



HOW TO READ A SCIENTIFIC STUDY

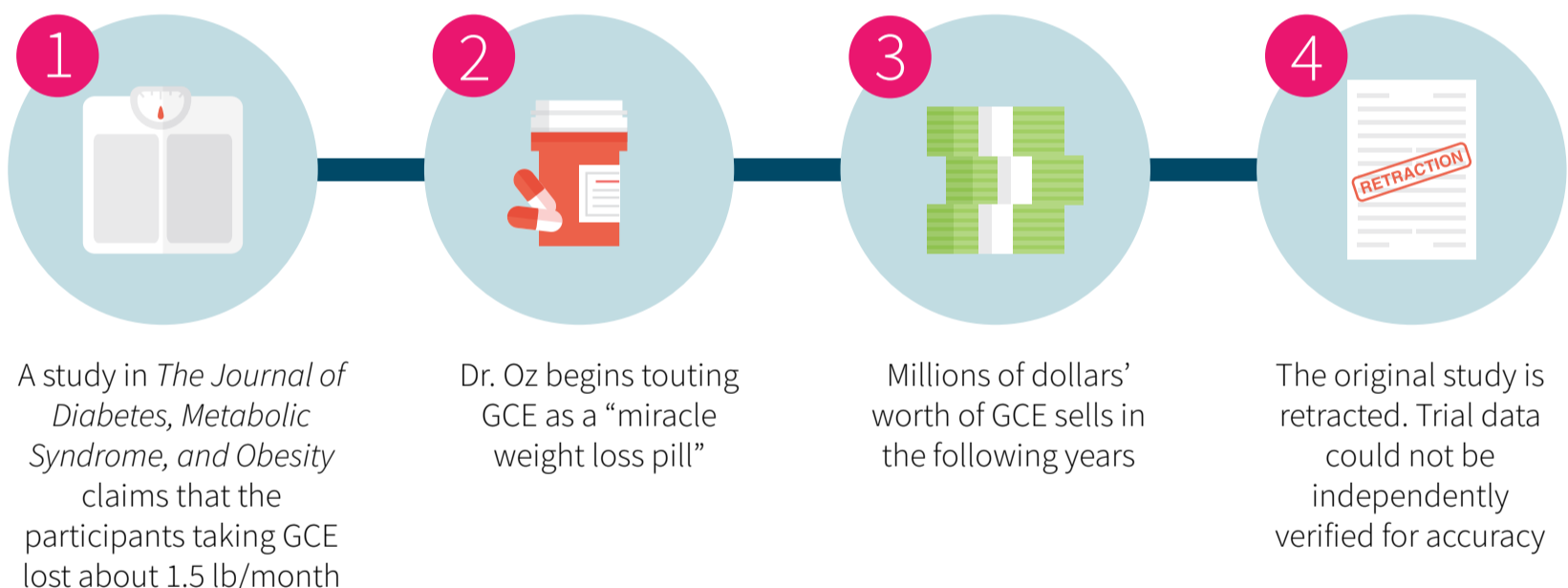
(and why it's pretty hard)



Why learn to read a scientific study?

Marketing claims backed by “scientific evidence” pervade the health and fitness industry. Supplement manufacturers sell compounds like [green coffee extract](#) (on which there is barely any human research) as if their effects were as well-established as those of [creatine](#) (on which there are hundreds of human trials). Sometimes, following the paper trail of a marketing claim does lead to a real, published study — but not all studies are created equal. To avoid wasting money on ineffective products, you need to be able to assess different aspects of a study, such as its credibility, its applicability, and the clinical relevance of the effects reported.

Figure 1: Green coffee extract: a cautionary tale



Poorly conducted trials can lead to exceptional results. It is usually best to wait and see if those results can be replicated before drawing a conclusion.

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 critically important when discussing new research, which is why abstracts are often misleading.

Types of studies

There exist numerous types of studies. This guide was designed to help you better understand all of them, with a special emphasis on experimental research.

Figure 2: Overview of study types

STUDY TYPES		DESIGN	STRENGTH	WEAKNESS
EVIDENCE SUMMARIES	Meta-analysis	Combs through all available literature on a topic to pool and analyze data	Can provide greater statistical power to a research question	Time consuming and requires advanced statistical knowledge
	Systematic review	Presents an expert review of the available evidence on a given topic	Can provide guidance in areas of limited research	Design differences can make comparing studies difficult
EXPERIMENTAL STUDIES	Randomized controlled trial (RCT)	Participants are randomly assigned to either an intervention group or a control group	Randomization can help eliminate population bias in sample	Can be very expensive and resource intensive
	Nonrandomized controlled trial	Participants are assigned to either an intervention group or a control group	Can blind participants to treatment	Not randomized
OBSERVATIONAL STUDIES	Cohort study	Follows a group to track habits and risk factors over time	Can be easier to conduct than an RCT	Can take years to conduct
	Case control study	Compares histories of groups with and without a specific disease or health outcome	Helps identify potential risk factors	Often confounded by recall bias
	Case report	Provides a detailed account of individual cases	Helps identify new trends	Not generalizable

Randomized, double-blind, placebo-controlled trials are commonly seen as the gold standard of biomedical research. In such trials, the participants are randomly assigned to either an intervention group (which will receive the intervention) or a control group (which will receive a placebo), and neither they nor the researchers running the experiment know which participants belong to which group.



Understanding the abstract and introduction

A paper is divided into sections. Those sections vary between papers, but they usually include an abstract, an introduction, a section on methods (which provides demographic information, presents the design of the study, and sometimes expounds on the chosen endpoints), and a conclusion (which is often split between “results” and “discussion”).

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 often make clear the limitations of an experiment or how applicable the results are to the real world. Before citing a study as evidence in a discussion, make sure to **read the whole paper**, because it might turn out to be weak evidence.

The introduction sets the stage. It should clearly identify the research question the authors hope to answer with their study. Here, the authors usually summarize previous related research and explain why they decided to investigate further.

For example, the non-caloric sweetener [stevia](#) showed promise as a way to help improve [blood sugar control](#), particularly in diabetics. So, researchers set out to conduct larger, more rigorous trials to determine if stevia could be an effective treatment for diabetes. Introductions are often a great place to find additional reading material since the authors will frequently reference previous, relevant, published studies.

One study is just one piece of the puzzle

Reading several studies on a given topic will provide you with more information — more data — even if you don’t know how to run a meta-analysis. For instance, if you read only one study that looked at the effect of [creatine on testosterone](#) and it found an increase, then 100% of your data says that creatine increases testosterone; but if you read ten studies that looked at the effect of creatine on testosterone and only one found an increase, then 90% of your data says that creatine does *not* increase testosterone.

(This is a simplified example, in which we used “vote-counting”: we compared the number of studies that found an effect to the number of studies that found no effect. Meta-analyses, however, are *a lot* more complicated than that: they must take into account various criteria, such as the design of the study, the number of participants, and the biases that affect the results, rather than reduce each study to a positive or negative result.)

Unsurprisingly, it is common for supplement companies to cherry-pick studies. If a company wants to sell you creatine as a testosterone booster, they’ll mention the one study that found an increase in testosterone, not the nine that found no increase.

Likewise, it is usually easy for warring camps to throw studies at each other to “prove” their point. If you seek one study that shows that a low-fat diet is better than a low-carb diet to promote weight loss, you’ll find one. If you seek one study that shows the contrary, you’ll find one too. It is therefore important, if you seek the truth (and not just some ammunition for a Twitter brawl) to look at the whole body of evidence, and to consider fairly the studies that don’t agree with your original opinion (if you have one, but most of us do).

On that note, keep in mind that companies are not alone in cherry-picking studies. Researchers sometimes do it too. If you know that a field is contentious yet a paper only mentions studies that support the authors’ conclusions, you may want to [make your own search for other papers on the topic](#) (always a good idea at any rate).

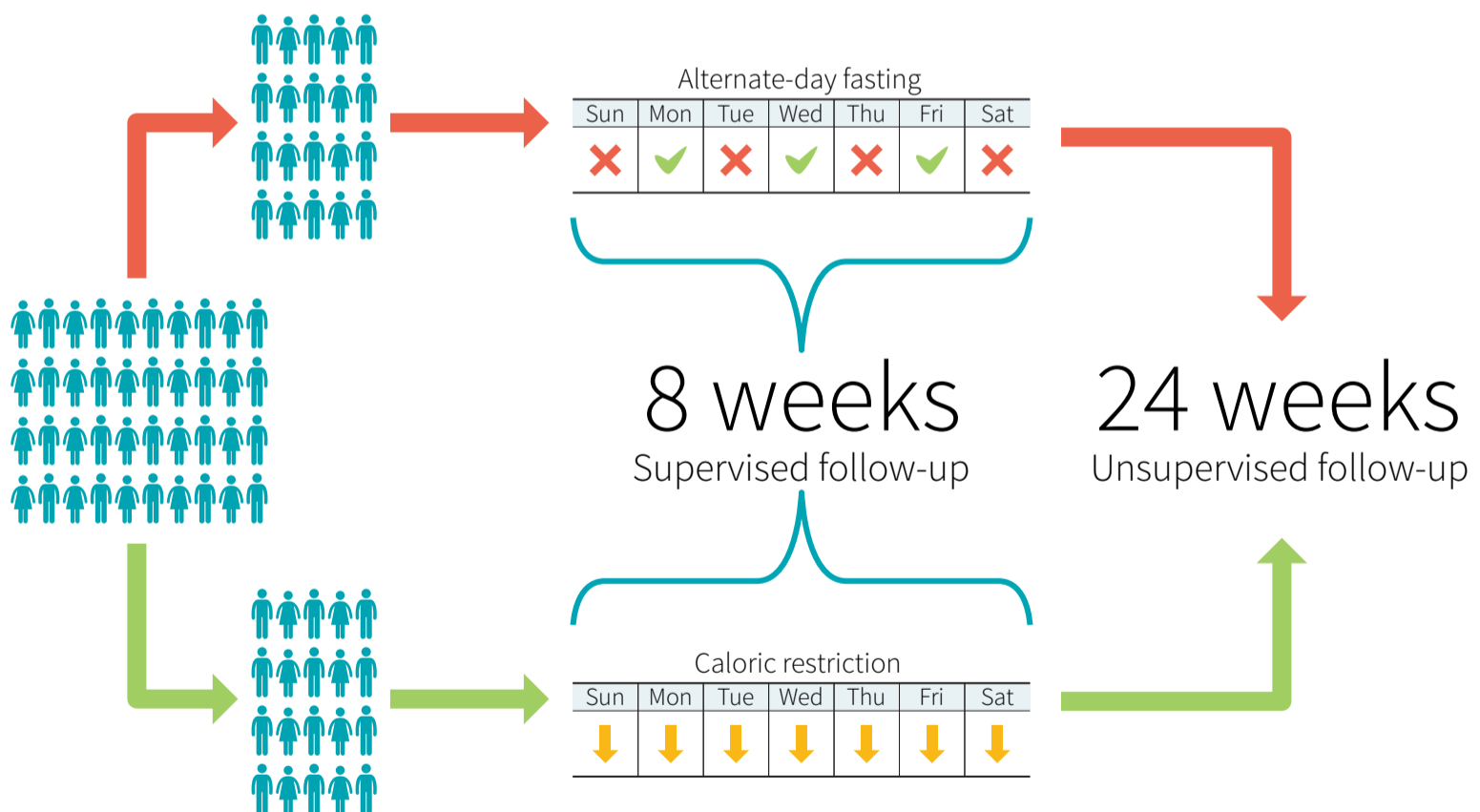
Methods: the most important part of the study

A paper's "Methods" (or "Materials and Methods") section provides information on the study's design and participants. Ideally, it should be so clear and detailed that other researchers can repeat the study without needing to contact the authors. You will need to examine this section to determine the study's strengths and limitations, which both affect how the study's results should be interpreted.

Demographics

The "Methods" section usually starts by providing information on the participants, such as age, sex, lifestyle, health status, and method of recruitment. This information will help you decide how relevant the study is to you, your loved ones, or your clients.

Figure 3: Example study protocol to compare two diets



The demographic information can be lengthy, you might be tempted to skip it, yet it affects both the **reliability** of the study and its **applicability**.

Reliability. The bigger the sample size of a study (i.e., the more participants it has), the more reliable its results. Note that a study often starts with more participants than it ends with; diet studies, notably, commonly see a fair number of dropouts.

Applicability. In health and fitness, *applicability* means that a compound or intervention (i.e., exercise, diet, supplement) 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 improve exercise performance.

Each individual is different, as the creatine example shows, yet a study’s demographic information can help you assess its applicability. If a trial [only recruited men](#), for instance, women reading the study should keep in mind that its results may be less applicable to them. Likewise, an intervention tested in college students may yield different results when performed on people from a retirement facility.

Figure 4: Some trials are sex-specific



Furthermore, different recruiting methods will attract different demographics, and so can influence the applicability of a trial. In most scenarios, trialists will use some form of “convenience sampling”. For instance, studies run by universities will often recruit among their students. However, some trialists will use “random sampling” to make their trial’s results more applicable to the general population. Such trials are generally called “augmented randomized controlled trials”.

Confounders

Finally, the demographic information will usually mention if people were excluded from the study, and if so, for what reason. Most often, the reason is the existence of a **confounder** — a variable that would confound (i.e., influence) the results.

For example, if you study the effect of a resistance training program on muscle mass, you don't want some of the participants to take muscle-building supplements while others don't. Either you'll want all of them to take the same supplements or, more likely, you'll want none of them to take any.

Likewise, if you study the effect of a muscle-building supplement on muscle mass, you don't want some of the participants to exercise while others do not. You'll either want all of them to follow the same workout program or, less likely, you'll want none of them to exercise.

It is of course possible for studies to have more than two groups. You could have, for instance, a study on the effect of a resistance training program with the following four groups:

- Resistance training program + no supplement
- Resistance training program + creatine
- No resistance training + no supplement
- No resistance training + creatine

But if your study has four groups instead of two, for each group to keep the same sample size you need twice as many participants — which makes your study more difficult and expensive to run.

When you come right down to it, any differences between the participants are variable and thus potential confounders. That's why trials in mice use specimens that are genetically very close to one another. That's also why trials in humans seldom attempt to test an intervention on a diverse sample of people. A trial restricted to older women, for instance, has in effect eliminated age and sex as confounders.

As we saw above, with a great enough sample size, we can have more groups. We can even create more groups *after* the study has run its course, by performing a subgroup analysis. For instance, if you run an observational study on the effect of [red meat](#) on thousands of people, you can later separate the data for “male” from the data for “female” and run a separate analysis on each subset of data. However, subgroup analyses of these sort are considered exploratory rather than confirmatory and could potentially lead to false positives. (When, for instance, a blood test **erroneously detects** a disease, it is called a **false positive**.)

Design and endpoints

The “Methods” section will also describe how the study was run. Design variants include single-blind trials, in which only the participants don’t know if they’re receiving a placebo; observational studies, in which researchers only observe a demographic and take measurements; and many more. (See [figure 2](#) above for more examples.)

More specifically, this is where you will learn about the length of the study, the dosages used, the workout regimen, the testing methods, and so on. Ideally, as we said, this information should be so clear and detailed that other researchers can repeat the study without needing to contact the authors.

Finally, the “Methods” section can also make clear the endpoints the researchers will be looking at. For instance, a study on the effects of a resistance training

program could use muscle mass as its primary endpoint (its main criterion to judge the outcome of the study) and fat mass, strength performance, and testosterone levels as secondary endpoints.

One trick of studies that want to find an effect (sometimes so that they can serve as marketing material for a product, but often simply because studies that show an effect are more likely to get published) is to collect many endpoints, then to make the paper about the endpoints that showed an effect, either by downplaying the other endpoints or by not mentioning them at all. To prevent such “data fishing” (a method whose devious efficacy was demonstrated through the hilarious [chocolate hoax](#)), many scientists push for the [preregistration of studies](#).

Sniffing out the tricks used by the less scrupulous authors is, alas, part of the skills you’ll need to develop to assess published studies.

“ Sniffing out the tricks used by the less scrupulous authors is, alas, part of the skills you’ll need to develop to assess published studies. ”

Interpreting the statistics

The “Methods” section usually concludes with a hearty statistics discussion. Determining whether an appropriate statistical analysis was used for a given trial is an entire field of study, so we suggest you don’t sweat the details; try to focus on the big picture.

First, let’s clear up two common misunderstandings. You may have read that an effect was **significant**, only to later discover that it was very small. Similarly, you may have read that **no effect** was found, yet when you read the paper you found that the intervention group had lost more weight than the placebo group. What gives?

The problem is simple: those quirky scientists don’t speak like normal people do.

For scientists, **significant doesn’t mean important** — it means *statistically significant*. An effect is significant if the data collected over the course of the trial would be unlikely if there really was no effect.

Therefore, an effect can be significant **yet very small** — 0.2 kg (0.5 lb) of weight loss over a year, for instance. More to the point, **an effect can be significant yet not clinically relevant** (meaning that it has no discernible effect on your health).

Relatedly, for scientists, **no effect** usually means *no statistically significant effect*. That’s why you may review the measurements collected over the course of a trial and notice an increase or a decrease yet read in the conclusion that no changes (or no effects) were found. There *were* changes, but they weren’t significant. In other words, there were changes, but so small that they *may* be due to random fluctuations (they *may* also be due to an actual effect; we can’t know for sure).

We saw earlier, in the “Demographics” section, that the bigger the sample size of a study, the more reliable its results. Relatedly, the bigger the sample size of a study, the greater its ability to find if small effects are significant. A small change is less likely to be due to random fluctuations when found in a study with a thousand people, let’s say, than in a study with ten people.

This explains why a meta-analysis may find significant changes by pooling the data of several studies which, independently, found no significant changes.

P-values 101

Most often, an effect is said to be significant if the statistical analysis (run by the researchers post-study) delivers a *P*-value that isn’t higher than a certain threshold (set by the researchers pre-study). We’ll call this threshold the *threshold of significance*.

Understanding how to interpret P -values correctly can be tricky, even for specialists, but here's an intuitive way to think about them:

Think about a coin toss. Flip a coin 100 times and you will get roughly a 50/50 split of heads and tails. Not terribly surprising. But what if you flip this coin 100 times and get heads every time? Now that's surprising! For the record, the probability of it actually happening is 0.00000000000000000000000000000008%.

You can think of P -values in terms of getting all heads when flipping a coin.

- A P -value of 5% ($p = 0.05$) is no more surprising than getting all heads on 4 coin tosses.
- A P -value of 0.5% ($p = 0.005$) is no more surprising than getting all heads on 8 coin tosses.
- A P -value of 0.05% ($p = 0.0005$) is no more surprising than getting all heads on 11 coin tosses.

Contrary to popular belief, the “ P ” in “ P -value” does not stand for “probability”. The probability of getting 4 heads in a row is 6.25%, not 5%. If you want to convert a P -value into coin tosses (technically called S -values) and a probability percentage, check out the converter [here](#).

As we saw, an effect is significant if the data collected over the course of the trial would be unlikely if there really was no effect. Now we can add that, the lower the P -value (under the threshold of significance), the more confident we can be that an effect is significant.

P -values 201

All right. Fair warning: we're going to get nerdy. Well, *nerdier*. Feel free to skip this section and resume reading [here](#).

Still with us? All right, then — let's get at it. As we've seen, researchers run statistical analyses on the results of their study (usually one analysis per endpoint) in order to decide whether or not the intervention had an effect. They commonly make this decision based on the **P -value** of the results, which tells you how likely a result at least as large as the one observed would be if the **null hypothesis**, among other assumptions, were true.

Ah, jargon! Don't panic, we'll explain and illustrate those concepts.

In every experiment there are generally two opposing statements: the **null hypothesis** and the **alternative hypothesis**. Let's imagine a fictional study testing the weight-loss supplement "Better Weight" against a placebo. The two opposing statements would look like this:

- **Null hypothesis:** compared to placebo, Better Weight **does not** increase or decrease weight. (The hypothesis is that the supplement's effect on weight is **null**.)
- **Alternative hypothesis:** compared to placebo, Better Weight **does** decrease or increase weight. (The hypothesis is that the supplement has an effect, positive or negative, on weight.)

Figure 5: Threshold for statistical significance



The purpose is to see whether the effect (here, on weight) of the intervention (here, a supplement called "Better Weight") is better, worse, or the same as the effect of the control (here, a placebo, but sometimes the control is another, well-studied intervention; for instance, a new drug can be studied against a reference drug).

For that purpose, the researchers usually set a threshold of significance (α) before the trial. If, at the end of the trial, the P -value (p) from the results is less than or equal to this threshold ($p \leq \alpha$), there is a significant difference between the effects of the two treatments studied. (Remember that, in this context, *significant* means *statistically significant*.)

The most commonly used threshold of significance is 5% (0.05). It means that if the **null hypothesis** (i.e., the idea that there was **no** difference between treatments) is true, then, after repeating the experiment an infinite number of times, the researchers would get a false positive (i.e., would detect a significant effect where there is none) at most 5% of the time ($p \leq 0.05$).

Generally, the P -value is a measure of consistency between the results of the study and the idea that the two treatments have the same effect. Let's see how this would play out in our Better Weight weight-loss trial, where one of the treatments is a supplement and the other a placebo:

- **Scenario 1:** The P -value is 0.80 ($p = 0.80$). The results are more consistent with the **null hypothesis** (i.e., the idea that **there is no difference** between the two treatments). We conclude that Better Weight had no significant effect on weight loss compared to placebo.
- **Scenario 2:** The P -value is 0.01 ($p = 0.01$). The results are more consistent with the **alternative hypothesis** (i.e., the idea that **there is a difference** between the two treatments). We conclude that Better Weight had a significant effect on weight loss compared to placebo.

While $p = 0.01$ is a significant result, so is $p = 0.000001$. So what information do smaller P -values offer us? They give us greater confidence in the findings. In our example, a P -value of 0.000001 would give us greater confidence that Better Weight had a significant effect on weight change.

Remember that a significant effect may not be clinically relevant. Let's say that we found a significant result of $p = 0.01$ showing that Better Weight improves weight loss. The catch: Better Weight produced only 0.2 kg (0.5 lb) more weight loss compared to placebo after one year — a difference too small to have any meaningful effect on health. In this case, though the result is significant, statistically, the real-world effect is too small to justify taking this supplement. (This type of scenario is more likely to take place when the study is large since, as we saw, the bigger the sample size of a study, the greater its ability to find if small effects are significant.)

Finally, we should mention that, though the most commonly used threshold of significance is 5% ($p \leq 0.05$), some studies require greater certainty. For instance, for genetic epidemiologists to declare that a genetic association is statistically significant (say, to declare that a gene is associated with weight gain), the threshold of significance is usually set at 0.0000005% ($p \leq 0.000000005$), which corresponds to getting all heads on 28 coin tosses. The probability of this happening is 0.00000003%.

P -values: Don't worship them!

Finally, keep in mind that, while important, **P -values aren't the final say on whether a study's conclusions are accurate.**

We saw that researchers too eager to find an effect in their study may resort to "data fishing". They may also try to lower P -values in various ways: for instance, they may run different analyses on the same data and only report the significant P -values, or they may recruit more and more participants until they get a statistically significant result. These bad

scientific practices are known as “p-hacking” or “selective reporting”. (You can read about a real-life example of this [here](#).)

While a study’s statistical analysis usually accounts for the variables the researchers were trying to control for, *P*-values can also be influenced (on purpose or not) by study design, hidden confounders, the types of statistical tests used, and much, much more. When evaluating the strength of a study’s design, imagine yourself in the researcher’s shoes and consider how you can torture a study to make it say what you want and advance your career in the process.



Reading the results

To conclude, researchers discuss the primary outcome, or what they were most interested in investigating, in a section commonly called “Results” or “Results and Discussion”. 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 first reading the “Methods” section; **knowing how researchers arrived at a conclusion is as important as the conclusion itself.**

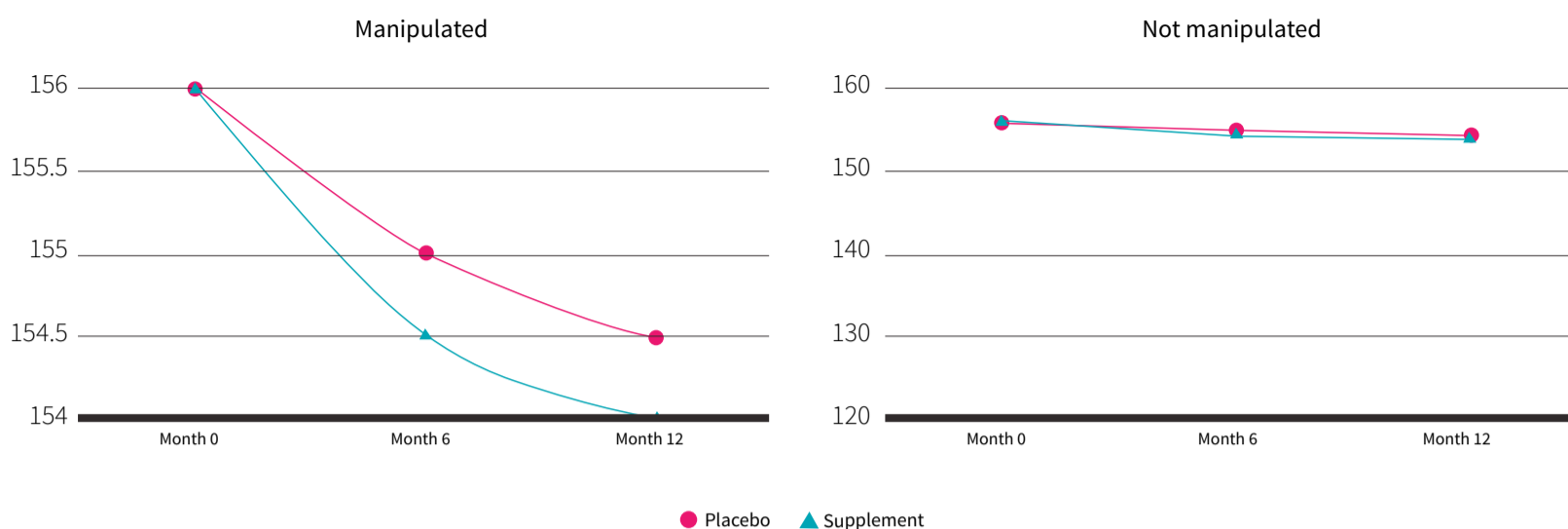
One of the first things to look for in the “Results” section is a comparison of characteristics between the tested groups. Big differences in baseline characteristics after randomization may mean the two groups are not truly comparable. These differences could be a result of chance or of the randomization method being applied incorrectly.

Researchers also have to report dropout and compliance rates. Life frequently gets in the way of science, so almost every trial has its share of participants that didn’t finish the trial or failed to follow the instructions. This is especially true of trials that are long or constraining (diet trials, for instance). Still, too great a proportion of dropouts or noncompliant participants should raise an eyebrow, especially if one group has a much higher dropout rate than the other(s).

Scientists use questionnaires, blood panels, and other methods of gathering data, all of which can be displayed through charts and graphs. Be sure to check on the vertical axis (y-axis) the scale the results are represented on; what may at first look like a large change could in fact be very minor.

In our Better Weight weight-loss trial, the supplement produced only 0.2 kg (0.5 lb) more weight loss compared to placebo after one year. By altering the y-axis, though, we can make this lackluster result look a lot more impressive:

Figure 6: Manipulation of the y-axis (lb)




The “Results” section can also include a secondary analysis, such as a **subgroup analysis** or a **sensitivity analysis**.

Subgroup analysis. As we saw at the end of our “Confounders” section, it consists in performing the analysis again but only on a subset of the participants. For instance, if your trial included both men and women of all ages, you could perform your analysis only on the “female” data or only on the “over 65” data, to see if you get a different result.

Sensitivity analysis. You may want to check if the results stay the same when you perform a different analysis or when, as in a subgroup analysis, you exclude some of the data (you could, in a meta-analysis, remove one study and run the meta-analysis again, for instance).

As we saw in the “Demographics” section, the reliability of a study depends on its sample size. If you exclude some of the participants from your analysis, the sample size decreases, and the risk of false positives can increase. It also means that if you play enough with the data, you may eventually get a positive result.

Let’s make up an extreme example: let’s say a researcher is paid to prove that “Better Weight” works. He tested Better Weight in 20 participants of both sexes, whose ages ranged from 21 to 87. Alas, of those 19 participants, only one lost weight. It happened to be a woman aged 65. The researcher could decide to perform a subgroup analysis excluding all men as well as all people not aged 65. He could then conclude that “Better Weight” is efficacious in women aged 65.



Clarifying the conclusion

Sometimes, the conclusion is split between “Results” and “Discussion”.

In the “Discussion” section, the authors expound the value of their work. They may also clarify their interpretation of the results or hypothesize a mechanism of action (i.e., the biochemistry underlying the effect). Often, they will compare their study to previous ones and suggest new experiments that could be conducted based on their study’s results. It is critically important to remember that a single study is just one piece of an overall puzzle. Where does this one fit within the body of evidence on this topic?

The authors should lay out what the strengths and weaknesses of their study were. Examine these critically. Did they do a good job of covering both? Did they leave out a critical limitation? You needn’t take their reporting at face value — analyze it.

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



Conflicts of interest

Conflicts of interest (COIs), if they exist, are usually disclosed after the conclusion. COIs can occur when the people who design, conduct, or analyze research have a motive to find certain results. The most obvious source of a COI is financial — when the study has been sponsored by a company, for instance, or when one of the authors works for a company that would gain from the study backing a certain effect.

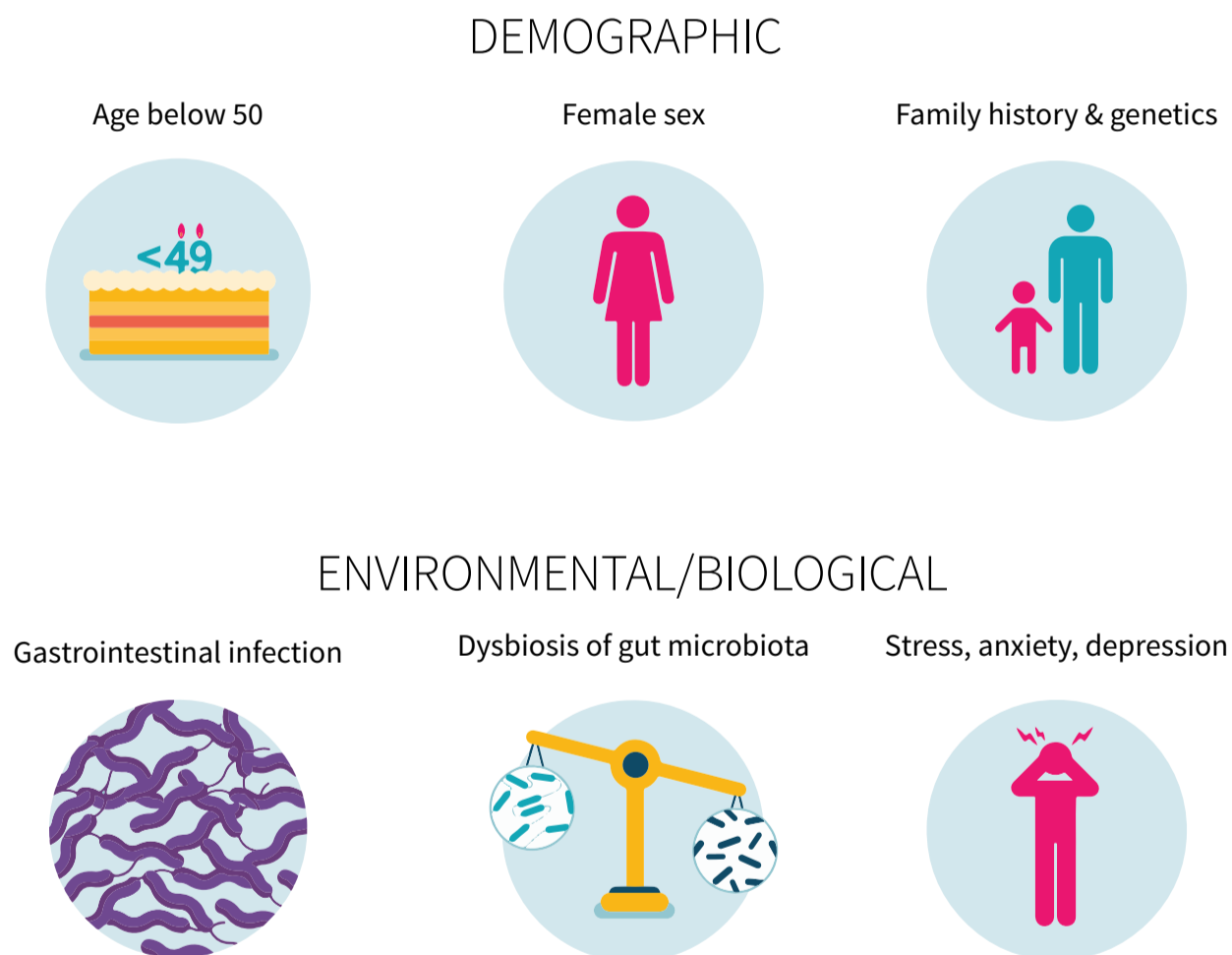
Sadly, [one study](#) suggested that nondisclosure of COIs is somewhat common. Additionally, what is considered a COI by one journal may not be by another, and some journals can themselves have COIs, though they don't have to disclose them. A journal from a country that exports a lot of a certain herb, for instance, may have hidden incentives to publish studies that back the benefits of that herb — so it isn't because a study is about an herb in general and not a specific product that you can assume there is no COI.

COIs must be evaluated carefully. Don't automatically assume that they don't exist just because they're not disclosed, but also don't assume that they necessarily influence the results if they do exist.

Digging down to the truth

As we saw in the “Demographics” section, the results of a study seldom 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. Subsequent studies showed that people who are not deficient in glutamine would not experience the same benefits as burn victims.

Figure 6: Many factors can influence the applicability of study results



Intentionally selecting a certain demographic makes sense for researchers who are looking for a way to help a specific kind of patient, but it can also be a strategy to promote certain results, which is why it isn't uncommon for new “fat burners” to be supported by studies that only recruited overweight postmenopausal women. When this type of information is left out of the abstract and then journalists skip the “Methods” section (or even the whole paper), people end up misled.

Never assume the media have read the entire study. A 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.

In conclusion, there's no substitute for appraising the study yourself, so when in doubt, re-read its "Methods" section to better assess its strengths and potential limitations.



Why you need a team

Going over and assessing just one paper can be a lot of work. Hours, in fact. Knowing the basics of study assessment is important, but we also understand that people have lives to lead. No single person has the time to read *all* the new studies coming out, and certain studies can benefit from being read by professionals with different areas of expertise.

With degrees in public health, exercise science, kinesiology, nutrition, pharmacology, toxicology, microbiology, molecular biophysics, biomedical science, neuroscience, chemistry, and more, the members of our team are all accredited experts, but with very different backgrounds, so that when we review the research, we get the full picture. Furthermore, we each have our own network to call upon whenever we need to contact the top specialists in any given field.

Professionals whose livelihoods depend on their getting reliable information trust in Examine.com to keep them abreast of the latest nutrition research; they trust in us to examine each study with the utmost care and report on it clearly, concisely, and accurately. But even if you're not a health professional, you can benefit from Examine.com's expertise by visiting our website (with its hundreds of free articles) or reading one of our practical guides — on fitness, keto, whey, and more.



Basic checklist

We've covered a lot in this guide, so here's a simplified checklist to keep handy the next time you want to dive into a nutritional paper.

- **What is the main hypothesis? (What question was the study trying to answer?)**
- **Does the paper clearly and precisely describe the design of the study?**
 - What type of study is it?
 - How long did the study last?
 - What were the primary and secondary endpoints?
- **If it is a trial, could you reproduce it with the information provided in the paper?**
 - Was the trial randomized? If so, how?
 - Was the trial blinded? If yes, was it single, double, or triple blinded?
 - What treatments were given? (Are sufficient details provided on what both the intervention and control groups did and did not receive?)
- **What demographic was studied?**
 - What is the sample size? (How many participants were recruited?)
 - Are the inclusion and exclusion criteria clearly laid out?
 - How were the participants recruited?
- **What did the analysis show?**
 - How many dropouts were there in each group?
 - Were the results statistically significant?

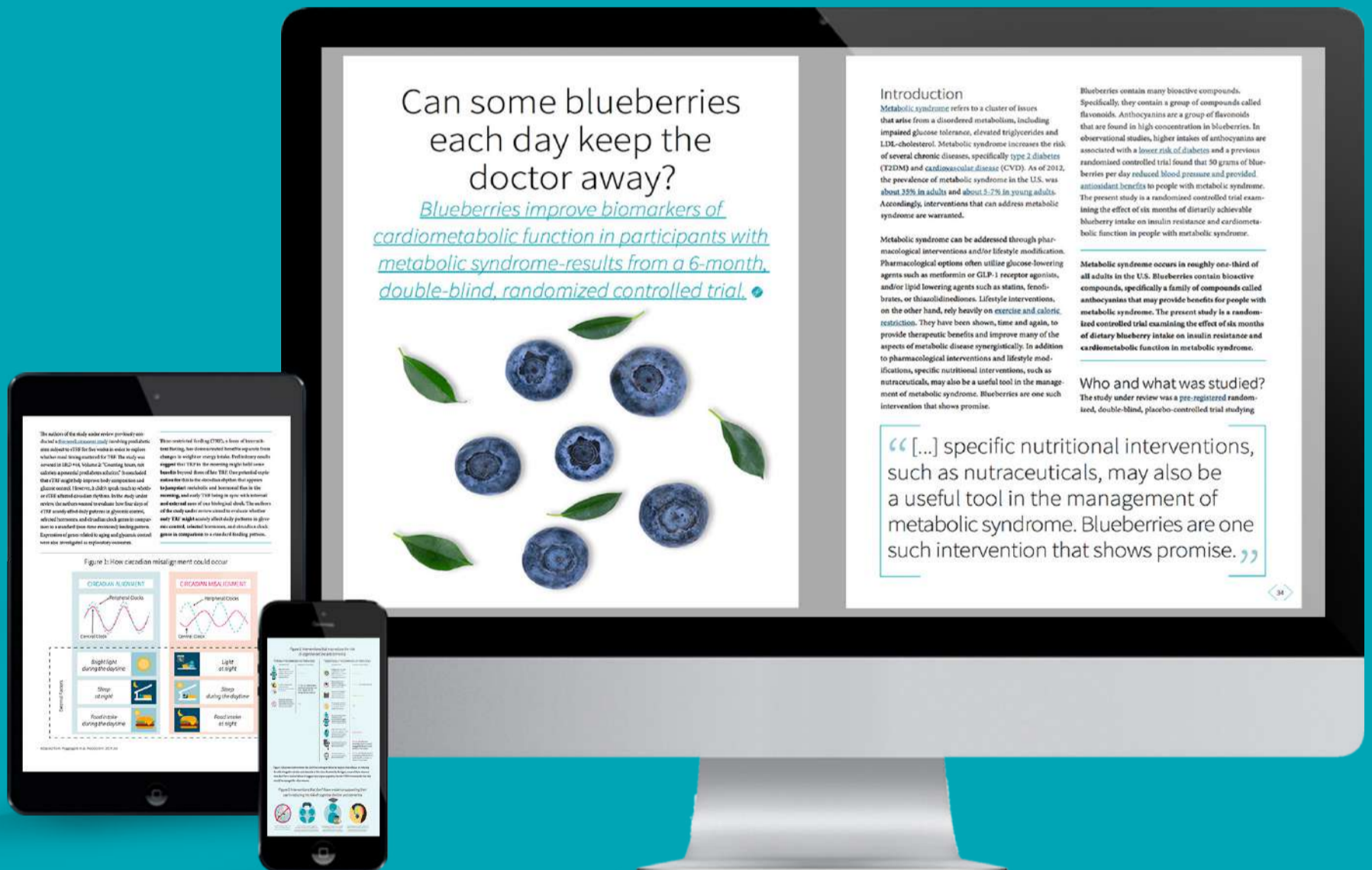
- **Are the results applicable to the real world?**
 - Were the results clinically relevant?
 - Based on the demographic studied, who might the results apply to?
 - Were the dosages realistic?

- **Were there any side effects or adverse events?**
 - If so, how severe were they?
 - If so, how frequently did they occur?

- **What were the sources of potential bias?**
 - Were there very unequal dropouts between groups? Why?
 - Did the intervention group actually follow the intervention?
 - Was the study pre-registered, to prevent “data fishing”?
 - What were the conflicts of interest, if any?

Looking for the most accurate deep-dive analysis of the science behind nutrition headlines?

If you want to stay up-to-date on the most important research, the **Examine.com Research Digest** is for you.



YES! SIGN ME UP