

somatic research

BY RAVENSARA S. TRAVILLIAN



THE RESULTS SECTION

By getting more familiar with the statistics commonly used in research studies, you'll grow more comfortable with the conventions articles use to present study outcomes.



In a well-written research article, the groundwork for the Results section was laid in the Introduction and Methods sections—the researcher explained what outcomes the team was looking for, why they were looking for them, and how they would detect them. Those outcomes are reported in the Results section.

In the Results section, the researchers report their results without interpretation. Later on, in the Discussion section, they will discuss their interpretation of those results and their meaning—but in the Results section itself, they simply lay out those results, uninterpreted, for everyone else to see and draw their own conclusions.

We are now far enough along in the discussion of massage research that you will see some things you can already recognize from previous articles and there will be other things we have not yet discussed. I'll provide some tips to help illustrate unfamiliar concepts, without presenting them in as much detail as the essential concepts. However, when reading articles on your own without additional guidance, a good strategy is to follow as much of the article as you can understand, without sweating the stuff you haven't been exposed to yet. Get what you can from the articles, and trust that you are laying the foundation for more learning. Be patient with your learning process! Don't let the parts you are not yet familiar with discourage you.

REPORTING STATISTICAL SIGNIFICANCE IN THE RESULTS SECTION

Our first example looks at a study by Hernandez-Reif et al. on the effects of massage on mood and emotion in children with cystic fibrosis and their parents. In cystic fibrosis, a life-threatening genetic lung disease, mucus, fluid secretions, and infections build up and block the airways of affected children. In addition to the physical effects of the disease, its serious implications for continued quality of life for the affected child can create a state of chronic stress for caregivers.

In this example, STAI, STAIC, and POMS are questionnaires used to measure parents' and children's self-reports of their anxiety and mood states, as explained in the article's Methods section. The F-values analyze complicated mixtures or interactions of cause and effect. F-values of 6.28, 5.73, 7.30, and 8.04 are considered large; large F-values provide increased confidence in the results reported.

For the STAI and STAIC, group by session interaction effects for parents in the massage group, $F(1, 17) = 6.28, p < .05$, and their children, $F(1, 17) = 5.73, p < .05$, revealed a reduction in anxiety following the session on the first day for the parents and for the children ... For the POMS, a group by session interaction effect, $F(1, 17) = 7.30, p < .05$, revealed improved mood for the massaged children following the first and last day sessions ... A group by days interaction effect, $F(1, 17) = 8.04, p < .05$, revealed an increase in peak air flow readings for the massage therapy group by day 30 of the study.¹

Hernandez-Reif's team reports that the measure of parents' anxiety showed a reduction after the first day's session, the measure of the children's mood showed an improvement after the first and last days' sessions, and the measure of the children's peak air flow readings showed an increase after the 30th day of the study. All of these measurements have a *p*-value of less than .05, or 5%, indicating that there is less than a 5% probability of a false positive due to these effects occurring by chance. By the level established in the Methods section, these effects are considered statistically significant.

In that same Methods section, she also explains that peak air flow measures the amount of air that flows through larger airways, and is a measure of effort on the children's part. An improvement in that measurement, then, should indicate that the children feel better and are making more

effort—an interpretation she will return to in the Discussion section, but which she fastidiously avoids here where she reports only the raw results. In the Discussion section, and in the title of the article, her interpretation centers on the idea of “benefit” to the children and their parents—but here, she lets the numbers speak for themselves.

APPLYING NEW KNOWLEDGE

In a study on the effects of trigger-point massage on chronic tension headaches, Christopher Quinn and his team report several different types of results. The term *mean*, used in the results, is one of several different ways statistics can calculate averages for a group. (Notice that where Hernandez-Reif put the two numbers after an F in parentheses, Quinn uses subscript notation. The important point is not how the author typesets the two numbers, but the two numbers themselves, indicating that F is a ratio, or comparison of two numbers standing for the variation within a single group, and the variation among different groups.)

A repeated measures analysis of variance indicated a reliable change over time for number of headaches per week ($F_{7,21} = 3.69, P = .009$). Post hoc analysis confirmed that headache frequency was significantly lower during the weeks of massage treatment (weeks 5-8) than during the baseline weeks (weeks 1-4). A reduction in number of headaches per week was noted for all subjects within the first week of massage treatment. Additionally, the mean [average] number of headaches per week was reduced from 6.8 during baseline to 2.0 during the treatment period.²

All of the participants in the study experienced fewer headaches during the treatment weeks of the study than they did in the baseline weeks before the treatment began. While those results were about the results of frequency of headaches (how often they occurred), the team also looked at the duration of the headaches, or how long they lasted, over the course of the study.



The term *mean*, used in the results, is one of several different ways statistics can calculate averages for a group.



Research Terminology

ESSENTIAL TERMS

± (plus or minus): Indicates a range of values for results. If massage reduces average headache duration from 8.0 ± 4.3 hours to 4.3 ± 2.3 hours, that means the pre-massage headaches lasted anywhere from 12.3 hours to 3.7 hours, and the post-massage headaches lasted anywhere from 6.6 to 2.0 hours.

IMRaD: Mnemonic device for the functional sections of a research article: Introduction, Methods, Results, and Discussion.

Mean: A way to calculate an average, derived by adding all the numeric values of the results together, and dividing that sum by the number of results.

Null hypothesis: The hypothesis, or proposed explanation, that there is no difference between a treatment group and a non-treatment group in a study.

p: The probability that, if the null hypothesis is true, a study would produce particular results through chance alone, and researchers would then mistake a false positive result for a real effect.

Power: The probability that a study will detect an effect that is truly present and will not make a false negative error by mistaking the failure to detect a real effect for a true absence of any effect. Expressed as an amount between 0 and 1, meaning from 0% to 100%; for example power = 0.8 means that a real effect was detected 80% of the time.

Protocol: A standardized procedure or formula for a treatment during a study so that everyone receiving a treatment receives essentially the same treatment.

Sample size: Number of participants (n) in a research study.

Statistical significance: The threshold at which the researcher determines that the study has passed the test of not making a false positive error and mistaking an effect due to chance for a real effect, judged by the value of p for the results. Often, but not always, $p < 0.05$, or detecting an effect that is not really there less than 5% of the time.

SUPPLEMENTARY TERMS

F(number1, number2): F-value is a ratio between the variation in a group, and variation across groups, associated with performing an analysis of variance, or ANOVA. A high F-value, which often means a value greater than 1.0, indicates a low p , and thus increased confidence in the results being due to a real effect, rather than just chance.

Post-hoc analysis: a method of analyzing data and looking for patterns that occurs *after* (“post”) the conclusion of the study, rather than specifying in advance what the researchers *will be* looking for once the study begins.

Repeated measures analysis of variance: A type of analysis of variance (ANOVA), which in general are a set of statistical methods to analyze complicated mixtures of conditions and cause and effect by clustering results into components that are associated with different factors involved.

The mean duration of headaches during the baseline period was compared with the mean duration during the treatment period ... For the group as a whole, mean headache duration decreased from $8.0 \pm$ [plus or minus] 4.3 hours during the baseline period to 4.3 ± 2.3 hours during the treatment phase, and the duration of each headache also decreased for each subject. Because there were only 4 subjects in the study, however, the P value of this measure did not achieve statistical significance ($P=.058$).²

All of the participants experienced shorter headaches during the study individually, and in addition, the group's average length of headache duration decreased as well. However, as we discussed in this space in the last issue, because the study had a very small sample size, it had insufficient power to reliably detect a treatment effect. Even though the researchers observed a trend, with a *p*-value of .058 or 5.8%, that trend was not statistically significant, because it was greater than 5%.

Finally, the team examined the effects of their massage protocol [procedure] on the intensity of headaches the study participants experienced:

The intensity of the most severe headache that occurred each day was marked on a visual analog scale ranging from 0 to 100 mm (with 100 mm indicating the most severe headache). The mean headache intensities during the baseline period and the treatment period are compared in Figure 3. The effect of massage therapy on headache intensity was subject dependent. Two subjects showed marked improvement, one subject showed mild improvement, and one subject showed an increase in headache intensity. For the group as a whole, changes between the baseline period (44.2 ± 2.2 mm) and the treatment period (35.8 ± 8.2 mm) were not significant for headache



intensity as measured by a visual analog scale ($P = .19$ [19%, or almost 1 time out of 5 that you would see this result by chance]).³

USING STATISTICS, UNDERSTANDING RESULTS

These results are fairly straightforward to read in plain language—many articles will present their results that way, while others may use more complicated statistics, numbers, and tables. By practicing and getting more familiar with the statistics commonly used in research studies and the way they are usually presented, you'll start seeing things you recognize, and you'll grow more comfortable with the conventions articles use to present study outcomes. And that's a good thing, because the Results section is often full of tables, charts, and graphs used to present the study's results. In the next article, we'll talk about charts and other graphics—how to build on your developing skills to read them, to understand them, and to interpret them. **m&b**

6 Ravensara S. Travillian is a massage practitioner and biomedical informatician in Seattle, Washington. She has practiced massage at the former Refugee Clinic at Harborview Medical Center and in private practice. In addition to teaching research methods in massage since 1996, she is the author of an upcoming book on research literacy in massage. Contact her at researching.massage@gmail.com with questions and comments.

NOTES

1. M. Hernandez-Reif, Field T, Krasnegor J, Martinez E, Schwartzman M, and K Mavunda, "Children with cystic fibrosis benefit from massage therapy," *Journal of Pediatric Psychology* 24, no. 2 (April 1999): 175–81.
2. C. Quinn, Chandler C, and A Moraska, "Massage therapy and frequency of chronic tension headaches," *American Journal of Public Health* 91, no. 10 (October 2002): 1657–61.
3. *Ibid.*