

Understanding scientific studies

01 January 2008

Background

“Eat eggs and exercise to live longer!”

Obviously, this ‘headline’ is fictitious, but it got your attention, didn't it? Reports of new research findings in the area of food and health seem to get a lot of media coverage, after all, food and health issues are of interest to nearly everyone. Reporters, academics, and health professionals all want the latest information—as do consumers. A single study and the reports surrounding it can send many people in search of the latest food or supplement holding the promise of good health.

Frustrated and confused by the tremendous amount of food and health information reported nowadays, many people want simple certainties to help them protect their health through diet. The trouble is that single studies rarely provide such certainty, although they often get big headlines.

The media, health professionals, and educators are the gatekeepers of today's food and health information. They have a large influence on what consumers hear, read, and believe about food and health. Along with that comes the responsibility to provide the facts accurately, put them in perspective, and help people understand how the findings may affect their behaviour and health choices. Responsible media reporting means that new studies should be critically reviewed before being publicised. News releases and study abstracts, although helpful for the identification of interesting research, do not provide all the information necessary to accurately report findings to the public.

Fortunately the process of critically reviewing scientific research becomes easier and faster with practice. This EUFIC Review is designed to help the media, health professionals, and educators understand how to read and evaluate food and health-related scientific studies. It presents an overview of key information to look for, questions to ask, and other important considerations.

A process of discovery and debate

To consumers, it often seems that contradictory studies about food and health appear in the media almost weekly, leaving many to wonder why researchers can't get it right the first time. The answer is not straight forward, because to understand it means we must accept uncertainty.

The scientific process is a road of discovery. It is the process of gaining knowledge about the universe through the observation of measurable evidence. Contrary to what many people believe, this ‘road’ is not a straight, smooth motorway: researchers may take different directions of exploration, going down routes that twist, turn, and sometimes even backtrack or come to a dead end, before the facts are uncovered. Even then, the facts uncovered may be only part of a larger, partially understood phenomenon, which requires further research before we come to more complete answers.

As a result, the scientific process—how studies are designed, conducted, and reported—frequently generates a great deal of debate. Tracking the debate is often key to putting new research into context. With that in mind, new research studies published in scientific journals should be viewed as discussions among scientists. In these discussions, almost no one gets to have the final word, as it is rare that a study provides a final, complete answer. In fact, occasionally even old, accepted research results are revisited and discussed again. With the benefit of new information or technology, scientists sometimes see previous research results in a new light. The publication of research findings allows researchers to get opinions and critiques on their work by other experts, which not only confirms or contradicts their conclusions but also adds to the body of literature on a subject and so helps shape future research.

The bottom line is that dialogues characterised by cycles of revision, conjecture, assertion, and contradiction are frequently key to investigating a subject. Although such cycles often frustrate non-scientists and contribute to increasing public scepticism about advice on food and health, it is important to understand that science is evolutionary, not revolutionary. Because scientific research explores the unknown, uncertainty is an unavoidable part of current investigations. Only through repeated research and analyses can certainties emerge.

Types of research studies: What they are and when they are used

Research can be divided into two main categories: observational and experimental. Within these categories, there are three basic methods by which scientists investigate questions about food and health. It is essential to understand what each method can and cannot do.

Observational research

Observational research involves the examination of specific factors in defined groups of subjects to investigate the relationships between those factors and aspects of health or illness. For example, an observational study may focus on the body weight of healthy women aged 50 years old or older, and its relationship to blood pressure in that group. However, observational research can only suggest relationships between factors; it takes experimental research to determine cause and effect.

Epidemiological research is often observational, but it may also be experimental. It is the study of the determinants of disease or other health outcomes and their distribution in human populations. It seeks to expose possible associations between different aspects of health (such as cancer and heart disease) and diet, smoking habits or other lifestyle factors within populations.

Although epidemiological studies are useful for suggesting relationships between two factors, it is important to remember their basic limitation: they do not necessarily prove cause and effect. In fact, the associations that they indicate can actually turn out to be coincidental. A simple example would be a study which suggests that driving a car is associated with increased risk for cardiovascular disease. In this case, the fact of driving a car is a coincidence. The association revealed by the study should have been between driver characteristics (gender, age, weight) and the disease.

Observational epidemiological research may be most revealing when considered together with experimental research. For example, to assess whether an association discovered in an epidemiological study is real or the result of bias or confounding factors, researchers need to conduct a randomised clinical trial to confirm a suspected cause-and-effect relationship.

What is a double-blind placebo-controlled trial?

Considered the 'gold standard' of clinical research, the double-blind placebo-controlled study provides dependable findings without bias introduced by either the subject or the researcher.

In this type of study, neither the subject nor the researcher conducting the study knows whether the test substance or a placebo has been administered. For the results to be valid and to ensure that the subject cannot violate the 'blindness', the placebo and the test substance must be virtually identical (look, smell and taste similar).

The 'blindness' of the study is crucial. It eliminates the possibility that a participant's personal assumptions or beliefs undermine the validity of outcomes. It also prevents the researcher's expectations from influencing the test results. However, this is not always possible. For example, in studies about foods to compare the glycaemic index of bread to that of corn, it will be very difficult to hide the fact that the tested foods are different.

Experimental research

In experimental research, study subjects (whether human or animal) are selected according to relevant characteristics and are then randomly assigned to either an experimental group or a control group. This is called a controlled randomised trial. Random assignment ensures that factors (known as variables) that may affect the outcome of the study are distributed equally among the groups and therefore could not lead to differences in the effect of a treatment. The experimental group(s) is then given a treatment (sometimes called an intervention), and the results are compared with those for the control group, which does not receive a treatment (or which receives a placebo, or fake/dummy treatment). If the randomisation is done correctly, any differences in results between the groups can then be attributed to the treatment; the effect can be considered as caused by the treatment. Controlled experimental research can be fraught with errors, sometimes becoming 'uncontrolled'. For example, the population might have not been selected properly or the groups might know whether or not they are receiving the treatment. These flaws may be spotted easily, but in many cases it is worth asking experts—they know what to look for.

Basic research generates data by investigating biochemical substances or biological processes. It is often undertaken to confirm observations or discover how a particular process works. For example, an experiment might examine how vitamin E helps prevent oxidation of LDL (low-density lipoprotein) cholesterol, a process believed to play a role in the development of heart disease. This basic research is just part of a larger effort to understand how diet can help reduce risk for heart disease.

Basic research may be conducted *in vitro* (such as in test tubes) or *in vivo* (with animals). Research with animals is an important tool in determining how humans may react when exposed to particular substances. However, because of differences in physiology and the fact that animals are routinely exposed to far higher levels of compounds than those that human populations typically encounter, results from studies with animals cannot always be generalised to humans.

Clinical trials deal with the experimental study of human subjects. Trials attempt to determine whether the findings of basic research are applicable to humans, or to confirm the results of epidemiological research. Studies may be small, with a limited number of participants, or they may be large intervention trials that seek to discover the outcome of treatments on entire populations. 'Gold standard' clinical trials are double-blind, placebo-controlled studies that use random assignment of subjects to experimental or control groups.

Meet the meta-analysis

A meta-analysis is a statistical method of combining results from separate studies to get broader overall conclusions about a hypothesis. Meta-analyses are a useful way to reconcile differences of statistical power or sample size between studies, or to aggregate relevant findings across studies.

The procedure is most appropriate when examining studies that look at the same question and use similar methods to measure relevant variables. For example, scientists used a meta-analysis of observational studies to examine the relationship between the consumption of red meat, and the risk of bowel cancer. Although individual studies showed different results, pooling of data from 16 similar studies showed significant decreases in the risk of oral cancer with increased intakes of fruits and vegetables.

However, as with every study technique, the meta-analysis is not without limitations. Data from flawed studies may be included, or the analysis may include data from studies that use different methods to measure variables—resulting in a comparison of 'apples and oranges'.

Considerations for judging the validity of a meta-analysis include the following:

- Is the objective clearly stated?
- Are the criteria for inclusion or exclusion of studies explicit?
- Is the search mechanism for the determination of suitable studies adequate?
- Is the quality of the trials included assessed?
- Are all of the (intervention) trials randomised?
- Does the discussion include mention of limitations? Does it put results in context?
- Are the conclusions justified by the data?

What to look for when critically reviewing scientific studies

To enhance communication among scientists and to make the replication of a study easier, published research generally follows an established format. This section of this Review highlights important

information to look for, and questions to ask yourself or to pose to experts. This information should help you to understand each part of the study. However exactly where information appears in different articles can vary somewhat. In general, a scientific article is composed of:

Abstract

The abstract of a published study serves to describe briefly what was studied, how it was done, and the results. It allows readers to make a judgement of whether a study is of interest, without having to read the complete paper. If only we could just skim the abstract and consider our review of the study complete! Unfortunately, that is not the case. Abstracts do not provide sufficient detail to enable readers to assess the validity of a study, or put it into context. Only reading the rest of the study can do these.

Introduction

The introduction section of a report, 'sets the scene' for the reader of the study. It eases the reader into the research by presenting the issue that the researcher seeks to answer or the problem/hypothesis that the study addresses. It explains why the study was conducted, which gives the reader an impression of the potential importance of the research. It also expands on how the research was conducted. The introduction can be divided into two parts: the Background section and the Purpose section.

Background: The background information presented in the introduction of a study tells why the researchers think the study is important. It should reflect a comprehensive knowledge of the body of research on the subject and should brief the reader on both previous studies that support the concepts or theories of the current study and also those that do not. In essence, it brings the reader up to speed on current thinking and presents the researcher's rationale for pursuing the study.

Purpose: The purpose essentially defines the study. It states the question and briefly announces how it will be investigated.

In some instances you may find that the study does not seem to be appropriately designed or conducted to achieve its purpose. For example, the type of study undertaken may not yield the type of information required to answer the stated question, or the study population may not fit the purpose. Consultation with experts about these points can be very helpful in determining the validity of the study's conclusions.

Key questions to ask:

- What are the inherent limitations of this type of study?
- Does the research design fit the stated purpose of the study?
- Has the author omitted from the Background section important points that could affect the study design or interpretation of the results?

Peer review



Peer review (also known as refereeing), is the process of subjecting the research to the scrutiny of other experts in the same field. It is used by editors of academic journals to select and to screen submitted manuscripts. The peer review process aims to make authors meet the standards of their discipline, and of science in general, increasing the probability that weaknesses will be identified, and, with advice and encouragement, fixed.

Reviewers are typically anonymous and independent. In addition, since reviewers are normally selected from experts in the fields discussed in the article, the process of peer review is considered critical to establishing a reliable body of research and knowledge.

Publications that have not undergone peer review are likely to be regarded with suspicion by professionals in many fields. However, even refereed journals can contain errors.

Methodology

The key question of the methodology section is "How?" This section should enable critical readers to determine whether the research is valid: was it adequately designed to achieve its purpose? For this reason, the methodology section warrants careful review. It explains how the research was conducted and should give information in enough detail for the reader to evaluate the study. It should also enable the reader to understand to whom or what the study results apply. Important information featured in the methodology section includes the following:

- the setting of a study (in a clinic, laboratory, free-living population etc.)
- how variables were controlled (how did they adjust for specific subject qualities or outside influences that could affect the results?)
- the sample size
- the number of study groups
- the treatment or variables being observed (e.g. a vitamin supplement or specific diet)
- the length of the study
- how the data were collected
- how and by what statistical procedures the data were analysed

The methodology section also provides information about the sampling method and whether subjects were randomly assigned or not (in experimental studies). Pay specific attention to these points, because they are among the first steps in conducting research, and flaws that occur here can render the results invalid.

Talking with experts

Journalists, educators, and health professionals who need to quickly distil and understand the findings of a new study, should consider contacting the study authors or other scientists familiar with the body of research on the topic. Experts can answer questions and provide insight that both novice and experienced readers may not be able to unearth by themselves. In addition, interviewing scientists other than the study



author can bring valuable insights and contribute to a more balanced perspective of the study.

Questions you may want to ask a study's author or other experts include the following:

Could the study be interpreted to conclude something else?

Scientists often reach different conclusions from the same or similar data, so asking "What's your take on this study" is not unusual. The rationale for different conclusions may be important when putting a study's findings into context.

Are there any methodological flaws in the study that should be considered when making conclusions?

The more experts from different fields who review a study, the more likely potential flaws—such as confounding variables—will be discovered.

Can the study's results be applied to other groups?

Consumers want to know how research affects them. If study results are applicable only to a defined group of people, it should be reported as such.

How does this work fit with the body of research on the subject?

Even a well-written article may not include a discussion of all relevant research because of space limitations. Yet, it is extremely important to know—and communicate to consumers—whether a study confirms previous research, therefore adding weight to scientific beliefs, or whether the study's results and conclusions diverge from current thinking on the subject.

Randomness in selection and assignment: The term 'random sample' is familiar to most of us, but exactly how subjects (the 'sample') are selected for the study is of crucial importance. The sampling method can affect to whom the study results may be relevant.

If the subjects are selected randomly, that is, by procedures in which all individuals in a population being studied have an equal chance of being selected, then the study results may be generalisable to that population. True random selection may be done using a table of random numbers generated by a computer. Telephoning people picked randomly out of a telephone book between 13.00 hrs and 15.00 hrs is not random sampling of the entire population of Berlin in Germany, for instance. We can all think of reasons why this is not truly random: some people don't have telephones, and some people have unlisted phone numbers. In addition, the sample would be likely to have fewer workers in full-time employment, and would be weighted towards stay-at-home parents, elderly and unemployed people, students, people who are ill, and people who work the night shift, for example.

The term 'random' also applies to the assignment/ division of subjects into groups. Random assignment ensures that all subjects have an equal chance of being in the experimental or the control groups, and

increases the probability that unidentified variables systematically occur in both groups with the same frequency. Randomisation is crucial to controlling for variables that researchers may not be aware of, or cannot adequately control, but that could affect the outcome of an experimental study.

To determine the true effect of a treatment, researchers must carefully control for all known variables that could affect the outcome of a study. Some of the variables are obvious, such as age, body weight, and gender. To control for these differences, researchers match subjects in experimental and control groups so that they have similar characteristics. Some variables, such as heredity, are more difficult to control for. Still others may be unknown—because knowledge of human biology is still incomplete. By randomly assigning subjects to study groups, the influence of such variables is minimised, and any differences in results between groups can be attributed to the treatment.

A question of sample size: The simple question about sample size is "Was it big enough to find an effect?" The answer is not always as easy, and is often a matter of judgment. For example, when studying the effect of a weight-loss drug, a researcher may decide that a sample size of 100 people is adequate because the effect is easily measured: How many kilos did those who received the drug lose, compared with those who did not receive the drug? However, when assessing the average fruit and vegetable consumption among children who participated in a school-based intervention program, several thousand children may be needed, because the increase from such an intervention is likely to be relatively small. That is, the diets of the children in the experimental and the control groups may not differ much in terms of fruit and vegetable intake, and therefore, the effect of the intervention might not be noticed. It is easier to identify a small effect when you are looking at the results from a large sample.

A small sample size, however, does not necessarily mean that the study is flawed. For example, prospective clinical nutrition studies usually have a small number of subjects because there are so many variables that need to be controlled. When reading a study, be sure to look for the rationale that the researcher used to decide on the sample size.

One also should bear in mind the way in which the data collection is done. For example, in nutritional epidemiology investigating the role of food and nutrition in human health and disease, dietary consumption is frequently assessed. The complexity of the human diet makes this measurement a major challenge. The most accurate information on food intake can be obtained by food diaries or dietary records, where people record their daily dietary intake over short periods of time. Another method often used is a food frequency questionnaire. This has some limitations, as it relies on individuals recalling what they consumed previously over a period of time (e.g. months) and there may have been dietary changes in this time. These uncertainty aspects of dietary intake reports should be taken into account when epidemiological associations are reported.

OF ABSTRACTS AND NEWS

When deadlines loom, it is tempting to rely on an abstract and a press release for information about new research, instead of reading the original published study. Resist the temptation! Abstracts and releases are not substitutes for looking at original research. They do not provide enough information to clearly judge

the merits of a study or to accurately report the study results.

Medical journals, organisations and universities regularly issue press releases to stimulate media coverage of research or conferences. Press releases and abstracts are a useful way to identify ideas, story angles, quotes, potential interview subjects, and as a glimpse into research findings. Once you have a little background on the study, you can then review the original research article without spending too much additional time. Keep the 'Key Questions to Ask' in mind as you read the article, and note specific questions that you have for experts as you go along.

A word about methodological limitations: Often limitations are placed on researchers, such as limits to finances or the ethics of human testing, and these can restrict research progress and study results. Aside from these external limitations, there can also be internal limitations, such as gaps/limits in the current state of knowledge in a field (particularly as it relates to data collection). Any constraints, if they could affect the results of research, should be discussed in the methodology or discussion sections of the study.

Key questions to ask:

- Are there any major design flaws in this study?
- Are the data collection measures appropriate to answer the study questions?
- Were methodological limitations acknowledged and discussed?
- What influence might these limitations have had on the results?

Results

Nobody denies that reading a scientific study up to this point can be difficult and require focus. But then we finally get to the really interesting stuff: the answers. The results section of a study does indeed provide 'answers', what scientists would call data and the statistical analysis. Statistical measures can clearly and accurately convey the existence and strength of relationships observed in the study.

The field of statistics is based on the quantification of information. Descriptive statistics present the information in an organised way so that it is easier to interpret. Some of the more familiar descriptive statistics include percentage, frequency, mean, and standard deviation. Descriptive statistics, however, do not provide information about cause and effect; this requires the use of inferential statistics. As the name implies, inferential statistics often involve making inferences from the results for the sample studied and extrapolating them to a larger population.

Understanding statistical significance: Researchers generally calculate the statistical significance of differences observed between groups, and report it as a 'p-value'. A p-value is the probability of obtaining an effect or association in a study sample by chance. If the results of a study are statistically significant, then the study demonstrates some real association or effect. The study author should identify what p-value he or she has used in the analysis. A p-value of less than 5 percent ($p < 0.05$) is fairly common and would be considered statistically significant. This means that the result would occur less than 5 percent of the time by chance (if the treatment had no effect). More stringent levels of significance are $p < 0.01$ and $p < 0.001$.

If the results of a study are not statistically significant, the author may discuss the statistical power of the study. An in-depth discussion of statistical power is complex, however such information about a study will help the reader understand whether the study had a chance of finding the answer to the research questions in the first place.

It is easy to get wrapped up in discussions of statistical significance when reading research, but it is important to remember that a statistically significant result does not necessarily mean that the outcomes are important or relevant to the public. A statistically significant finding does not guarantee that the research is without bias or confounding factors that could make the statistical value irrelevant. Statistical significance is only part of the picture; to get the whole picture, one must consider the context of the study.

Communicating risk: Relative risk and absolute risk have different 'baselines', and muddle between these often leads to misunderstanding and miscommunication of research results.

Absolute risk refers to the actual risk of an occurrence: the chance that a specific outcome will occur. Relative risk puts risk in comparative terms: the outcome rate for people exposed to the factor in question, compared with the outcome rate for those not exposed to the factor. A relative risk of >1 indicates an increased risk of the outcome under investigation; one of <1 indicates a decreased risk of the outcome. Relative risks are the most commonly used measure of morbidity or mortality in the medical literature today. However, in many cases the absolute risk is a far more relevant statistic for the public.

For example, suppose that a study shows that a man who brushes his teeth only once a day is 50 percent more likely to have all his teeth fall out in the next 10 years, than others who brush their teeth twice per day. This is the relative risk. Yet the absolute risk that all of the man's teeth fall out may only be 1 percent. In this case, the relative risk makes the problem, a rare one anyway, seem more important than it really is. However relative risk can also make a problem appear less important than it actually is. Therefore, it is important to consider both relative risk and absolute risk when discussing study results.

Key questions to ask:

- What is the real and statistical significance of these results?
- To whom do these results apply?
- How do these results compare to those of other studies on the subject?

Discussion

The discussion section of a study gives the reader some insight into the study subject area and can shed new light on the results and their meaning. Alternative explanations for the results and the implications of the research may also be presented.

One of the most frequent errors in scientific research is drawing conclusions that are not adequately supported by the data. This may occur for a number of reasons: collection of insufficient or inadequate

data, over generalisation of results, methodological problems, or inherent limitations of the study design. This is why it is important to review the methodology section.

Sometimes researchers stray from the scientific method by reporting conclusions that are unrelated to the research question that was tested. Although conclusions made in this manner may have merit, it is important to take another look at whether the study was adequately designed and conducted to support the secondary conclusions.

Finally, be wary of absolute conclusions that profess to be the final word on a subject. Good research answers some questions and raises others. A call for more research to investigate particular issues that remain unclear or to replicate the study findings is a frequent conclusion in journal articles.

Key questions to ask:

- Are the conclusions supported by the data?
- Are the conclusions of the study related to the stated purpose of the study? If not, do the study design and results support the secondary conclusions?

References

Experts in the subject area can usually assess quickly if key research has been omitted from the reference list. If this is the case, the researchers may have failed to adequately review or evaluate prior research in the field that could have benefited their current study. A reference list that includes both older and newer relevant research can reassure the reader that the author has reviewed the entire body of research, and has not just considered the last few or first few studies conducted on the topic.

Also consider. . .

There are also other issues that merit attention in the critical review of studies, including the funding sources of a study and the appropriate use of editorials and letters to the editor.

Many words do not have the same meaning in a scientific context as in every day life. In the scientific context, "operational definitions" are used, which make the meaning of a term clear in the context of a particular study, but which may not be the universal or even the familiar meaning of the term. Terms used in a scientific context at times have a more restrictive meaning than what the lay public would understand by this term. This is important in order to avoid excessive generalisations.

Funding source: A study might be criticised, or its findings dismissed entirely, because industry or another interested party funded it. Many scientific journals today require that potential conflicts of interest be disclosed and sources of funding be referenced at the end of a paper. Although it is interesting to note the funding source of a study, it is unfair, and perhaps short-sighted, to simply negate results solely on the basis of the funding source.

The reason that studies are often funded by organisations that may benefit from the results is obvious. After all, who else but an interested party would allocate the large amounts of money that good research often requires? For example, when a company seeks approval for a new food ingredient, it is required by law to provide data to demonstrate the ingredient's safety. Government funding, from taxpayers, would not invest millions of Euros to study food ingredients or products that may never come to market.

Ethical researchers would not manipulate data or design studies to support funding interests. Indeed, most industry funders would not want a "tell them what they want to hear" researcher; they want real answers to their questions. A critical evaluation of research on its own merit is the best way to assess its validity and importance. If the study is good, its results will stand on their own, regardless of who supported the research.

Editorials and letters-to-the-editor: Editorials (commentary and opinion by experts in a field other than the authors of a study) are a very helpful way for readers to understand a study, its meaning, and its practical implications. Editorials often provide perspective on a study, discussing it in the context of other research, as well as identifying potential flaws that may affect the applicability or even the veracity of results.

Although letters to the editor usually appear in later issues of a journal following research publication, such letters can be very useful in the identification of problems with a study. They can be used as a continuing education tool on what to look for, when critically reviewing studies.

Study results that are reported via letters to the editor need to be considered with caution; they cannot substitute for peer-reviewed articles that provide full details necessary for readers to critically assess the research.

Summary

This EUFIC Review describes the assessment of food and health-related scientific studies, to help media, health professionals and educators assess original research. Such critical review is essential, to put study results into the context of other scientific literature on a subject and to accurately present the relevance of new research to the public.

Although the various sections of a study that have been discussed affect whether a piece of research can provide valid and relevant answers to a question being investigated, 'perfect' research does not exist. Economics, ethics, and gaps in the current state of knowledge may limit the ability of a study to find the answers sought. The scientific process is not linear but frequently moves in many different directions, generating questions, discussions, and debates along the way.

So how does the communicator manoeuvre through the maze of emerging scientific findings about food and health to deliver accurate, relevant information to the public? Firstly, by reserving judgement about a study until you have sought out other studies and experts to help assess the findings of the study and their importance—or unimportance. In other words, by putting new research into context. Secondly, by taking a moderate and cautious approach to communicating new information: what may seem to be a

revolutionary, life-altering study today, may turn out to be just the opposite tomorrow. The big steps and small steps and changes in direction are what make the journey of scientific research so exciting.

Key definitions

Bias – Problems in study design that can lead to effects that are not related to the variables being studied. An example is selection bias, which occurs when study subjects are chosen in a way that can misleadingly increase or decrease the strength of an association. Choosing experimental and control group subjects from different populations would result in a selection bias.

Blind, single or double – In a single-blind experiment, the subjects do not know whether they are receiving an experimental treatment or a placebo. In a double-blind experiment, neither the researchers nor the participants are aware of which subjects receive the treatment until after the study is completed.

Confounding variable or confounding factor – A 'hidden' variable that may cause an association that the researcher attributes to other variables.

Control group – The group of subjects in a study to whom a comparison is made to determine whether an observation or treatment has an effect. In an experimental study, it is the group that does not receive a treatment. Subjects are as similar as possible to those in the test group eg gender, age or health risk markers.

Correlation – An association (when one phenomenon is found to be accompanied by another) whose strength has been tested using statistical procedures. A correlation does not prove cause and effect.

Experimental group – The group of subjects in an experimental study that receives a treatment.

Generalisability – The extent to which the results of a study can be applied to the general population of people; demonstrating this requires an assessment of relevant features of a study population and whether these are comparable to those in other populations.

Incidence – The number of new cases of a disease during a given period of time in a defined population.

Meta-analysis – A quantitative technique in which the results of several individual studies are pooled to yield overall conclusions.

Outcomes research – Type of research that is increasingly used by the health industry, that provides information about how a specific procedure or treatment affects the subject (clinical safety and efficacy), the subject's physical functioning and lifestyle, and economic considerations such as saving or prolonging life and avoiding costly complications.

Placebo – Sometimes casually referred to as a 'sugar pill', a placebo is a 'fake' treatment that seems to be identical to the real treatment. Placebo treatments are used to eliminate bias that may arise from the

expectation that a treatment produces an effect.

Prevalence – The number of existing cases of a disease in a defined population at a specified time.

Prospective study – Epidemiological research that follows a group of people over time to observe the potential effects of diet, behaviour or other factors, on health or the incidence of disease. It is considered a more valid research design than retrospective research.

Randomisation, or random assignment – A process of assigning subjects to experimental or control groups in which the subjects have an equal chance of being assigned to either. Used to balance known, unknown, and difficult-to-control-for variables.

Random sampling – A method by which subjects are selected to participate in a study in which all individuals in a population have an equal chance of being chosen. This helps to ensure that the study results are generalisable.

Reliability – Whether the method used to collect data, such as a questionnaire, gives the same results, if repeated with the same person several times. A reliable test gives reproducible results.

Research design – How a study is set up to collect information or data. For valid results, the design must be appropriate to answer the question or hypothesis being studied.

Residual confounding – The effect that remains after attempts to statistically control for variables that cannot be measured perfectly. This is a particularly important concept in epidemiological studies because knowledge of human biology is incomplete. There may be unknown variables that could significantly change conclusions made on the basis of epidemiological research.

Retrospective study – Research that relies on recall of past data or on previously recorded information. This type of research is considered to have limitations because the number of variables cannot be controlled and because the memory of study participants is not always reliable.

Risk – A term encompassing a variety of measures of the probability of an outcome. It is usually used in reference to unfavourable outcomes such as illness or death. It is important to distinguish between absolute risk and relative risk:

- Absolute risk is defined as the chance of a person developing a specific outcome (e.g. disease) over a specified time-period. For example, a woman's lifetime absolute risk of breast cancer is one in nine. That is to say, one woman in every nine will develop breast cancer at some point in their lives. Absolute risk can also be expressed as a percentage (e.g. 11 per cent instead of one in nine) or as a decimal (0.11).
- Relative risk is used to compare risk in two different groups of people that differ in their exposure to a factor (e.g. alcohol intake). For example, researchers found that women who drank alcohol were at slightly increased risk of developing breast cancer compared with women who did not. If

we call the non-drinker's breast cancer risk '1', then a woman who drinks alcohol (e.g. 2-3 units per day) has a relative risk of 1.13, or expressed in other terms, has 13 per cent more chance to have breast cancer than a woman who does not drink alcohol.

Presentation of the results in terms of “relative risk” rather than “absolute risk” often makes the effects of the independent variable bigger than they actually are. Looking at the example above, how big is the increase in risk for the women who drink alcohol in absolute terms? An increase in relative risk of 13 per cent means that the absolute risk increases 1.43 per cent. Therefore, a woman who drinks two or three units of alcohol per day has an absolute lifetime risk of 12.43 per cent instead of 11 per cent ($11 + 1.43 = 12.43$) assuming she continues to drink alcohol at the same rate over her life time.

Risk factor – Anything statistically shown to have a relationship with the incidence of a disease. It does not necessarily mean a cause and effect.

Statistical power – A mathematical quantity that indicates the probability a study has of revealing a statistically significant effect. A high power of 80 percent, or 0.8, indicates that the treatment (or independent variable) used in the study, if conducted repeatedly, would reveal a statistically significant effect 80 percent of the time. On the other hand, a power of only 0.1 means that there would be a 90 percent chance that the research missed the effect – if one exists at all.

Statistical significance – The probability of revealing an effect or association in a study sample equal or higher than the one observed, if there was actually no effect in the population. In other words, a result is called “statistically significant” if it is unlikely to have occurred by chance. The significance of a result is also called its p-value; the smaller the p-value, the more significant the result is said to be. It works on the basis of the hypothesis that if there is no effect, the results of a treatment are unlikely to have occurred. A p-value of less than 5 percent ($p < 0.05$) means that the result would occur by chance less than 5 percent of the time, and is generally considered evidence of a true treatment effect or a true relationship. A “statistically significant difference” means there is statistical evidence that there is a difference; it does not mean the difference is necessarily large, important or significant in the usual sense of the word

Validity – The extent to which a study or investigation instrument measures what it is intended to measure. Refers to the accuracy or truthfulness in regard to a study's conclusions.

Variable – Any characteristic that may vary in study subjects, such as gender, age, body weight, diet or behaviour such as smoking. In an experiment, the treatment is called the independent variable; it is the factor being investigated. The variable that is influenced by the treatment is the dependent variable; it may change as a result of the effect of the independent variable.

Additional reading

1. Swinscow TDV. Statistics at Square One (9th ed) 1997, British Medical Association, London.
2. Coggon D, Rose G, Barker DJP. Statistics for the Uninitiated (4th ed) 1997, British Medical Association, London.

3. [WCRF/AICR \(2007\) Food, Nutrition, Physical Activity and the Prevention of Cancer - a Global Perspective.](#)
4. Langseth L. Nutritional Epidemiology: Possibilities and Limitations. ILSI Europe Concise Monograph Series; 1996.
5. Angell M, Kassirer J. Editorials and conflicts of interest. N Engl J Med. 1996;335(14):1055-1056.
6. Eastman M. Nutritional Epidemiology. Chapter in: Principles of Human Nutrition. 1997, Chapman & Hall, London.