HOW TO READ A SCIENTIFIC STUDY

Examine.com

HOW TO READ A SCIENTIFIC STUDY

Marketing claims backed by so-called scientific evidence pervade the health and fitness industry. Supplement manufacturers sell compounds like **raspberry ketones** as if their effects were as proven as **creatine**'s. Sometimes, following the paper trail of a marketing claim does lead to a real, published study. Unfortunately, not all studies are created equal. To navigate the vast world of health and fitness marketing, to avoid wasting money on ineffective products, it is vital to be able to assess the credibility of a study.

To understand a study, as well as how it relates to previous research on the topic, you need to read more than just the abstract. Context is very important when discussing new research, which is why abstracts that promise too-good-to-be-true results are often misleading.

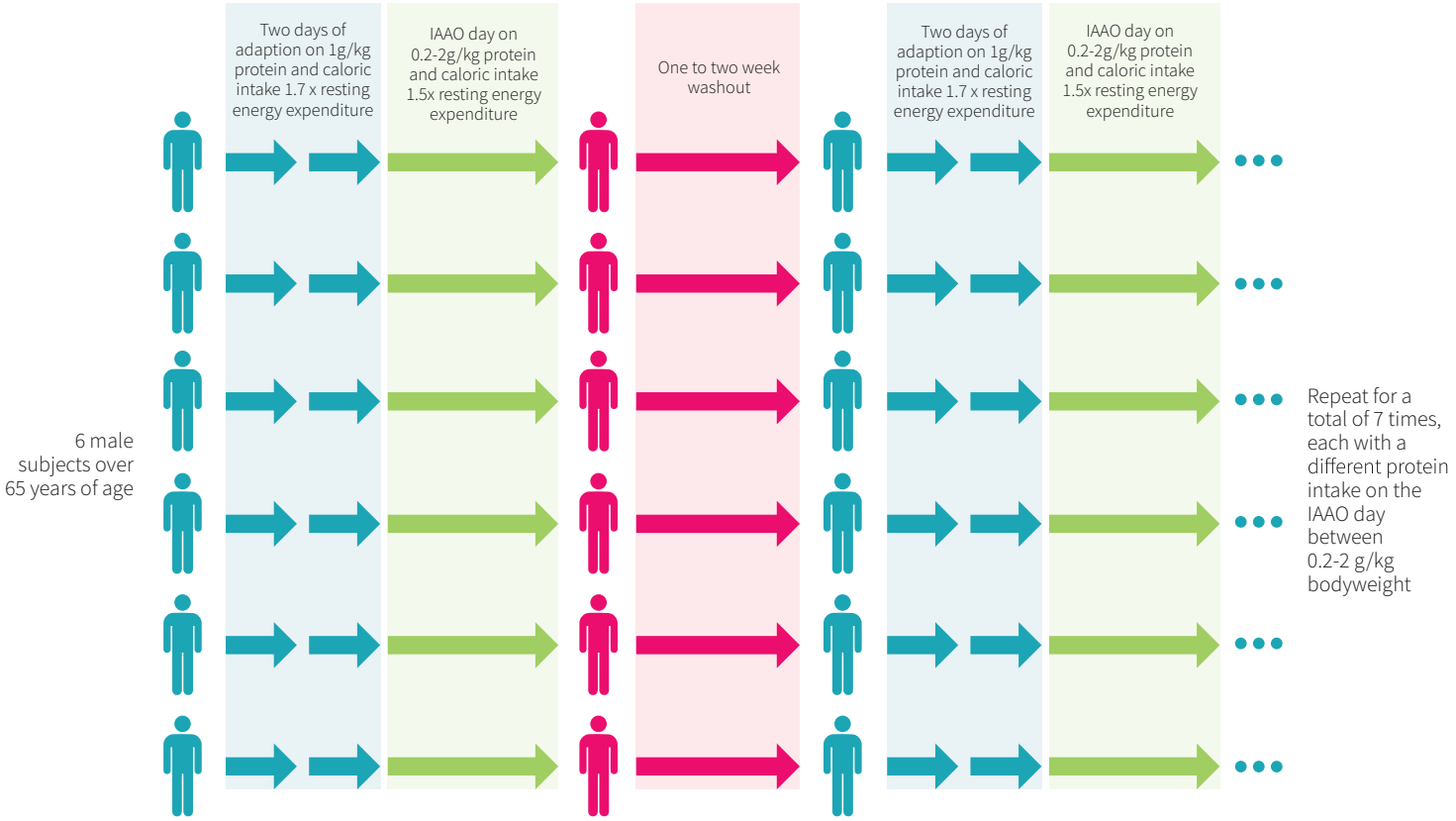
Understanding the Abstract and Introduction

An abstract is a brief summary that covers the main points of a study. Since there's a lot of information to pack into a few paragraphs, an abstract can be unintentionally misleading. Because it does not provide context, an abstract does not make clear the mechanisms of an experiment or its applicability. Before citing a study as evidence in a discussion, make sure to read the whole thing, because it might turn out to be weak evidence.

The introduction sets the stage. The authors usually summarize previous related research and explain why they chose to investigate further. For example, the compound HMB was found to increase bone mineral density and skeletal muscle mass when supplemented by the elderly, so researchers set out to determine if it also worked in younger people. Introductions are often a great place to find additional reading material, since the authors will frequently reference other relevant, published studies.

The Most Important Part of the Study: Methodology

Figure 1: Example protocol for a protein study



Scrutinizing the methodology section of a study (often called "Methods" or "Materials and Methods") is vital for determining the strength of the evidence and its degree of applicability. This section contains demographic information like age, sex, health status, and lifestyles of the participants, how they were recruited, and details about the intervention itself. Ideally, the methodology is presented in such detail that other researchers can repeat the study without needing to contact the authors of the paper.

The demographic information can be lengthy, you might be tempted to skip it, yet it impacts both the reliability of the study (the larger the sample size, the more reliable the results) and its applicability. In health and fitness, applicability means that a compound that is useful for one person may be a waste of money — or worse, a danger — for another. For example, while **creatine** is widely recognized as safe and effective, there are "nonresponders" for whom this supplement fails to significantly improve exercise performance.

The Most Important Part of the Study: Methodology

The creatine example shows that each individual is different, yet the demographic information works as an indicator of applicability. If a trial only recruited women, for instance, men reading the study should keep in mind that its results are less likely to apply to them. Likewise, a study performed on college students may yield different results when performed on subjects recruited from a retirement facility. How the participants were recruited also matters: Different methods will attract different demographics, which can influence the outcome of the study.

As a last piece of demographic information, the methodology section will usually detail what factors excluded someone from a study, and why. These "exclusion factors" help ensure that potential variables that may confound the results are eliminated from the test sample.

The actual design of the study is another important part of the methodology section. The gold standard is a randomized, double-blind, placebo-controlled trial. The participants are randomly distributed between the intervention group and the control group, and neither they nor the researchers administering the experiment know who belongs to which group.

Figure 2: Some trials are sex-specific

Other study variants include single-blind trials, in which only the participants don't know if they're receiving a placebo, and observational studies, in which researchers observe a population and take measurements. Keep in mind that an observational study cannot show causation, since the scientists conducting the study are not controlling any variables.

Finally, primary and secondary outcomes can also be discussed in the methodology section, to make clear the endpoints the researchers will be looking at.

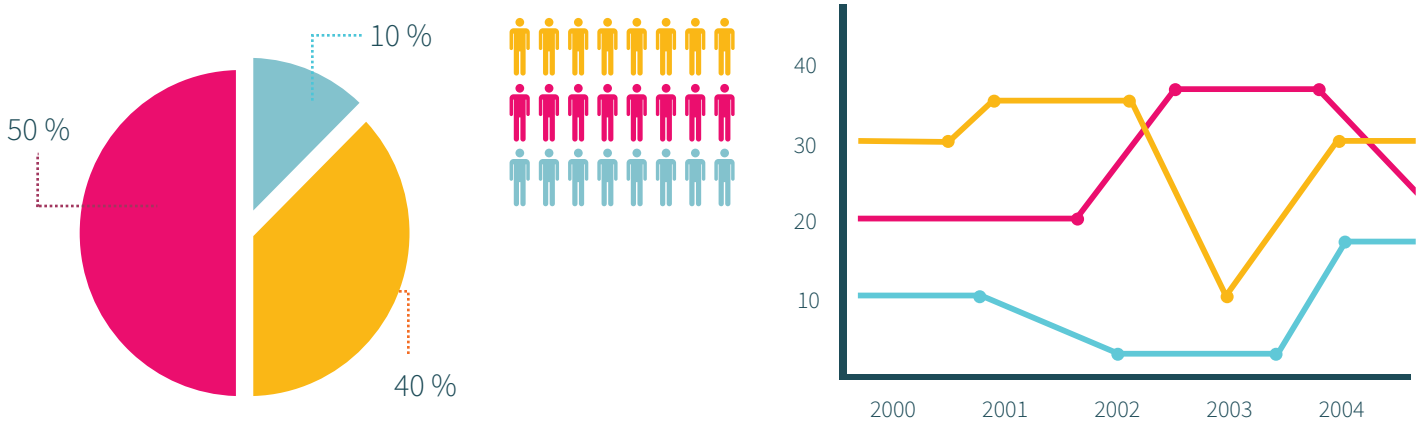
Interpreting Statistics and P-values

The methodology section usually concludes with a hearty statistics discussion. Determining whether an appropriate statistical analysis was used is an entire field of study, so when reading statistics, try to focus on the big picture.

Researchers run a statistical analysis on the results of their study in order to determine if those results are significant. A p-value, which stands for probability value, tells you how likely a given result would be if the null hypothesis were true. (The null hypothesis is usually along the lines of "XYZ supplement/intervention does not work any better than the control/placebo".)

In most studies, the cutoff for statistical significance is 0.05. A p-value of 0.0482 means there is a 4.8% chance a given result is compatible with the null hypothesis, while a 0.000001 (0.0001%) value means there's only one chance in a million — both numbers indicate statistically significant results, but the smaller p-value warrants greater confidence in the findings.

Figure 3: Graphically impressive results don't equate to significant results



Note that a statistically significant result can be clinically insignificant. Let's say researchers find that daily supplementation of a certain compound helps people burn an extra half pound of fat every year; if the p-value is 0.0003, the result is reliable … but the half-pound is probably too small to justify supplementing the compound.

Finally, keep in mind that, while important, p-values aren't the final say on significance. While the statistical analysis accounts for the variables the researchers were controlling, p-values can also be influenced by confounding variables. When evaluating the strength of a study's design, imagine yourself in the researcher's shoes and consider what flaws a study might have outside of what was written down in the paper.

HOW TO READ A SCIENTIFIC STUDY

Reading the Results

Researchers discuss the primary outcome, or what they were most interested in investigating, in the results section (sometimes the "Results and Discussion" section). Skipping right to it after reading the abstract might be tempting, but that often leads to misinterpretation and the spread of misinformation. Never read the results without reading the methodology section first; knowing how researchers arrived at a conclusion is as important as the conclusion itself.

One of the first things to look for in the results discussion is a comparison of characteristics between the tested groups. Big differences in baseline characteristics after randomization may mean the two groups are not truly comparable. Researchers also have to report dropout and compliance rates. Life frequently gets in the way of science, so almost every study has its share of participants that didn't finish the trial or failed to follow the instructions; too great a proportion of dropouts or non-compliant participants should raise a red flag, though, especially if they are significantly more numerous in one group than the other(s).

Scientists use questionnaires, blood panels, and other methods of gathering data, all of which are displayed here through charts, graphs, and figures. Be sure to examine the y-axis of any graph or figure to see what scale the results are being represented on; what may seem like a large change on the graph may in fact be very minor.

The results section can also include a secondary analysis, such as looking specifically at the results for a subgroup (like women only, or only those participants over age 65) or doing a sensitivity analysis (which checks if the results stay the same when some of the data are not included). While subgroup analyses can be interesting, seldom is a trial large enough for statistically strong conclusions to be drawn from them. As for sensitivity analyses, they can reinforce or further impair the confidence one can place in the results of studies that include suspect data, and so are potentially very important.

Clarifying the Conclusion

The conclusion (or sometimes the “Discussion”, if detached from the “Results”) wraps up the paper. Here, the researchers expound the value of their work. If they think it necessary, they may also clarify their interpretation of the results they obtained. They will often compare their study to previous ones on the same topic, as well as discuss potential experiments based on the new results: They might hypothesize a mechanism of action, or point out ways future trials could improve on their own design.

Conflicts of interest, if they exist, are sometimes disclosed after the conclusion. Studies funded by companies that stand to benefit from positive findings aren’t automatically invalid, but they do deserve extra scrutiny.

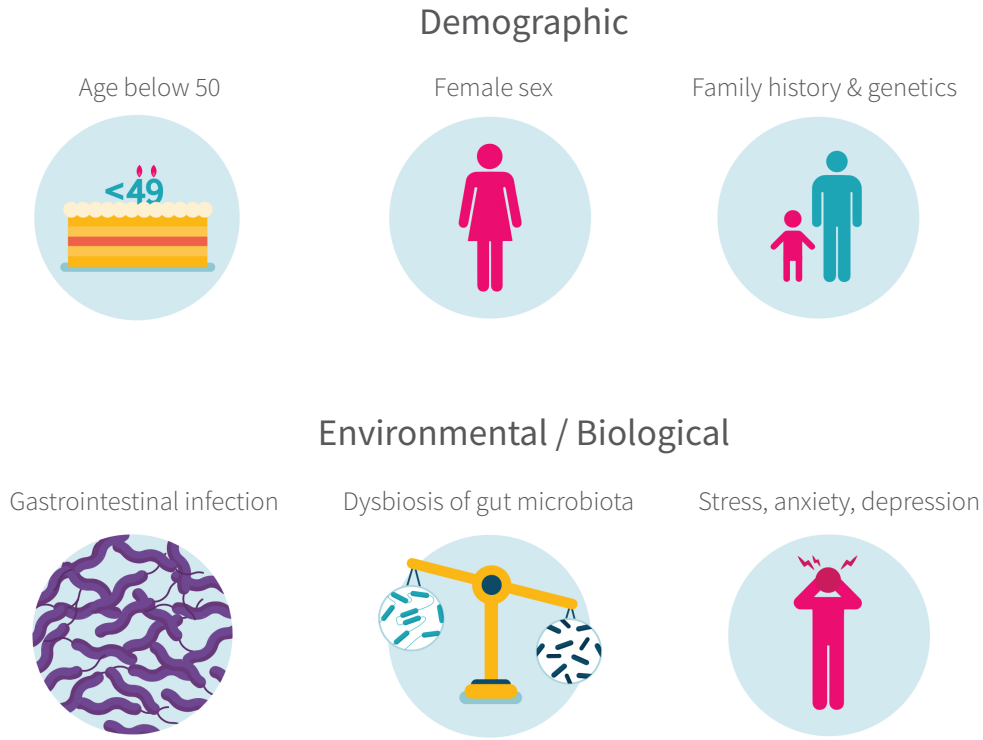
Like the introduction, the conclusion provides valuable context and insight. If it sounds like the researchers are extrapolating to populations beyond the scope of their study, or are overstating its results, don’t be afraid to read it again (especially the methodology section).

But we’re not done yet....

Digging Down to the Truth

Just because a study is published doesn't mean the results apply to everyone. For example, the first studies on **glutamine** were conducted on burn victims, who are deficient in this amino acid due to their injury. People who are not deficient in glutamine would not experience the same benefits as burn victims.

Figure 4: Many factors can influence applicability



Intentionally selecting demographics so as to maximize the effects of an intervention makes sense for hospital researchers who are looking for a way to help a specific kind of patient. For other reasons, this strategy is also popular in weight loss trials, which is why it isn't uncommon for new "fat burners" to be supported by studies that only recruited obese postmenopausal women. When this type of information is left out of the abstract and then journalists skip the methodology section, people end up misled.

Never assume the media have read the entire study. A recent survey assessing **the quality of the evidence for dietary advice given in UK national newspapers** found that between 69% and 72% of health claims were based on deficient or insufficient evidence. To meet deadlines, overworked journalists frequently rely on study press releases, which often fail to accurately summarize the studies' findings. That's why we have a dozen researchers review each trial we analyze for the Examine.com Research Digest (ERD). With human health being such a complex topic, there really is no substitute for reading the full text of the paper, often more than once!