necessarily wrong... but is certainly not as clear as other interpretations. And such an interpretation may lead one to lose sight of what a CI really means.

**d)** The article states that the relative risk (*RR*) of cephalic presentation at birth for the intervention group is 1.21. Interpret this value for a nonstatistician. Compute a 95% Confidence Interval for the Relative Risk (Given in the article as 1.02-1.43). Why might we be interested in the Relative Risk (*RR*) rather than the risk difference (RD as in b) above)?

**5d)** RR= $\frac{98/130}{81/130} = \frac{.754}{.623} = 1.21$

| | Cphal | NO | |
|---|---|---|---|
| T | 98 | 32 | 130 |
| C | 81 | 49 | 130 |

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

$= 1.2099 e^{\pm 1.96 \sqrt{(\frac{32}{98}/130 + \frac{49}{81}/130)}}$

$= 1.2099 e^{\pm 1.96 (.084647)}$

$= (1.02, 1.43)$

Patients in the acupuncture *Treatment* group were 1.21 times as likely to have the preferable cephalic birth compared to those in the *Control* group. In other words, the *Treatment* group was 21% more likely to have a cephalic presentation compared to the *Control* group. The confidence interval for the RR is (1.02, 1.43).
[Correct interpretations: "The true relative risk is likely between 1.02 and 1.43." "About 95% of CI computed in this way contain the true RR, so we can be reasonably confident the true population RR is between 1.02 and 1.43". Incorrect interpretation: "The RR will be between 1.02 and 1.43 about 95% of the time." ]

The relative risk measures STRENGTH of association, while risk difference measures public health IMPACT. In our example, we are probably more interested in assessing the relationship and knowing the strength of the association, so RR is probably a reasonable measure to use.

Risk difference may be preferable if we were more interested in "how many cases of breech birth could be prevented if we use this treatment?" Often, therefore, RR is used in the beginning of research to see if there is a relationship between two factors… and Risk Difference is used later, after an association has been established, to show the potential impact of treatment.

**e)** Suppose there were another intervention group, who received *Factor X* (perhaps a sham moxibustion). Suppose that "….90 (69.2%) of 130 fetuses in the intervention *Factor X* group were cephalic at birth vs. 81 (62.3%) of the 130 fetuses in the *Control* group" Investigators are interested in whether the proportion of cephalic presentations at birth is the same for these two groups.

- Compute the 'risk of cephalic birth' for the *Factor X* treatment group, the 'risk of cephalic birth' for the *Control* group, and then compute the 'risk difference'.
- Test whether the risk difference is equal to zero in the population. (State the null and alternative hypotheses, the test statistic, precisely interpret the *p*value in words a nonstatistician could understand.)
- Do you think that women in the population who receive *Factor X* have a different proportion of *Cephalic* births compared to women in the *Control* group?

**5e)**

$\hat{p}_T = 90/130 = 0.692$

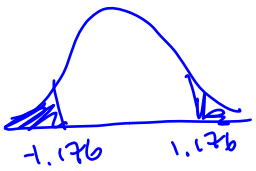
$\hat{p}_C = 81/130 = 0.623$

$H_0: p_T = p_C \qquad H_a: p_T \neq p_C$

Pooled $\hat{p} = \frac{171}{260} = .658$

$SE_{pp} = \sqrt{\hat{p}(1-\hat{p})\left(\frac{1}{n_1}+\frac{1}{n_2}\right)} = \sqrt{(.658)(1-.658)\left(\frac{1}{130}+\frac{1}{130}\right)} = 0.05885$

$Z = \frac{\hat{p}_T - \hat{p}_C - (p_T - p_C)}{SE_{pp}} = \frac{.692-.623-0}{.05885} = 1.176$




p-value $= 2P(z > 1.176) = 0.238$

If there were no difference in the proportion of cephalic births in the *Factor X Treatment* group and the *Control* group populations, then we'd expect sample proportions this different ( 0.692 and 0.623 ) about 24% of the time. Our sample results are not unusual if the population proportions are the same. So we have no reason to suspect that the assumption "no difference" in the proportion of head-down birth in the two groups is not true.
We don't know if the proportions of cephalic birth in the two groups are the same. Our results are not unusual if the proportions are the same in the two groups. However, there may be a small difference in the proportions in the two populations, and we just were unable to detect it.

INCORRECT: Since p> 0.05, we can conclude the proportions in the groups are the same. OR The probability that that proportions are the same is 0.238.