

1. **OPERA Study, 2012 *JAMA*.** Postoperative atrial fibrillation or flutter (AF) is one of the most common complications of cardiac surgery and significantly increases morbidity and health care utilization. A total of 1516 patients scheduled for cardiac surgery were randomized to receive fish oil or placebo. The main outcome was occurrence of postoperative AF lasting longer than 30 seconds. The primary end point occurred in 233 (30.7%) patients assigned to placebo and 227 (30.0%) assigned to fish oil (odds ratio, 0.96 [95% CI, 0.77-1.20]; $p=0.74$).

- (a) What are the null and alternative hypotheses?
- (b) How do you think they tested the null hypothesis?
- (c) How do you interpret their results?

2. **Exercise and Nicotine Craving Trial, 2012 *Journal of Smoking Cessation*.**

Cigarette cravings are one of the most often expressed difficulties related to quitting. The effect of acute exercise on craving and withdrawal symptoms during a pharmacological based smoking cessation intervention is unknown. Methods: Participants completed the Shiffman-Jarvik scale (which provides a continuous measure of craving intensity) immediately before and after a scheduled exercise session. Significant reductions in cigarette craving were demonstrated following the exercise session.

- (a) What are the null and alternative hypotheses?
- (b) How do you think they tested the null hypothesis?
- (c) How do you interpret their results?

3. **Birth Defects in Fallujah, Al-Sabbak et al, 2012.** In late summer, the world media widely reported results of a scientific study that “showed levels of lead in hair were 5 times higher” in hair samples of children with birth defects compared to levels in children without birth defects. The paper also made this statement and provided data that of 44 children with birth defects, the mean lead level was 56454 $\mu\text{g}/\text{kg}$ with a standard deviation of 217,705 $\mu\text{g}/\text{kg}$; of 11 children without birth defects, the mean lead level was 11277 $\mu\text{g}/\text{kg}$ with a standard deviation of 27,781 $\mu\text{g}/\text{kg}$. Is the statement reported in the media justified? Why or why not?

4. **Urinary Incontinence, 2012 *NEJM*.** Urgency urinary incontinence is characterized by unpredictable loss of urine; it is a prevalent condition that occurs disproportionately in women, affecting up to 19% of older women in the United States. Duke researchers conducted a randomized trial comparing an oral drug regimen with a single botox injection into bladder muscle to assess the reduction in episodes of urgency urinary incontinence over the course of 6 months, improvement in quality of life, and side effects. Patients were randomly assigned to the drug or botox in a 1:1 ratio. Of 249 women who underwent randomization, 247 were treated, and 241 had data available for the primary outcome analyses. The mean reduction in episodes of urgency urinary incontinence per day over the course of 6 months, from a baseline average of 5.0 per day, was 3.4 in the drug group and 3.3 in the botox group ($P=0.81$). Complete resolution of urgency urinary incontinence was reported by 13% and 27% of the women, respectively ($P=0.003$). Quality of life improved in both groups, without significant between-group differences. The drug group had a higher rate of dry mouth (46% vs. 31%, $P=0.02$) but lower rates of catheter use at 2 months (0% vs. 5%, $P=0.01$) and urinary tract infections (13% vs. 33%, $P<0.001$).

- (a) How many hypothesis tests are described in these results?
- (b) Consider the outcome of mean reduction in episodes of urgency urinary incontinence per day over the course of 6 months.
 - i. What are the null and alternative hypotheses?
 - ii. How do you think they tested the null hypothesis?
 - iii. How do you interpret their results?
- (c) Consider the outcome of complete resolution of urgency urinary incontinence.
 - i. What are the null and alternative hypotheses?
 - ii. How do you think they tested the null hypothesis?
 - iii. How do you interpret their results?

5. **Propecia and Depression, Irwig et al, 2012.** A researcher interviewed 61 users of the hair-loss drug Propecia and 29 men with male pattern baldness who were nonusers, finding that 39 of the users had moderate to severe symptoms of depression and that none of the non-users had moderate to severe depression symptoms. The researcher published the results, reporting a statistically significant association between Propecia use and depression.

- (a) What are the null and alternative hypotheses?
- (b) How do you think they tested the null hypothesis?
- (c) How do you interpret their results?

6. **Success in Graduate School.** A researcher is interested in factors related to success in graduate school. Using data from a variety of sources, the researcher obtains a continuous measure of graduate school success y , the outcome of interest. In addition, the researcher has continuous measures of intellectual ability x_1 and of work ethic x_2 . All three of these variables are *standardized* to have mean zero and standard deviation 1. This means that a predictor effect β can be interpreted as the expected increase in y in standard deviation units for a 1 standard deviation increase in x . All variables are coded so that higher values are better (i.e., more success, smarter, better work ethic). The researcher fits the model

$$y_i = \beta_0 + \beta_1 x_{1i} + \beta_2 x_{2i} + \varepsilon_i.$$

The parameter estimates are $\hat{\beta}_0 = 0.00$, $\hat{\beta}_1 = 0.608$, $\hat{\beta}_2 = 0.614$. The p-values for testing $H_{01} : \beta_1 = 0$ versus $H_{A1} : \beta_1 \neq 0$ and for testing $H_{02} : \beta_2 = 0$ versus $H_{A2} : \beta_2 \neq 0$ are both < 0.001 . Describe the relationship between intellectual ability, work ethic, and graduate school success in language that could be used in a journal publication.

7. **Blood Pressure and Age.** Consider the following quote from a research project from a basic statistics class. “Age was significantly positively correlated with systolic blood pressure; each one-year increase in age was associated with a 1.22 mm Hg increase in systolic blood pressure ($p < 0.05$).”

Suppose the researcher based this statement on a linear regression model. Which 95% CI is most consistent with the quote?

- (a) (-0.17, 2.62)
- (b) (0.80, 1.43)
- (c) (0.80, 1.63)
- (d) (1.01, 1.45)

8. **Toxicology.** A laboratory experiment compared the relative potencies of four cardiac substances. In the experiment, a suitable dilution of each substance was slowly infused into an anesthetized guinea pig, and the dosage at which the pig died was recorded. Ten guinea pigs were used for each substance for a total of 40 guinea pigs. The main research goal was to determine whether any differences existed among the potencies of the four substances, and, if so, to quantify those differences. The researchers used the ANOVA model $y_{ij} = \mu_i + \varepsilon_{ij}$, $i = 1, 2, 3, 4$, $j = 1, \dots, 10$ to address this research question.

What is H_0 for testing whether any differences exist among the potencies of the four substances?

Descriptive statistics are below.

Substance	Parameter	N	\bar{y}_i	s_i^2
1	μ_1	10	25.9	9.4
2	μ_2	10	22.2	12.2
3	μ_3	10	20.0	8.7
4	μ_4	10	19.6	8.7

The F statistic corresponding to the overall F test is 8.545 ($p < 0.001$). Does the p-value provide evidence of any difference between the groups, or not?

In order to investigate differences between species, the investigators tested differences between each pair of substances and have provided the following table.

Parameter	Estimate	95% CI
$\mu_1 - \mu_2$	3.7	(-0.2, 7.6)
$\mu_1 - \mu_3$	5.9	(2.0, 9.8)
$\mu_1 - \mu_4$	6.3	(2.4, 10.2)
$\mu_2 - \mu_3$	2.2	(-1.7, 6.1)
$\mu_2 - \mu_4$	2.6	(-1.3, 6.5)
$\mu_3 - \mu_4$	0.4	(-3.5, 4.3)

Using this table, describe differences in potencies using language suitable for journal publication.

