REPORT TO THE BOARDS OF HEALTH
Jennifer Morse, M.D., Medical Director

Mid-Michigan District Health Department, Wednesday, January 25, 2017
Central Michigan District Health Department, Wednesday, January 25, 2017
District Health Department 10, Friday, January 27, 2017

Evidence-Based Public Health Practice

Evidence-based medicine (EBM) has been around in it modern form since the late twentieth century[1]. It is not always clear what types of health care practices are associated with better clinical outcomes and EBM was developed in response to this. EBM uses the best available research evidence to guide clinical decision making in the care of patients[2]. The amount of research that is being done and evidence available to guide clinical decisions and practice guidelines continues to grow at a rapid pace, which has allowed for improving medical care.

Evidence-Based Public Health Practice has its roots in EBM as well as clinical epidemiology[3].

Evidence-Based Public Health practice is "development, implementation, and evaluation of effective programs and policies in public health through application of principles of scientific reasoning, including systematic uses of data and information systems, and appropriate use of behavioral science theory and program planning models"[4]. Given the limited time and resources available in public health, there is a great need to identify evidence of what interventions are effective prior to writing policy, recommendations, and creating programs. Healthy People 2020 set quantitative health promotion and disease prevention goals and objectives and proposed that these be tied to evidence-based interventions[3].

There are at least 455 journals related to public health world-wide, and 124 in the United States (http://www.scimagojr.com/journalrank.php?category=2739&type=j&country=US.) In 1984, the U.S. Preventative Services Task Force (USPSTF) (https://www.uspreventiveservicestaskforce.org) was formed as an independent, volunteer panel of national experts in evidence-based medicine and prevention to make recommendations about clinical preventative services[3]. The Cochrane Library (www.cochranelibrary.com), named after Dr. Archibald Cochrane, considered one of the fathers of EBM, contains countless reviews of medical evidence covering a wide variety of medical topics, including public health. Other useful evidence-based resources for public health include The Campbell Collaboration (https://www.campbellcollaboration.org/), The Community Guide (https://www.thecommunityguide.org/), and The Rural Health Information Hub Evidence-Based Toolkits (https://www.ruralhealthinfo.org/community-health/toolkits). These resources can aid with the interpretation and utilization of such a large amount of available evidence.

Using scientific evidence is very important, but can be challenging to understand. Included with this report are two articles: one with instructions for the non-scientist to read scientific papers; the second explains the different types of research studies and tips to analyzing studies.

Recommendations:

1. Support evidence-based medicine and evidence-based public health practices; ask your health care providers what the research shows to be the recommended treatments or screening you need or use the websites listed above to find the standard of care.
2. Be *very leery* of any non-expert (e.g. news, magazine, bloggers, etc.) interpretation of medical research. Research study findings are often misrepresented and the limitations of studies are almost never discussed. Review the original research yourself or discuss the findings of the study with your health care provider if you are interested or concerned.

---

[1] Claridge, J. A., & Fabian, T. C. (2005). History and development of evidence-based medicine. *World journal of surgery*, *29*(5), 547-553.
[2] Evans, A, et al. Evidence-based Medicine. In: UpToDate, Post TW (Ed), UpToDate, Waltham, MA. (Accessed on January 9, 2017).
[3] Secretary's Advisory Committee on National Health Promotion and Disease Prevention Objectives for 2020 (July, 2010). Evidence-Based Clinical and Public Health: Generating and Applying the Evidence.
[4] Brownson, Ross C., Elizabeth A. Baker, Terry L. Leet, and Kathleen N. Gillespie, Editors. *Evidence-Based Public Health*. New York: Oxford University Press, 2003.

How to Read and Understand a Scientific Paper: A Guide for Non-Scientists

By: Jennifer Raff August 25, 2013

"Be skeptical. But when you get proof, accept proof." – Michael Specter

What constitutes enough proof? Obviously everyone has a different answer to that question. But to form a truly educated opinion on a scientific subject, you need to become familiar with current research in that field. And to do that, you have to read the "primary research literature" (often just called "the literature"). You might have tried to read scientific papers before and been frustrated by the dense, stilted writing and the unfamiliar jargon. I remember feeling this way! Reading and understanding research papers is a skill which every single doctor and scientist has had to learn during graduate school. You can learn it too, but like any skill it takes patience and practice.

I want to help people become more scientifically literate, so I wrote this guide for how a layperson can approach reading and understanding a scientific research paper. It's appropriate for someone who has no background whatsoever in science or medicine, and based on the assumption that he or she is doing this for the purpose of getting a *basic* understanding of a paper and deciding whether or not it's a reputable study.

The type of scientific paper I'm discussing here is referred to as a **primary research article**. It's a peer-reviewed report of new research on a specific question (or questions). Another useful type of publication is a **review article**. Review articles are also peer-reviewed, and don't present new information, but summarize multiple primary research articles, to give a sense of the consensus, debates, and unanswered questions within a field. (I'm not going to say much more about them here, but be cautious about which review articles you read. Remember that they are only a snapshot of the research at the time they are published. A review article on, say, genome-wide association studies from 2001 is not going to be very informative in 2013. So much research has been done in the intervening years that the field has changed considerably).

Before you begin: some general advice

Reading a scientific paper is a completely different process than reading an article about science in a blog or newspaper. Not only do you read the sections in a different order than they're presented, but you also have to take notes, read it multiple times, and probably go look up other papers for some of the details. Reading a single paper may take you a very long time at first. Be patient with yourself. The process will go much faster as you gain experience.

Most primary research papers will be divided into the following sections: Abstract, Introduction, Methods, Results, and Conclusions/Interpretations/Discussion. The order will depend on which journal it's published in. Some journals have additional files (called Supplementary Online Information) which contain important details of the research, but are published online instead of in the article itself (make sure you don't skip these files).

Before you begin reading, take note of the authors and their institutional affiliations. Some institutions (e.g. University of Texas) are well-respected; others (e.g. the Discovery Institute) may appear to be legitimate research institutions but are actually agenda-driven. *Tip: google "Discovery Institute" to see why you don't want to use it as a scientific authority on evolutionary theory.*

Also take note of the journal in which it's published. Reputable (biomedical) journals will be indexed by Pubmed**.** Beware of questionable journals.

As you read, write down **every single word** that you don't understand. You're going to have to look them all up (yes, every one. I know it's a total pain. But you won't understand the paper if you don't understand the vocabulary. Scientific words have extremely precise meanings).

Step-by-step instructions for reading a primary research article

1. Begin by reading the introduction, not the abstract. The abstract is that dense first paragraph at the very beginning of a paper. In fact, that's often the only part of a paper that many non-scientists read when they're trying to build a scientific argument. (This is a terrible practice—don't do it.). When I'm choosing papers to read, I decide what's relevant to my interests based on a combination of the title and abstract. But when I've got a collection of papers assembled for deep reading, I always read the abstract last. I do this because abstracts contain a succinct summary of the entire paper, and I'm concerned about inadvertently becoming biased by the authors' interpretation of the results.

2. Identify the BIG QUESTION. Not "What is this paper about", but "What problem is this entire field trying to solve?"

This helps you focus on why this research is being done. Look closely for evidence of agenda-motivated research.

3. Summarize the background in five sentences or less. Here are some questions to guide you:

What work has been done before in this field to answer the BIG QUESTION? What are the limitations of that work? What, according to the authors, needs to be done next?

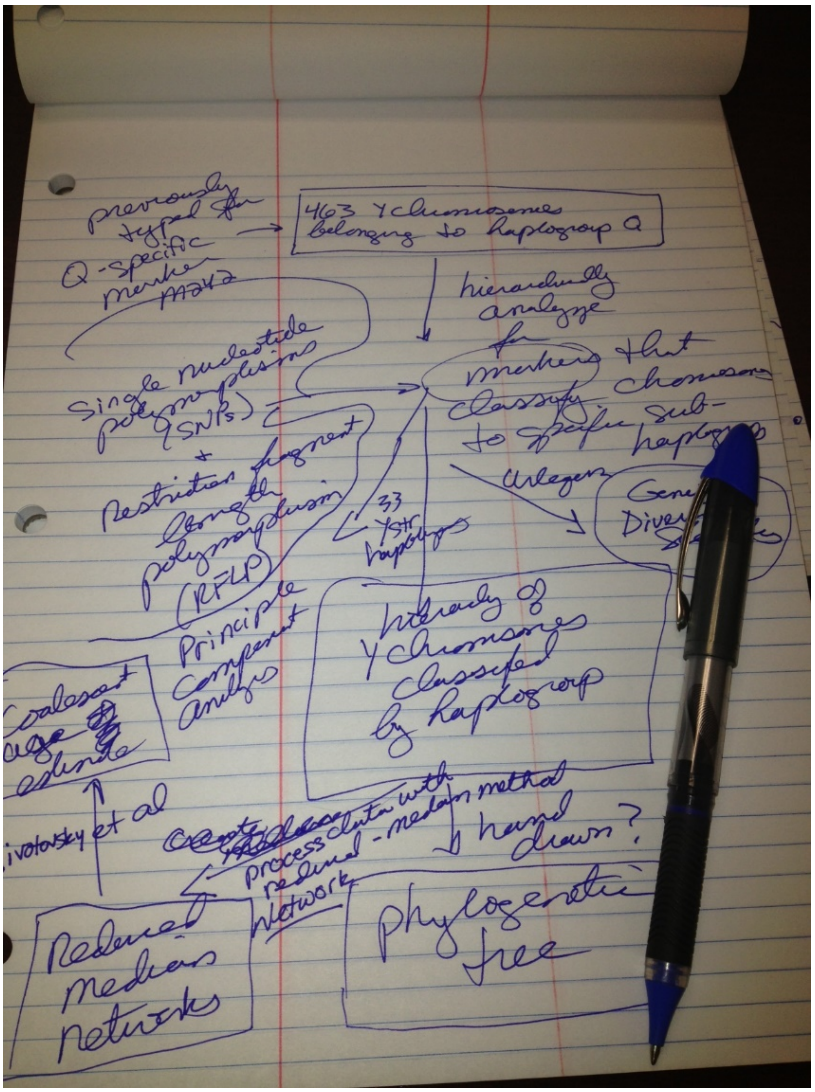
The five sentences part is a little arbitrary, but it forces you to be concise and really think about the context of this research. You need to be able to explain *why* this research has been done in order to understand it.

4. Identify the SPECIFIC QUESTION(S). What **exactly** are the authors trying to answer with their research? There may be multiple questions, or just one. Write them down. If it's the kind of research that tests one or more null hypotheses, identify it/them.

Not sure what a null hypothesis is? Go read this, then go back to my last post and read one of the papers that I linked to (like this one) and try to identify the null hypotheses in it. Keep in mind that not every paper will test a null hypothesis.

5. Identify the approach. What are the authors going to do to answer the SPECIFIC QUESTION(S)?

6. Now read the methods section. Draw a diagram for each experiment, showing exactly what the authors did. I mean *literally* draw it. Include as much detail as you need to fully understand the work. As an example, here is what I drew to sort out the methods for a paper I read today (Battaglia et al. 2013: "The first peopling of South America: New evidence from Y-chromosome haplogroup Q"). This is much less detail than you'd probably need, because it's a paper in my specialty and I use these methods all the time. But if you were reading this, and didn't happen to know what "process data with reduced-median method using Network" means, you'd need to look that up.



You don't need to understand the methods in enough detail to replicate the experiment—that's something reviewers have to do—but you're not ready to move on to the results until you can explain the basics of the methods to someone else.

7. Read the results section. Write one or more paragraphs to summarize the results for each experiment, each figure, and each table. Don't yet try to decide what the results *mean*, just write down what they *are.* You'll find that, particularly in good papers, the majority of the results are summarized in the figures and tables. Pay careful attention to them! You may also need to go to the Supplementary Online Information file to find some of the results.

It is at this point where difficulties can arise if statistical tests are employed in the paper and you don't have enough of a background

to understand them. I can't teach you stats in this post, but here, here, and here are some basic resources to help you. I STRONGLY advise you to become familiar with them.

THINGS TO PAY ATTENTION TO IN THE RESULTS SECTION:

- Any time the words "significant" or "non-significant" are used. These have precise statistical meanings. Read more about this here.
- If there are graphs, do they have error bars on them? For certain types of studies, a lack of confidence intervals is a major red flag.
- The sample size. Has the study been conducted on 10, or 10,000 people? (For some research purposes, a sample size of 10 is sufficient, but for most studies larger is better).

8. Do the results answer the SPECIFIC QUESTION(S)? What do you think they mean? Don't move on until you have thought about this. It's okay to change your mind in light of the authors' interpretation—in fact you probably will if you're still a beginner at this kind of analysis—but it's a really good habit to start forming your own interpretations before you read those of others.

9. Read the conclusion/discussion/Interpretation section. What do the authors think the results mean? Do you agree with them? Can you come up with any alternative way of interpreting them? Do the authors identify any weaknesses in their own study? Do you see any that the authors missed? (Don't assume they're infallible!) What do they propose to do as a next step? Do you agree with that?

10. Now, go back to the beginning and read the abstract. Does it match what the authors said in the paper? Does it fit with your interpretation of the paper?

11. FINAL STEP: *(Don't neglect doing this)* What do other researchers say about this paper? Who are the (acknowledged or self-proclaimed) experts in this particular field? Do they have criticisms of the study that you haven't thought of, or do they generally support it?

Here's a place where I do recommend you use Google! But do it last, so you are better prepared to think critically about what other people say.

12. Optional – but Critical. This step may be optional for you, depending on why you're reading a particular paper. But for me, it's critical! I go through the "Literature cited" section to see what other papers the authors cited. This allows me to better identify the important papers in a particular field, see if the authors cited my own papers (KIDDING! . . . mostly,) and find sources of useful ideas or techniques.)

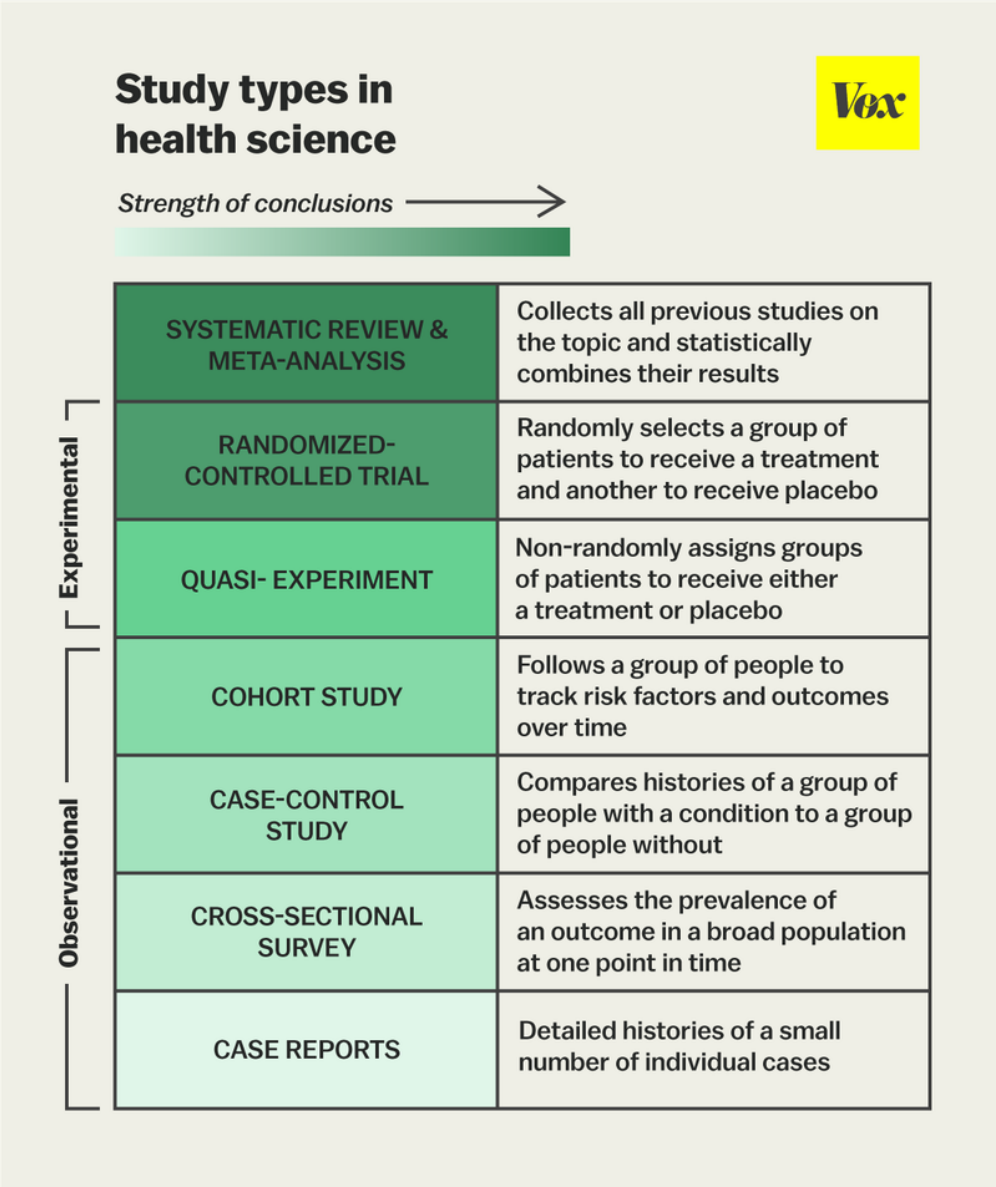
Now brace for more conflict– next week we're going to use this method to go through a paper on a controversial subject! Which one would you like to do? Shall we critique one of the papers I posted last week?

https://violentmetaphors.com/2013/08/25/how-to-read-and-understand-a-scientific-paper-2/

From: Vox

The One Chart You Need to Understand Any Health Study

By Julia Belluz and Steven Hoffman Jan 5, 2015, 11:30am EST



1) Much of health research can be broken down into two types: Observational and experimental studies

Much of health research — especially the kind that makes the news headlines — can be broken down into two basic types: observational and experimental.

In observational studies, scientists observe and gather data on some phenomenon that's already happening: patterns of olive oil consumption, who tends to take vitamin D supplements, how much people exercise, and so on. But they don't intervene at all to change anything in people's lives; they merely gather descriptive information on habits, beliefs, or events.

With experimental research, on the other hand, scientists do intervene, or at least use statistical methods to mimic intervention: they give some people a drug, they perform an operation on others. In the best-designed experiments, study participants are randomly divided into at least two groups: those who get the intervention (i.e., treatment) and those who don't (i.e., placebo). Random allocation ensures that the groups are statistically comparable with potential "confounding factors" equally distributed among them. The only difference between the groups is the intervention, which allows researchers to tease out what effect that intervention causes. This is why conclusions from experiments are generally considered to be more reliable and trustworthy.

2) There are four basic types of observational studies

There are many different types of observational studies, but here are the four most common that you need to know about: cross-sectional surveys, cohort studies, case-control studies, and case reports.

"Cross-sectional surveys" take a random sample of people and record information about them at one point in time. For example, researchers might survey randomly selected inhabitants of Washington, DC to figure out how many have heart disease (i.e., an epidemiological survey) or how they think about the quality of green space for outdoor exercise (i.e., a public opinion poll).

"Cohort studies" are just like surveys but they track the same groups of people over an extended period of time. That's why they are often called "longitudinal" and "prospective" studies. Instead of just gathering data on heart disease in Washington DC at one point in time, a cohort study would follow groups (or cohorts) of study participants over a period of, say, 10 years, and see how many people in each of the groups develop heart disease. This allows researchers to record changes in the health of the participants over time and compare the levels of health in different groups of people.

"Case-control studies" are often called "retrospective studies." That's because researchers start with an end point and work backward, figuring out what might have caused that outcome. For example, researchers could take two groups of people who live in Washington, DC: those who have been diagnosed with heart disease and those who haven't. They could then work backwards and survey the two groups about their earlier health behaviors to figure out what might have caused the disease to develop or not. They may ask about saturated fat consumption or exposure to disease-inducing viruses. From there, they would note any differences in risk factors or exposures that emerge between the two groups which can help suggest what may have led to heart disease in some people.

"***Case reports***" are basically detailed stories about a particular patient's medical history. If a doctor writes up case reports about a cluster of patients with the same condition or disease, this is a "case series." Though these are considered the weakest kind of observational studies, they can still be very helpful for rare diseases and powerful for advocacy. Sometimes they can be a bellwether in medicine. Early case reports, for example, led to the tragic discovery that mothers who were taking thalidomide for morning sickness were having babies with missing limbs. These reports surfaced long before a randomized trial could ever be done — and spared thousands of babies.

3) Observational studies have limits you need to understand

From a single observational study, researchers will only be able to suggest whether there's an association between a risk like fat consumption and an outcome like heart disease — and not that one *caused* the other. That's because the research participants were already eating fat or already had heart disease (or not) when the study began. What if people who eat lots of fat happen to be less health conscious? What if they are poorer and therefore more stressed? What if this particular group of fat-eaters just happened to be chubbier than those who stick to a low-fat diet? These things are called "confounding factors," or the difficult-to-predict variables that are associated with both the cause (e.g., saturated fat) and potential effect (e.g., heart disease) under study.

Sometimes confounding factors are knotty and wholly misleading. In 1991, the authors of a commentary published in the New England Journal of Medicine suggested that left-handed people had a higher risk of mortality. For their retrospective case-control study, researchers looked at death certificates from two counties in southern California and then asked family members of the deceased about their beloved ones' handedness. They found that being left-handed is associated with dying younger. "The mean age at death in the right-handed sample was 75 years, as compared with a mean age at death of 66 years in the left-handers," they wrote.

After publication, the journal editor was inundated with angry letter-writers. That's because the researchers failed to account for the cultural context: there was a time in the US when left-handed children were forced to become right-handed children. The reason there were few older left-handers was not because the hand you write with spells an early end, but because the would-be elderly lefties had converted when they were young and appeared as right-handed people in the study.

4) There are two basic types of experimental research

Now let's move on to experimental research. There are two basic types: randomized controlled trials and quasi-experimental designs.

"***Randomized controlled trials***" are considered the gold standard of medical evidence, though as you will probably surmise by now, they aren't necessarily the best study design for every research question. The reason they're so powerful, when they're well done, is because they are designed to tease out cause-and-effect relationships; randomization means treatment groups are comparable, and the only difference between them is the intervention (i.e., whether they received the drug or not) so any difference in outcome between the two groups can be attributed to the intervention.

When these experiments are blinded, they're even more powerful: blinding means either the study participants, the doctors, or both ("double-blinded") do not know whether they are receiving/giving the real treatment or a placebo. So blinded studies account for any placebo effects that may arise.

Lastly, there's a type of study design that lies somewhere between experimental and observational research: that's the "***quasi-experiment***." These are essentially a type of unplanned or uncontrolled experiment that uses statistics and human ingenuity to mimic the conditions of an experiment. Scientists have found many ways of undertaking these. One example would be comparing tobacco consumption before and after a border town is subjected to new state smoking regulations with its neighboring town in a different state that keeps the old regulations. Another example would be to evaluate the effects of GPA-based university scholarships by comparing those students who were just above and just below the grade point cut-off for receiving them.

The classical hierarchy of evidence. (From the MS Trust Information)

5) The king of all evidence: systematic reviews

Researchers often rank study designs in hierarchies (see above) to describe the relative weight of their conclusions. At the top of the hierarchy are syntheses of evidence that identify and integrate all sources of high-quality information relevant for a particular question coming from different contexts, settings, and methods.

These reviews address that problem of the single study puzzle piece. Rather than relying on just one person's

experience or even just one randomized controlled trial, synthesized evidence draws on multiple sources and weighs their contributions to arrive at a more fully-supported conclusion according to each study's rigor and relevance. This kind of research is regarded as the highest form of evidence — the king of all evidence if you will — and the best science to inform decision-making.

The idea is that many studies, done on thousands of people and taken together as a whole, can get us closer to the truth than any single study or anecdote ever could. (That is, unless a single study or anecdote is the only evidence available.) Reviews are less biased than a selective sampling of smaller studies that they might summarize.

Within synthesized evidence, the most reliable type for evaluating health claims are called "systematic reviews." These studies represent the best available syntheses of global evidence about the likely effects of different decisions, therapies and policies.

As their name suggests, systematic reviews use particular methods for finding helpful information, assembling it, and assessing its quality and applicability to the question you're interested in answering. Following this approach to the evidence — which is usually independently repeated at least twice by separate reviewers — reduces the bias that can creep into single studies. This process also helps to make sure results are not skewed or distorted by an individual author's preconceptions or cognitive biases. Finally, such transparency means that readers can know what the authors did to arrive at their conclusions and can easily evaluate the quality of the review itself.

You can log into a place like the Cochrane Library, Health Systems Evidence, or PubMed Health and read systematic reviews about everything from the effects of acupuncture for migraines and premenstrual syndrome, to the efficacy of cranberry juice for bladder infections. The hard-working people behind these studies are even starting to translate their conclusions into "plain language summaries," written in the way most people actually speak. This means these reviews and databases are more accessible than ever before. But then again, not all systematic reviews are created equally, either. And systematic reviews are only a starting point.

Even with the best available evidence from around the world at our disposal, we have to analyze it and apply it to our particular circumstances. A personal experience with the success or failure of a drug, like an allergic reaction, is more informative for you than the most rigorous study on the drug ever could be.

Just remember that one person's experiences are merely anecdotes — the least helpful type of evidence — for others. And one study, like the latest on whole grains, is only one piece of the puzzle.

OTHER TIPS:

1. **How can you tell if scientific evidence is strong or weak?**

The world abounds with evidence and studies, some of it good and some of it poor. How can you know what to trust? This card stack aims to guide you through tricky issues that can cloud your understanding of scientific findings.

One major problem is that scientific lingo often means something different in common parlance. And these words can insidiously sneak into media coverage. Simple words such as theory, significant, and control have totally different meanings in the realm of science.

2. **How can you tell if scientific evidence is strong or weak?**

The world abounds with evidence and studies, some of it good and some of it poor. How can you know what to trust? This card stack aims to guide you through tricky issues that can cloud your understanding of scientific findings.

One major problem is that scientific lingo often means something different in common parlance. And these words can insidiously sneak into media coverage. Simple words such as theory, significant, and control have totally different meanings in the realm of science.

3. Watch out for selection bias

If a psychologist (say) could run a single test on every person in the world, that would lead to some powerful results. But that's not practical. So scientists do the next best thing: they select a smaller group to study.

However, they always have to be careful about the particular group they are selecting. Studies can suffer from selection bias, in which the chosen subset isn't random enough and therefore somehow biased in favor of a certain outcome of the study.

Selection bias can happen in a number of ways. Perhaps certain types of people are more likely to want to be involved — or are committed enough to not quit, say, a longer multi-year experiment.

Consider a yearlong study of a weight-loss drug in which half the participants dropped out of the study before it was over. The ones who stayed in the game might have all lost weight, but it's also important to consider those who quit. Maybe those people were seeing no progress. So an apparent success rate of 100 percent was actually more like 50 percent.

Another issue to think about is whether the people participating are representative enough of the group that the paper or article is talking about.

This is why, for example, polling data from a nationally representative sample of people (such as the ones conducted by Pew and Gallup) can be more informative about national opinion than an informal poll that's open to anyone on the internet, even if far more people participated in the latter.

Another scenario that's a common problem with psychology studies is that they tend to primarily enroll American undergraduates because they are easy to recruit on college campuses. But undergrads aren't necessarily representative of the average American.

Similarly, studies in the US or Europe that look at people from "WEIRD" countries (that is, Westernized, Educated, Industrialized, Rich, and Democratic) might not apply to people from other cultures. For more, check out Bethany Brookshire's great piece about psychology's WEIRD problem at Slate.

4. Don't confuse correlation and causation

Often scientists will find that two different variables are correlated — for example, they both increase together over time. That's a hint that they might be related, but it doesn't necessarily mean that one is causing the other. Perhaps it's just a coincidence. Or perhaps a third variable is causing both of the other two. Further testing is typically required.

Over time, lots of correlative evidence — combined with systematically ruling out other possible causes — can lead to a stronger case that something is causing something else. But the best way to show causation is to perform a carefully controlled experiment.

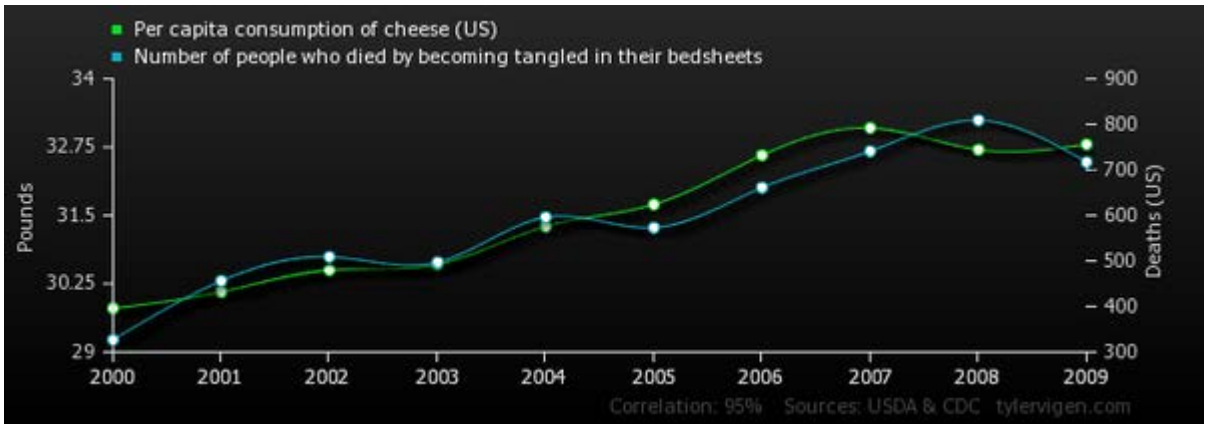
Here's an example: One study found that doctors were more likely to overprescribe antibiotics in the afternoon. That is, there was a correlation between antibiotics prescriptions and the time of day. The authors guessed that a phenomenon called decision fatigue could be the cause — that the brain gets tired after making too many decisions. But other possible causes could be too little sugar in the body (glucose fatigue) or general fatigue.

In order to find out if decision fatigue is a cause, they would need to set up an experiment in which they randomly had some doctors make more decisions than others. And to make the experiment as controlled as possible, those making fewer decisions would have to do some other kind of mentally fatiguing task, instead.

What does it mean that an experiment is controlled? In science, a control group is a group used for comparison. Control groups in medical studies often receive a placebo — a fake medicine, device, or procedure.

For example, many health problems will naturally get better (or worse) on their own. If you didn't also have a control group, you might think that, for example, you've invented a cure for the common cold. But the truth is that the common cold gets better in a week or two anyway.

It's also worth watching out for odd correlations. If a correlation seems really weird or too good to be true, then it's possible that there's nothing meaningful behind it. Tyler Vigen creates a bunch of what he calls spurious correlations from real data, like this one between per capita cheese consumption and people who have died by getting tangled in their bed sheets:



(tylervigen.com)

One fun correlation that people like to cite is the decrease in the number of pirates worldwide and the increase in global temperatures. However, it's highly unlikely that the loss of pirates has increased temperatures or that increasing temperatures has killed pirates off. It's also unlikely that some underlying cause is affecting both. There's a correlation, but it's meaningless.

5. **Look for the gold standard: double-blind, placebo-controlled, randomized tests**

The most reliable type of study — especially for clinical trials — is generally thought to be the randomized, placebo-controlled, double-blind study.

If you are looking at a clinical trial, a psychology study, or an animal study, and it hasn't been designed like this — and there isn't a good reason that it couldn't have been — then you might want to question the results.

Let's break down this terminology:

1) Randomized: This means that the participants in the study were randomly placed into the experimental group and the comparison group. This is important because if people get to choose, they might be more likely to pick one or the other because of some unexpected factor.

As a hypothetical example, maybe people who are more optimistic are more likely to want to try a new drug for anxiety rather than an old drug that's being used for comparison. And maybe optimism is linked to better outcomes for generalized anxiety disorder. The researchers could end up thinking that these people got better because of the drug when it was actually because they were innately going to do better anyway.

Similar problems can be introduced if researchers choose who goes into which category. That's why random is best.

2) Placebo-controlled: A controlled study has an appropriate comparison group, also called a control group. In medical studies, one comparison group usually gets a placebo — a fake intervention such as a sugar pill. This is in order to distinguish what the drug actually did from what a participant's psychological expectations did. (Placebo effects can be surprisingly strong — so strong that they can oftentimes relieve pain, among other health problems. And they've been getting stronger in recent decades, according to Steve Silberman's in-depth placebo story from Wired.)

A good placebo group should be as similar to the experimental group as possible. So, for example, if you were testing out a drug that's a large, red pill, you'd ideally want to give the people in your comparison placebo group a large, red pill that's the same in every way, but doesn't contain the drug. (Yes, even a pill's color and size can have a placebo effect.) Some studies go as far as to do sham surgeries, including anesthesia, incisions, stitches — the works.

3) Double-blind: A study is "blind" if the participants don't know whether they are in the experimental group or the control group. For example, you don't want someone knowing if she's received a real drug or a fake drug because her expectations could change the outcome of the study.

A study is "double-blind" if the researchers in personal contact with participants *also* don't know which treatment they are administering. You don't want the nurse giving out pills to know if they're real or not because then subtle differences in his behavior could influence patients — and therefore the results.

6. Understand "significance"

In everyday language, significant means that something is important or large. But a scientific finding that's considered "statistically significant" isn't necessarily either of those things. Scientists generally say that something is statistically significant if the effect can be picked up with a particular statistical tool called a p-value.

What's considered a good p-value is arbitrary and can vary somewhat between scientific fields. Often the cut-off for what's considered "statistically significant" is a p-value of 0.05.

It's important to keep in mind that p-values aren't the only relevant numbers in a study. For example, a treatment for a disease could have a statistically significant effect of changing the survival rate from 43 to 44 percent. That's a tiny change that probably isn't all that meaningful for how the disease will be treated in the future.

In fact, some people think that scientific papers should do away with p-values altogether and instead clearly and prominently show both the size of the effect and the range of the effect, both of which can be exceptionally important.

Another hazard: if you run a study many, many times or do a whole bunch of different statistical analyses on the same data, you could end up with results that look meaningful purely by chance. And then publishing only those meaningful-looking results would be likely to make the public draw misleading conclusions about your research. For more, Charles Seife has a good overview of various p-value pitfalls up at Scientific American.

7. Be aware of conflicts of interest

Conflicts of interest can come in many forms. The one that's generally most of concern within science and medical publishing these days is financial.

For example, this could be someone who received funding from a company that has a vested interest in the outcome of her own study. Or maybe that person has a relationship with the company — such as sitting on its board or acting as an unpaid consultant — that could lead to benefits in the future.

For example, one type of conflict would be a council that promotes a certain type of food then funding a study about that food's health benefits. Another would be a researcher who accepted travel money for a conference from a drug company and also researches that company's drugs or a competing company's drugs.

A recent analysis found that from 7 to 32 percent of randomized trials in top medical journals were completely funded by medical industry sources. And that's just those with full, direct funding. Presumably, the percent with any type of conflict of interest would be far higher.

One solution might be to ban such conflicts of interest. But what many journals have opted for instead are various requirements about disclosure, such as this one from the journal *Science,* which instructs people submitting papers to reveal "any affiliations, funding sources, or financial holdings that might raise questions about possible sources of bias." (The actual form that authors fill out is even more detailed.)

The editor then determines which relationships should be printed publicly with the scientific paper. And then whoever's reading the paper can draw their own conclusions about whether those relationships have — knowingly or unknowingly — influenced the data.

What, if anything, needs to be disclosed is up to the journal (and in some cases people's employers, too). Many journals publish conflict-of-interest policies on their websites. And if you scrutinize a paper closely, you may find some of these disclosures, too.

8. Know that peer review isn't perfect

Peer review is the system in which a couple of independent experts read over a paper that's been submitted to a journal. Generally, a journal isn't considered high quality if the papers aren't peer reviewed.

It's usually the case that reviewers are chosen by the journal and kept anonymous so that the review can be as impartial as possible. These people can recommend revisions to the text, new experiments that should be added,

or even that the journal shouldn't publish the paper. Then, the paper's authors will generally look at those reviews and incorporate them into a revised paper, if necessary.

But reviewers *aren't* asked to do everything within their power to make sure that the results are absolutely correct. (That would simply take too much time and be impractical. A paper can sometimes take years to put together already.) For example, reviewers are not expected to try the experiments themselves. And they don't generally look at raw data or re-run calculations.

They *do* look at the manuscript to see if the experiments were properly designed, if the data supports the paper's conclusions, and if the findings seem important enough to warrant publication.

So, peer review is generally beneficial, but not perfect. The scientific process really isn't complete until someone else replicates what's in the paper. And that's something that happens (if at all) after publication, not before.

In addition, sometimes papers get retracted. It's rare, but it does happen. Ivan Oransky and Adam Marcus's blog Retraction Watch is a great place to hear stories of the biggest, most important, and most dramatic retractions. (And they can be quite dramatic. In 2014, the *Journal of Vibration and Control*, for example, retracted 60 papers all at once.)

There are a few odd exceptions to the general peer review process, including the journal the *Proceedings of the National Academy of Sciences*, which allows members of its exceptionally prestigious academy to choose their own reviewers for up to four papers a year. Peter Aldhous has a good story about this controversial process over at *Nature*. (*PNAS* also accepts many papers through a more traditional peer-review system.)

9. Realize that not all journals are good

Just because it's in a journal doesn't mean it's a fantastic study. Journals and papers both range from great to mediocre to downright fraudulent. And even a top-notch journal can sometimes publish a flawed study.

The most commonly used metric to assess a scientific journal's influence is the Impact Factor. The Impact Factor is essentially a measure of popularity. It counts the number of times a journal's papers have been mentioned in other papers, relative to the journal's own volume of article output.

The more of these citations that appear, the more influence the journal seems to have on people's work. (Specifically, the IF is calculated from citations in the Thomson Reuters Journal Citation Reports database.)

How do you find a journal's Impact Factor? If you belong to a good library, some will have a subscription that can get you into the Journal Citation Reports analysis that comes out each year. If not, many journals and journal publishers will proudly list their rating somewhere on their websites. Just search for "impact factor." For comparison's sake, some of the most prestigious journals around, such as *Science*, *Nature*, and *JAMA*, have Impact Factors in the high-20s to mid-30s. (And the *New England Journal of Medicine* has an astounding Impact Factor in the 50s.)

The Impact Factor is controversial. It's a handy tool, but it's not the only way to look at things.

Some fields of science naturally generate more citations than others, but that doesn't necessarily mean that they're really better or more influential. And at least one study has found that Impact Factors don't correlate well with expert opinion.

Another thing to look out for are predatory, for-profit journals that will publish just about anything (and without peer review). In recent sting operations, several people have gotten such journals interested in publishing flawed or incoherent papers. (Added note: list of list of questionable, predatory, scholarly open-access standalone journals can be found on the Scholarly Open Access website, maintained by an academic librarian, at https://scholarlyoa.com/individual-journals/).

Also keep in mind if the study is in an appropriate journal for its subject matter. Sometimes junk science can end up in a peer-reviewed journal, especially if it's outside the journal's area of expertise. Its reviewers and editors might be less able to accurately assess the paper's quality.