Key Concepts for Informed Health Choices:
A framework for enabling people to think critically about health claims

Introduction

There are endless claims about treatments in the mass media, advertisements, and everyday personal communication. Some are true and some are false. Many are unsubstantiated: we do not know whether they are true or false. Unsubstantiated claims about the effects of treatments are often wrong. Consequently, people who believe and act on these claims suffer unnecessarily and waste resources by doing things that do not help and might be harmful, and by not doing things that do help.

In response to these challenges, we developed the IHC Key Concepts as the first step in the Informed Health Choices project, an initiative supported by the Research Council of Norway. The aim of the project is to help people make informed health choices.

A treatment is any intervention (action) intended to improve health, including preventive, therapeutic and rehabilitative interventions, and public health or health system interventions. Although we have developed and framed the Key Concepts to address treatment claims, people in other disciplines have also found them relevant; for example, for assessing claims about the effects of educational interventions or environmental measures. [1] Adaptations of the IHC Key Concepts for other disciplines can be found here.

The Informed Health Choices (IHC) Key Concepts

The concepts serve as the basis for developing learning resources to help people understand and apply the concepts when claims about the effects of treatments (and other interventions) are made, and when they make health choices.[2] They are also the basis for an item bank of multiple-choice questions (the Claim Evaluation Tools item bank) that can be used for assessing people's ability to apply the IHC Key Concepts.[3]

The IHC Key Concepts are principles for evaluating the trustworthiness of treatment claims, comparisons, and choices. The concepts can help people to:

1. Recognise when a **claim** about the effects of treatments has an untrustworthy basis
2. Recognise when evidence from **comparisons** of treatments is trustworthy and when it is not
3. Make well-informed **choices** about treatments

They can help anyone, not just researchers, to think critically about whether to believe a treatment claim and what to do. This is sometimes referred to as critical health literacy. We have not included concepts that are only relevant for researchers or that require a research background. The Key Concepts are intended for people using research, not for doing research.

What's new?

We started to develop this list of concepts in 2013. We published the first version of the list in 2015 (original version), with 32 concepts in six groups. We published revised lists in October 2016 with 34 concepts in three groups and in 2017 with 36 concepts. In 2018, the list included 44 concepts. We reorganised these within each of the three main groups and added three subgroups to each of the three main groups (Table 1).[4]

---

[1] Aronson JK, Barends E, Boruch R, et al. Key concepts for making informed choices. Nature 2019; 572:303-6.

[2] Chalmers I, Oxman AD, Austvoll-Dahlgren A, et al. Key Concepts for Informed Health Choices: A framework for helping people learn how to assess treatment claims and make informed choices. BMJ Evid Based Med 2018; 23:29-33.

[3] Austvoll-Dahlgren A, Semakula D, Nsangi A, et al. Measuring ability to assess claims about treatment effects: the development of the 'Claim Evaluation Tools'. BMJ Open 2017; 7:e013184.

[4] Oxman AD, Chalmers I, Austvoll-Dahlgren A and Informed Health Choices group. Key Concepts for assessing claims about treatment effects and making well-informed treatment choices. F1000Research 2019, 7:1784

Table 1. Nine high-level concepts

1. Claims *Claims about effects that are not supported by evidence from fair comparisons are not necessarily wrong, but there is an insufficient basis for believing them.*	**2. Comparisons** *Studies should make fair comparisons, designed to minimize the risk of systematic errors (biases) and random errors (the play of chance).*	**3. Choices** *What to do depends on judgements about a problem, the relevance of the evidence available, and the balance of expected benefits, harms, and costs.*
1.1 It should not be assumed that treatments are safe or effective - or that they are not. **1.2 Seemingly logical assumptions are not a sufficient basis for claims.** **1.3 Trust in a source alone is not a sufficient basis for believing a claim.**	**2.1 Comparisons of treatments should be fair.** **2.2 Syntheses of studies should be reliable.** **2.3 Descriptions should clearly reflect the size of effects and the risk of being misled by the play of chance.**	**3.1 Problems and options should be clear.** **3.2 Evidence should be relevant.** **3.3 Expected pros should outweigh cons.**

We did this partially in response to feedback that the organisation of concepts within the three main groups was not logical, and that having long lists of concepts was overwhelming. The subgroups of concepts provide a more transparent logic for how the concepts are organised in each main group. Having just three high level concepts for each group may also make it easier to get the gist of the concepts and make the list less overwhelming and easier to remember.

This year, we have added five new concepts:

- (1.2c) Assumptions that fair comparisons are not relevant can be misleading.
- (1.3a) Your own prior beliefs may be wrong.
- (3.3b) Consider the baseline risk or the severity of the symptoms when estimating the size of expected effects.
- (3.3c) Consider how important each advantage and disadvantage is when weighing the pros and cons.
- (3.3e) Important uncertainties about the effects of treatments should be reduced by further fair comparisons.

The list now includes 49 concepts (Table 2). In response to feedback, we have also edited the list of concepts to make their descriptions more consistent, and we have edited some of the explanations.

Competences and dispositions

In addition to modifying the Key Concepts in 2018, we added lists of competences (required skills, knowledge, or capacity to do something) and dispositions (frequent and voluntary habits of thinking and doing) for thinking critically about treatments. This year we have clarified our goal, increased the competences needed to achieve that goal from 10 to 20 (Table 3), and the dispositions from 10 to 15 (Table 4).

How we developed this list of Key Concepts

We developed the IHC Key Concepts by searching the literature and checklists written for the public, journalists, and health professionals; and by considering concepts related to assessing the certainty of evidence about the effects of treatments. We have revised the Key Concepts yearly, based on feedback and suggestions; and learning from using the IHC Key Concepts, other relevant frameworks, and adaptation of

the IHC Key Concepts by other disciplines. This year's update was based primarily on a systematic review of other relevant frameworks,[5] and adaptation of the IHC Key Concepts by other disciplines.[1]

We have tried to include all concepts that are important for people to consider when they assess treatment claims and make health choices. At the same time, we have tried to limit the number of concepts by minimising redundancy. We have organised the concepts in a way that makes sense to us and others who have provided feedback. They are not organised based on how complex or difficult they are to understand and apply, or in the order in which they should be learned.

Although we have written the concepts and explanations in plain language, some of them may be unfamiliar and difficult to understand. The list is not designed as a teaching tool. It is a framework, or starting point, for teachers, journalists and other intermediaries for identifying and developing resources (such as longer explanations, examples, games and interactive applications) to help people understand and apply the concepts.

When will the list of concepts next be updated?

A working group at the Centre for Informed Health Choices in Oslo review the list yearly. Astrid Dahlgren, Iain Chalmers, and Andy Oxman led the development of the original list of IHC Key Concepts and have amended it in the light feedback and suggestions since then. They are responsible for the final decisions about amendments and additions. Many other people have contributed to this work, including other members of the IHC team and people from around the world with various kinds of expertise. The next update will be made available in October 2020. Please send any comments or suggestions to: contact@informedhealthchoices.org.

Where you can find more information

More information about the IHC Key Concepts, their development, and their use can be found on the IHC website.

Suggested citation: Oxman AD, Chalmers I, Dahlgren A, and the Informed Health Choices Group. Key Concepts for assessing claims about treatment effects and making well-informed treatment choices. Version: 2019.

[5] Oxman AD, Martínez GL. Comparison of the Informed Health Choices Key Concepts to other frameworks that are relevant to learning how to think critically about treatment claims, comparisons, and choices: protocol for a mapping review. IHC Working Paper, 2018.

Table 2. Overview of the IHC Key Concepts

1. Claims

Claims about effects that are not supported by evidence from fair comparisons are not necessarily wrong, but there is an insufficient basis for believing them.

1.1 It should not be assumed that treatments are safe or effective - or that they are not.

a) Treatments can cause harms as well as benefits.
b) Large, dramatic effects are rare.
c) It is rarely possible to be certain about the effects of treatments.

1.2 Seemingly logical assumptions are not a sufficient basis for claims.

a) Treatment may not be needed.
b) Beliefs alone about how treatments work are not reliable predictors of the presence or size of effects.
c) Assumptions that fair comparisons of treatments in research are not applicable in practice can be misleading.
d) An outcome may be associated with a treatment but not caused by it.
e) More data is not necessarily better data.
f) Identifying effects of treatments depends on making comparisons.
g) The results of one study considered in isolation can be misleading.
h) Widely used treatments or those that have been used for decades are not necessarily beneficial or safe.
i) Treatments that are new or technologically impressive may not be better than available alternatives.
j) Increasing the amount of a treatment does not necessarily increase its benefits and may cause harm.
k) Earlier detection of 'disease' is not necessarily better.
l) It is rarely possible to know in advance who will benefit, who will not, and who will be harmed by using a treatment.

1.3 Trust in a source alone is not a sufficient basis for believing a claim.

a) Your existing beliefs may be wrong.
b) Competing interests may result in misleading claims.
c) Personal experiences or anecdotes alone are an unreliable basis for most claims.
d) Opinions alone are not a reliable basis for claims.
e) Peer review and publication by a journal do not guarantee that comparisons have been fair.

2. Comparisons

Studies should make fair comparisons, designed to minimize the risk of systematic errors (biases) and random errors (the play of chance).

2.1 Comparisons of treatments should be fair.

a) Comparison groups should be as similar as possible.
b) Indirect comparisons of treatments across different studies can be misleading.
c) The people being compared should be cared for similarly apart from the treatments being studied.
d) If possible, people should not know which of the treatments being compared they are receiving.
e) Outcomes should be assessed in the same way in all the groups being compared.
f) Outcomes should be assessed using methods that have been shown to be reliable.
g) It is important to assess outcomes in all (or nearly all) the people in a study.
h) People's outcomes should be counted in the group to which they were allocated.

2.2 Syntheses of studies need to be reliable.

a) Reviews of studies comparing treatments should use systematic methods.
b) Failure to consider unpublished results of fair comparisons may result in estimates of effects that are misleading.
c) Treatment claims based on models may be sensitive to underlying assumptions.

2.3 Descriptions should clearly reflect the size of effects and the risk of being misled by the play of chance.

a) Verbal descriptions of the size of effects alone can be misleading.
b) Relative effects of treatments alone can be misleading.
c) Average differences between treatments can be misleading.
d) Small studies may be misleading.
e) Results for a selected group of people within a study can be misleading.
f) The use of p-values may be misleading; confidence intervals are more informative.
g) Deeming results to be "statistically significant" or "nonsignificant" can be misleading.
h) Lack of evidence of a difference is not the same as evidence of "no difference".

3. Choices

What to do depends on judgements about a problem, the relevance of the evidence available, and the balance of expected benefits, harms, and costs.

3.1 Problems and options should be clear.

a) Be clear about what the problem or goal is and what the options are.

3.2 Evidence should be relevant.

a) Attention should focus on all important effects of treatments, and not surrogate outcomes.
b) Fair comparisons of treatments in animals or highly selected groups of people may not be relevant.
c) The treatments compared should be similar to those of interest.
d) There should not be important differences between the circumstances in which the treatments were compared and those of interest.

3.3 Expected advantages should outweigh expected disadvantages.

a) Weigh the benefits and savings against the harms and costs of acting or not.
b) Consider the baseline risk or the severity of the symptoms when estimating the size of expected effects.
c) Consider how important each advantage and disadvantage is when weighing the pros and cons.
d) Consider how certain you can be about each advantage and disadvantage.
e) Important uncertainties about the effects of treatments should be addressed in further fair comparisons.

Table 3. Goals, Competences and Dispositions for Informed Health Choices

Goal

To enable people to make good decisions* about which claims to believe about the effects of things they can do for their health, the health of others or for other reasons, and about what to do to achieve their goals.

Competences

To achieve this goal, people should be able to:

1. **Recognise when a claim has an untrustworthy basis by:**
   a) recognising claims about the effects of treatments
   b) questioning the basis for treatment claims
   c) thinking carefully about treatment claims before believing them
   d) recognising when a treatment claim is relevant and important, and warrants reflection

2. **Recognise when evidence used to support a treatment claim is trustworthy or untrustworthy by:**
   a) recognising the assumptions, evidence and reasoning behind treatment claims
   b) recognising unfair treatment comparisons
   c) recognising unreliable summaries of treatment comparisons
   d) recognising when a statistical model and its assumptions are used to support a treatment claim
   e) recognising misleading ways of presenting treatment effects
   f) understanding how systematic errors (the risk of bias), random errors (the play of chance), and the relevance (applicability) of treatment comparisons can affect the degree of confidence in estimates of treatment effects
   g) understanding the extent to which evidence does or does not support a treatment claim

3. **Make well-informed decisions about treatments by:**
   a) being aware of cognitive biases when making decisions
   b) clarifying and understanding the problem, options, and goals when making a decision
   c) recognising when decisions have irreversible consequences
   d) judging the relevance of evidence used to inform decisions about treatments
   e) weighing the advantages and disadvantages of treatments, taking into account the size of treatment effects, how important each outcome is, the costs, and the certainty of the evidence
   f) communicating with others about the advantages and disadvantages of treatments

4. **Reflect on people's competences and dispositions by:**
   a) monitoring how they decide which treatment claims to believe and what to do
   b) monitoring how people adjust the processes they use to decide what to believe and do to fit the relevance, importance, and nature of different types of treatment claims and choices
   c) being aware of when people are making treatment claims themselves

* A good decision is one that makes effective use of the information available to the decision maker at the time the decision is made. A good outcome is one that the decision maker likes. The aim of thinking critically about treatments is to increase the probability of good outcomes (and true conclusions), but many other factors affect outcomes aside from critical thinking.[6]

[6] Baron J. Thinking and deciding (4th ed). Cambridge, UK: Cambridge University Press, 2008.

Table 4. Dispositions

People should be in the habit of thinking critically about:

1. **Claims by**
   a) being aware of treatment claims (including those you make yourself) and choices
   b) questioning the basis for treatment claims
   c) being aware of cognitive biases and going from fast to slow thinking before forming an opinion about a treatment claim, making a claim, or taking a decision
   d) seeking evidence to reduce uncertainty when considering a relevant and important treatment claim or decision

2. **Evidence used to support claims by:**
   a) questioning the trustworthiness of evidence used to support treatment claims
   b) being alert to misleading presentations of treatment effects
   c) acknowledging and accepting uncertainty about the effects of treatments
   d) being willing to admit errors and modify their judgements when warranted by evidence or a lack of evidence

3. **Choices by:**
   a) clarifying and understanding the problem, options, and goals when making decisions about treatments
   b) preferring evidence-based sources of information about treatment effects
   c) considering the relevance of the evidence used to inform decisions about treatments
   d) considering effect estimates, baseline risk, the importance of each advantage and disadvantage, the costs, and the certainty of the evidence when making decisions about treatments
   e) making informed judgements about the certainty of estimates of treatment effects
   f) making well-informed decisions
   g) Being aware of how people decide which treatment claims to believe and what to do

4. **People's own thinking by:**
   a) Being aware of how people decide which treatment claims to believe and what to do

1. Claims

Claims about effects that are not supported by evidence from fair comparisons are not necessarily wrong, but there is an insufficient basis for believing them.

Concepts	*Short titles*[*]	Explanations	Implications
1.1 It should not be assumed that treatments are safe or effective - or that they are not.			
	Too good to be true		
a) Treatments can cause harms as well as benefits.	*"100% safe!"*	People often exaggerate the benefits of treatments and ignore or downplay potential harms. However, few effective treatments are 100% safe. Similarly, people in need or desperation hope that treatments will work and ignore potential harms. As a result, they may waste time and money on treatments that have never been shown to be useful and may cause harm.	Always consider the possibility that a treatment may have harmful effects.
b) Large, dramatic effects are rare.	*"100% effective!"*	Large effects (where everyone or nearly everyone treated experiences a benefit or a harm) are easy to detect without fair comparisons, but few treatments have effects that are so large that fair comparisons (designed to reduce the effects of biases and the play of chance) are not needed. Treatments that do not have large, dramatic effects may be helpful, but fair comparisons are needed to determine how safe and helpful they are.	Claims of large effects are likely to be wrong. Expect treatments to have moderate, small or trivial effects, rather than dramatic effects. Do not rely on claims of small or moderate effects of a treatment, which are not based on systematic reviews of fair comparisons of treatments.
c) It is rarely possible to be certain about the effects of treatments.	*"100% certain!"*	Fair comparisons of treatments can provide a basis for being confident about the probability of beneficial and harmful effects of treatments. However, it is rarely, if ever, possible to be 100% certain about the size of treatment effects or to know exactly what will happen if a treatment is used. This is especially true for treatments that are intended to prevent something happening a long time in the future. Fair comparisons of such treatments - for example changes in diet or exercise - are difficult, because people need to be followed-up for a very long time and it is difficult to ensure that people adhere to whatever advice they are given. Consequently, claims about the effects of such treatments are often based on associations and explanations. Some people argue that there should be different standards for judgements about the trustworthiness of claims when fair comparisons are difficult. However, it is dishonest not to acknowledge uncertainty, even when there are important limitations on the potential to reduce that uncertainty.	Recognise that there is some uncertainty about the effects of nearly all treatments, and that there is likely to be more uncertainty about some types of treatments. Choices still must be made, but it is better to acknowledge and accept uncertainty than to deny it and make misinformed or poorly informed decisions.
1.2 Seemingly logical assumptions are not a sufficient basis for claims.			
a) Treatment may not be needed.	*"Treatment needed!"*	Effective treatments can prevent health problems and premature death, and improve the quality of life. However, nature is a great healer and people often recover from illness without treatment. Likewise, some health problems may get worse despite treatment, or treatment may actually make things worse. Not using a treatment is not the same as "no treatment". Waiting to see what happens ("letting nature take its course"), with or without treating symptoms such as pain, is a treatment option.	Always consider the usual course of a health problem when considering treatments other than waiting to see what happens. Sometimes treatment is not needed and may even make things worse.

[*] Used on That's a claim! Key Concepts for thinking critically about health claims

Concepts	*Short titles**	Explanations	Implications
b) Beliefs alone about how treatments work are not reliable predictors of the presence or size of effects.	*"It works like this!"*	Treatments that should work in theory often do not work in practice, or may turn out to be harmful. A plausible explanation of how or why a treatment might work does not prove that it actually does work, or that it is safe. And even if there is plausible evidence that a treatment works in ways likely to be beneficial, the size of any such treatment effect, and its safety, cannot be predicted. For example, most drugs in a class of heart medicines called beta-blockers have beneficial effects in reducing recurrence of heart attacks; but one of the drugs in the class – practolol – caused unpredicted serious complications in patients' eyes and abdomens. Similarly, it cannot be assumed that a treatment works or does not work based on the type of treatment; for example, assuming that all alternative medicines or that all modern medicines do or do not work, or that all vaccines do or do not work. On the other hand, not understanding how a treatment works does not mean that it does not work.	Do not assume that claims about the effects of treatments based on an explanation of how they might work are correct if the treatments have not been assessed in systematic reviews of fair comparisons of treatments.
c) Assumptions that fair comparisons of treatments in research are not applicable in practice can be misleading.	*"Not relevant!"*	People often claim that evidence from fair comparisons of treatments cannot be applied to people who are not exactly like the participants in studies. This is only likely to be true if there are important differences. That is, unless there are compelling reasons why you or the people of interest are so different from the study participants that the treatments cannot work in the same way, the effects of treatments are unlikely to differ substantially. It should be noted that most often the relative effect will be similar. Differences in baseline risk will, however, often lead to differences in the absolute effect.	Do not assume fair comparisons are not applicable because of differences between study participants and the people of interest, unless there are compelling reasons why treatments would work differently.
d) An 'outcome' may be associated with a treatment, but not caused by it.	*"Associated with!"*	The fact that a possible treatment outcome (i.e. a potential benefit or harm) is associated with a treatment does not mean that the treatment caused the outcome. The association or correlation could instead be due to chance or some other underlying factor. For example, people who seek and receive a treatment may be healthier and have better living conditions than those who do not seek and receive the treatment. Therefore, people receiving the treatment might appear to benefit from the treatment, but the difference in outcomes could be because they are healthier and have better living conditions, rather than because of the treatment.	Do not assume that an outcome was caused by a treatment unless other reasons for an association have been ruled out by a fair comparison.
e) More data is not necessarily better data.	*"Lots of data!"*	Claims that are based on "big data" (data from large databases) or "real world data" (routinely collected data) can be misleading. More data simply gives a more statistically precise estimate of whatever biases there might be in a treatment comparison that uses routinely collected data. When using routinely collected data It is only possible to control for confounders that are already known and have been measured. Unfortunately, routinely collected data often do not include sufficient detail to confidently conclude that any association found between a treatment and an outcome means that the treatment caused the outcome. Describing routinely collected data as "real world data" implies that data collected in carefully designed fair comparisons of treatments do not come from the real world. Databases of routinely collected data may indeed include a broader spectrum of people than data collected in fair comparisons of treatments that have narrow eligibility criteria. However, routine collection of data is rarely planned to include the information that is needed to ensure fair comparisons, and these can be designed to have wide eligibility criteria.	Do not assume that an association between a treatment and an outcome found using "big data" or "real world data" means that the treatment caused the outcome unless other reasons for the association have been ruled out.

Concepts	*Short titles**	Explanations	Implications
f) Identifying effects of treatments depends on making comparisons.	*"No comparison needed!"*	Unless a treatment is compared to something else, it is not possible to know what would happen without the treatment, so making it difficult to attribute outcomes to the treatment. Whenever comparative terms are used to describe a treatment – for example "fastest relief" or "best" - ask "compared to what?" Sometimes people argue that a fair comparison of a treatment is not possible, for example because it is 'holistic', 'individualised', or 'complex'. However, as with any other treatment, claims about the effects of such treatments depend on comparing them to something else. How trustworthy those claims are depends on how fair the comparisons were.	Always ask what the comparisons are when considering claims about the effects of treatments. Claims that are not based on fair comparisons are not reliable.
g) The results of one study considered in isolation can be misleading.	*"A study shows!"*	A single comparison of treatments rarely provides conclusive evidence; and results are often available from other comparisons of the same treatments. Systematic reviews of all the similar comparisons (replications) may yield different results from those based on the initial studies, and these should help to provide more reliable and precise estimates of treatment differences. Even so, obtaining reliable estimates from treatment comparisons must always consider that important studies may remain unpublished, or inaccessible for other reasons.	The results of single comparisons of treatments can be misleading. Consider all the relevant fair comparisons.
h) Widely used treatments or those that have been used for decades are not necessarily beneficial or safe.	*"Old is better!"*	Treatments that have not been properly evaluated but are widely used or have been used for a long time are often assumed to work. Sometimes, however, they may be unsafe or of doubtful benefit.	Do not assume that treatments are beneficial or safe simply because they are widely used or have been used for a long time, unless this has been shown in systematic reviews of fair comparisons of treatments.
i) Treatments that are new or technologically impressive may not be better than available alternatives.	*"New is better!"*	New treatments are often assumed to be better simply because they are new, more expensive, or technologically impressive. However, on average, they are only very slightly likely to be better than other available treatments. Some side effects of treatments, for example, take time to appear and without long term follow-up it may not be possible to know whether they will appear.	Do not assume that a treatment is better or safer simply because it is new, brand-named, expensive, or technologically impressive.
j) Increasing the amount of a treatment does not necessarily increase its benefits and may cause harm.	*"More is better!"*	Increasing the dose or amount of a treatment (e.g. how many vitamin pills you take) often increases harms without increasing beneficial effects.	If a treatment is believed to be beneficial, do not assume that more of it is better.
k) Earlier detection of 'disease' is not necessarily better.	*"Early is better!"*	People often assume that early detection of disease and 'treating' people who are at statistical risk of disease lead to better outcomes. However, screening people to detect disease or treating people at statistical risk of disease is only helpful if two conditions are met. First, there must be an effective treatment. Second, people who are treated before the disease becomes apparent must do better than people who are treated after the disease becomes apparent. Screening and treating people at statistical risk of a disease can lead to overdiagnosis and overtreatment. Screening tests can be inaccurate (e.g. misclassifying people who do not have a disease as if they do have the disease). Screening or treating a statistical risk factor as if it is a 'disease' can also cause harm by labelling people as being sick when they are not, and because of side effects of the tests and treatments.	Do not assume that early detection of disease is worthwhile if it has not been assessed in systematic reviews of fair comparisons between people who were screened and people who were not screened.

Concepts	*Short titles**	Explanations	Implications
l) It is rarely possible to know in advance who will benefit, who will not, and who will be harmed by using a treatment.	*"Personalised medicine!"*	For some kinds of health problems, fair treatment comparisons can be made by giving different treatments to a patient at different times, and then comparing the outcomes associated with each of the different treatment periods. These are called n-of-1 trials because they compare the effects of alternative treatments in one specific patient. Most uncertainties about the effects of treatments cannot be compared in this way, however. For example, this person-specific approach could not be used to compare a surgical treatment with a drug treatment. Most treatment comparisons involve comparing similar groups of patients assigned to alternative treatments. Fair comparisons of treatments usually tell us what happened, on average, in groups of similar people. Treatments very rarely have exactly the same effect on everyone, however. Usually, in a group of people that have used a treatment, some benefit, some do not, and some may even be harmed. It is rarely possible to know in advance which people will benefit from which treatment, which will not benefit, or who will be harmed. Paradoxically, the only way to know whether "personalised medicine"- customising treatment for individuals works - is to test it in fair comparisons. Unless the customisation is 100% effective and 100% safe, it is still not possible to know in advance who will benefit from "personalised care" and who will not. Beyond n-of-1 trials, "personalised medicine" is not really personalised; it is simply an effort to identify subgroups of people who are most likely to benefit from specific treatments.	Fair treatment comparisons provide the best basis for making well-informed decisions about treatments, but there is almost always some uncertainty about who will benefit, who will not, and who will be harmed.
1.3 Trust in a source alone is not a sufficient basis for believing a claim.			
a) Your existing beliefs may be wrong.	*"Just as I thought!"*	People often look for and use information to support their own beliefs, including beliefs about the effects of treatments. This is sometimes called 'confirmation bias'. Confirmation bias can occur when people want a claim about treatment effects to be true. By focussing on evidence or arguments that support their existing beliefs and ignoring evidence or arguments that challenge these, people believe claims that confirm what they believe or wanted to be true without thinking critically about the basis for the claims.	Don't be misled by your own beliefs or rely on them, unless they are based on the results of systematic reviews of fair comparisons of treatments.
b) Competing interests may result in misleading claims.	*"As advertised!"*	People with an interest in promoting a treatment (in addition to wanting to help people) - for example, to make money - may promote treatments by exaggerating benefits, ignoring potential harmful effects, cherry picking which information is used, or making false claims. Conversely, people may be opposed to a treatment for a range of reasons, such as cultural practices.	Ask if people making claims that a treatment is effective have conflicting interests. If they do, be careful not to be misled by their claims about the effects of treatments.

Concepts	*Short titles**	Explanations	Implications
c) Personal experiences or anecdotes alone are an unreliable basis for most claims.	*"It worked for me!"*	People often believe that improvements in a health problem (for example, recovery from a disease) resulted from having received a treatment. Similarly, they might believe that an undesirable health outcome was due to having received a treatment. However, the fact that an individual recovered after receiving a treatment does not mean that the treatment caused the improvement, or that other people receiving the same treatment will also improve. The improvement (or the undesirable health outcome) might have occurred even without treatment. One reason that personal experiences - including a series of personal experiences - are sometimes misleading is that experiences, such as pain, fluctuate and tend to return to a more normal or average level. This is sometimes referred to as "regression to the mean". For example, people often treat symptoms such as pain when they are very bad and would improve anyway without treatment. The same applies to a series of experiences. For example, if there is a spike in the number of traffic crashes someplace, traffic lights may be installed to reduce these. A subsequent reduction may leave the impression that the traffic lights caused this change. However, it is possible that the number of crashes would have returned to a more normal level without the traffic lights.	If an individual improved after receiving a treatment it does not necessarily mean that the treatment caused the improvement, or that other people receiving the same treatment will also improve.
d) Opinions alone are not a reliable basis for claims.	*"Recommended by experts!"*	People often disagree about the effects of treatments, including doctors, researchers, and patients. This may be because their opinions are not always based on systematic reviews of fair comparisons of treatments. Who makes a treatment claim, how likable they are, or how much experience and expertise they have are not a reliable basis for assessing how reliable their claim is. This does not mean that conflicting opinions should be given equal weight - or that the existence of conflicting opinions means that no conclusion can be reached. How much weight to give an opinion should be based on the strength of the evidence supporting it.	Do not rely on the opinions of experts or other authorities about the effects of treatments, unless they have taken account of the results of systematic reviews of fair comparisons of treatments.
e) Peer review and publication by a journal do not guarantee that comparisons have been fair.	*"Peer reviewed!"*	Even though a comparison of treatments – whether in a single study or in a review of similar studies - has been published in a prestigious journal, it may not be a fair comparison and the results may not be reliable. Peer review (assessment of a study by others working in the same field) does not guarantee that published studies are reliable. Assessments vary and may not be systematic. Similarly, just because a study is widely publicised does not mean that it is trustworthy.	Always consider whether a published comparison of the effects of treatments is fair and whether the results are reliable. Peer-review is a poor indicator of reliability.

2. Comparisons

Studies should make fair comparisons, designed to minimize the risk of systematic errors (biases) and random errors (the play of chance).

Concepts	*Short titles**	Explanations	Implications
2.1 Comparisons of treatments should be fair.			
a) Comparison groups should be as similar as possible.	*Dissimilar comparison groups*	If people in the treatment comparison groups differ in ways other than the treatments being compared, the apparent effects of the treatments might reflect those differences rather than actual treatment effects. Differences in the characteristics of the people in the comparison groups at the beginning of the comparison might result in estimates of treatment effects that appear either larger or smaller than they actually are. A method such as allocating people to different treatments by assigning them random numbers (the equivalent of flipping a coin) is the best way to ensure that the groups being compared are similar in terms of both measured and unmeasured characteristics. If people are not randomly allocated to treatment comparison groups, differences between the groups other than the treatments may result in estimates of treatment effects appearing larger or smaller than they actually are because of confounders or other differences. For example, patients who are most ill (e.g. have severe pain) may be more likely to be given a new treatment than patients who are less ill. There may appear to be a sharp response to treatment in the most ill patients because of regression to the mean. If they are compared to patients who are less ill and receive an older treatment, the new treatment may appear to be more effective than it actually is compared to the older treatment. Differences in recall ("recall bias") can also lead to over- or under-estimates of effects in case-control and retrospective cohort studies that are based on recollection of exposure to the treatment.	Be cautious about relying on the results of non-randomized treatment comparisons (for example, if the people being compared chose which treatment they received). Be particularly cautious when you cannot be confident that the characteristics of the comparison groups are similar. If people were *not* randomly allocated to treatment comparison groups, ask if there were important differences between the groups that might have resulted in the estimates of treatment effects appearing either larger or smaller than they actually are.
b) Indirect comparisons of treatments across different studies can be misleading.	*Indirect comparisons*	For many conditions (e.g. depression) there are more than two possible treatments (for example, different drugs, or types of psychotherapy). Only very rarely are all the possible treatments for a condition compared in a single study, so it may be necessary to consider indirect comparisons among treatments. For example, there may be comparisons of drug A with placebo and comparisons of drug B with placebo, but no studies that compare drug A with drug B directly. In this case indirect comparisons among studies may be needed to inform a decision about whether to use drug A or drug B. However, there can be important differences between the studies examined in addition to the treatments they assessed, for example, differences in characteristics of the participants, or the way the comparisons were done, or in the outcome measures used. These differences can result in misleading estimates of treatment effects.	Indirect comparisons are sometimes needed to inform treatment choices. In these circumstances, careful consideration should be given to differences between the studies besides the treatments that were compared.
c) The people being compared should be cared for similarly apart from the treatments being studied.	*Dissimilar care*	If people in one treatment comparison group receive more care and attention than people in the comparison group, differences in outcomes may reflect differences in the amount of attention patients in each group has received rather than due to the treatments being compared. One way of preventing this is to keep treatment providers unaware of ("blind" to) which people have been allocated to which treatment.	Be cautious about relying on the results of treatment comparisons if people in the groups that are being compared were not cared for similarly (apart from the treatments being compared). The results of such comparisons can be misleading.

Concepts	*Short titles**	Explanations	Implications
d) If possible, people should *not* know which of the treatments being compared they are receiving.	*Dissimilar expectations*	People in a treatment group may behave differently or experience improvements or deterioration as a result of knowing the treatment to which they have been assigned. If this phenomenon is associated with an improvement in their symptoms it is known as a placebo effect; if it is associated with a harmful effect it is known as a nocebo effect. If individuals know that they are receiving a treatment that they believe is either better or worse than an alternative (that is, they are not "blinded"), some or all of the apparent effects of treatments may be due either to placebo or nocebo effects..	Be cautious about relying on the results of treatment comparisons if the participants knew which treatment they had received. This may have affected their expectations or behaviour. The results of such comparisons can be misleading.
e) Outcomes should be assessed in the same way in all the groups being compared.	*Dissimilar measurement*	If a possible treatment outcome is assessed differently in two treatment comparison groups, differences in that outcome may be due to *how* the outcome was assessed rather than *because* of the treatments received by people in each group. For example, if outcome assessors believe that a particular treatment works and they know which patients have received that treatment, they may be more likely to record better outcomes in those who have received the treatment. One way of preventing this is to keep outcome assessors unaware of ("blind" to) which people have been allocated to which treatment. This precaution is less important for "objective" outcomes, like death, than for "subjective" outcomes, like pain.	Be cautious about relying on the results of treatment comparisons if outcomes were not assessed in the same way in the different treatment comparison groups. The results of such comparisons can be misleading.
f) Outcomes should be assessed using methods that have been shown to be reliable.	*Unreliable assessment of outcomes*	Some outcomes are easy to assess, such as births and deaths. Others are more difficult, such as depression or quality of life. Treatment comparisons to be meaningful, outcomes that are meaningful to people should be assessed using outcomes and methods that have been shown to be reliable.	Be cautious about relying on the results of treatment comparisons if outcomes have not been assessed using methods that have been shown to be reliable.
g It is important to assess outcomes in all (or nearly all) the people in a study.	*Lots of missing people*	People in treatment comparisons who are not followed up to the end of the study may have worse outcomes than those who completed follow up. For example, they may have dropped out because the treatment was not working or because of side effects. If those people are excluded from the comparison, the findings of the study may be misleading.	Be cautious about relying on the results of treatment comparisons if many people were lost to follow-up, or if there was a big difference between the comparison groups in the proportions of people lost to follow-up. The results of such comparisons can be misleading.
h) People's outcomes should be counted in the group to which they were allocated.	*Outcomes counted in the wrong group*	Random allocation to treatment comparison groups helps to ensure that people in the comparison groups have similar characteristics before they receive treatment. However, people sometimes do not receive or take the treatment allocated to them. The characteristics of such people often differ from those who do take the treatments allocated to them. Excluding from the analysis people who did not receive the treatments allocated to them may mean that like is no longer being compared with like. This may lead to an underestimate of treatment differences relative to what would have been the case if everyone had received treatment that had been intended for them.	Be cautious about relying on the results of treatment comparisons if patients' outcomes have not been counted in the group to which they were allocated. For example, in a comparison of surgery and drug treatments, people who die while waiting for surgery should be counted in the surgery group, even though they did not receive surgery.

Concepts	*Short titles**	Explanations	Implications
2.2 Syntheses of studies should be reliable.			
a) Reviews of studies comparing treatments should use systematic methods.	*Unsystematic summary*	Reviews that do not use systematic methods may result in biased or imprecise estimates of the effects of treatments because the selection of studies for inclusion may be biased, or the methods may result in some studies not being found. In addition, the appraisal of the quality of some studies may be biased, or the synthesis of the results of the selected studies may be inadequate or inappropriate. To avoid these problems, systematic reviews of fair comparisons begin with protocols, which should be registered and searchable in registries such as Prospero. Even reviews that purport to be systematic may not be.	Whenever possible, use up-to-date systematic reviews of fair comparisons to inform decisions rather than non-systematic reviews of fair comparisons of treatments.
b) Failure to consider unpublished results of fair comparisons may result in estimates of effects that are misleading.	*Selective reporting*	Many fair comparisons are never published, and outcomes are sometimes left out. Those that are published are more likely to report favourable results. Consequently, reliance on published reports alone sometimes results in the beneficial effects of treatments being overestimated and the adverse effects being underestimated. Biased under-reporting of research is a major problem that is far from being solved. It is scientific and ethical malpractice, and wastes research resources. Selective reporting is an important reason why fair comparisons of treatments should begin with protocols that are registered and searchable in registries such as clinicaltrials.gov. This can also help to reduce selective reporting of some outcomes but not others in published reports, depending on the nature and direction of the results.	Be aware of the possibility of biased underreporting of fair comparisons, and whether or not the authors of systematic reviews have addressed this risk
c) Treatment claims based on models may be sensitive to underlying assumptions.	*Unfounded assumptions*	Sometimes treatment claims are based on chains of evidence, or models. For example, the effects of using a diagnostic test may depend on how accurate the test is, assumptions about what will be done based on the its results, and evidence of the effects of what is done. Similarly, evidence of the effects of public health and health system policies sometimes comes from models that combine different types of studies and assumptions; and assumptions are sometimes made when fair comparisons are combined in systematic reviews. When treatment comparisons depend on assumptions, it is important to consider their basis and to test how sensitive the results are to plausible changes in the assumptions made. For example, a model used to compare the effects of using different diagnostic tests on outcomes that are important to patients might require assumptions about what actions doctors or patients will take based on test results. If that is uncertain, it is important to consider whether changing the assumptions has a substantial impact on the estimated difference in outcomes important to patients.	Whenever treatment comparisons depend on assumptions, consider whether the assumptions are well-founded and how sensitive the results are to plausible changes in the assumptions that are made.
2.3 Descriptions should clearly reflect the size of effects and the risk of being misled by the play of chance.			
a) Verbal descriptions of the size of effects alone can be misleading.	*Just words*	A treatment effect (a difference in outcomes in a comparison) is a numerical concept, but it is difficult for some people to understand quantitative information about the effects of treatments. Qualitative (descriptive) labels may be easier to understand and can be helpful. However, qualitative descriptions of effects mean different things to different people, for example, saying that a treatment will 'slightly reduce', 'reduce', or 'greatly reduce' the likelihood of an undesirable outcome; or that a side effect is 'frequent' or 'rare'. In addition, verbal descriptions of treatments can be manipulative; e.g. promising 'amazing results' or describing treatments as 'natural', implying that they are safe because of that.	A verbal description of a treatment effect can be helpful, but it should be considered together with quantitative information about the size of the effect. Be wary of manipulative use of language in descriptions of treatment effects.

Concepts	*Short titles**	Explanations	Implications
b) Relative effects of treatments alone can be misleading.	*Relative effects*	Relative effects (for example, the ratio of the probability of an outcome in one treatment group compared with that in a comparison group) are insufficient for judging the importance of the difference (between the frequencies of the outcome). A relative effect may give the impression that a difference is more important than it actually is when the likelihood of the outcome is small to begin with. For example, if a treatment reduces the probability of getting an illness by 50% but also has harms, and the risk of getting the illness is 2 in 100, receiving the treatment may be worthwhile. If, however, the risk of getting the illness is 2 in 10,000, then receiving the treatment may not be worthwhile even though the *relative* effect is the same. The absolute effect of a treatment is likely to vary for people at different baseline risk.	Always consider the absolute effects of treatments – that is, the difference in outcomes between the treatment groups being compared. Do not make a treatment decision based on relative effects alone.
c) Average differences between treatments can be misleading.	*Average effects*	Average effects do not apply to everyone. For outcomes that are assessed using scales (for example, to measure weight, or pain) the difference between the average among people in one treatment group and the average among those in a comparison group may not make it clear how many people experienced a big enough change (for example, in weight or pain) for them to notice it, or that they would regard as important. In addition, many scales are difficult to interpret and are reported in ways that make them meaningless. This includes not reporting the lower and upper 'anchor', for example, whether a scale goes from 1 to 10 or 1 to 100; whether higher numbers are good or bad; and whether someone experiencing an improvement of, say, 5 on the scale would barely notice the difference, would consider it a meaningful improvement, or would consider it a large improvement.	When outcomes are assessed using scales, it cannot be assumed that every individual in the treatment comparison groups experienced the average effect. Be wary of differences on scales that are not explained or easily understood.
d) Small studies may be misleading.	*Few people or events*	When there are few outcome events, differences in outcome frequencies between the treatment comparison groups may easily have occurred by chance and may mistakenly be attributed to differences in the effects of the treatments.	Be cautious about relying on the results of treatment comparisons with few outcome events. The results of such comparisons can be misleading.
e) Results for a selected group of people within a study can be misleading.	*Subgroup analyses*	Average effects do not apply to everyone. However, comparisons of treatments often report results for selected groups of participants in an effort to assess whether the effect of a treatment is different for different types of people (e.g. men and women or different age groups). These analyses are often poorly planned and reported. Most differential effects suggested by these "subgroup results" are likely to be due to the play of chance and are unlikely to reflect true treatment differences.	Findings based on results for subgroups of people within treatment comparisons may be misleading.
f) The use of p-values may be misleading; confidence intervals are more informative.	*No confidence interval*	The observed difference in outcomes is the best estimate of how relatively effective and safe treatments are (or would be, if the comparison were made in many more people). However, because of the play of chance, the true difference may be larger or smaller than this. The confidence interval is the range within which the true difference is likely to lie, after taking into account the play of chance. Although a confidence interval (margin of error) is more informative than a p-value, the latter is often reported. P-values are often misinterpreted to mean that treatments have or do not have important effects.	Understanding a confidence interval may be necessary to understand the reliability of estimates of treatment effects. Whenever possible, consider confidence intervals when assessing estimates of treatment effects. Do not be misled by p-values.

Concepts	*Short titles**	Explanations	Implications
g) Deeming results to be statistically significant or "nonsignificant" can be misleading.	*Statistically significant*	"Statistical significance" is often confused with "importance". The cut-off for considering a result as statistically significant is arbitrary, and statistically non-significant results can be either informative (showing that it is very unlikely that a treatment has an important effect) or inconclusive (showing that the relative effects of the treatments compared are uncertain).	Claims that results were 'significant' or 'non-significant usually mean that they were 'statistically significant' or 'statistically non-significant'. This is not the same as 'important' or 'not important'. Do not be misled by such claims.
h) Lack of evidence of a difference' is not the same as 'evidence of no difference"	*No evidence*	Systematic reviews sometimes conclude that there is "no evidence of a difference" when there is uncertainty about the difference between the effects of treatments. This is often misinterpreted as meaning that there is "no difference" between the treatments compared. However, studies can never show that there is "no difference" ("no effect"). They can only rule out, with specific degrees of confidence, differences of a specific size.	Don't be misled by statements of "no difference" between treatments ("no effect"). Consider instead the degree to which it is possible to confidently rule out a difference of a specified size.

3. Choices

What to do depends on judgements about a problem, the relevance of the evidence available, and the balance of expected benefits, harms, and costs.

Concepts	*Short titles**	Explanations	Implications
3.1 Problems and options should be clear.			
a) Be clear about what the problem or goal is and what the options are.	*What is your problem and what are your options?*	Good decisions depend on correctly identifying the problems and considering an appropriate set of options to address the problems. For personal health choices, this means starting with a correct diagnosis (or assessment of risk) and then identifying the treatments that are available. For public health and health system policy decisions, this means describing the problem correctly and identifying the policy options relevant for that problem. Changing how a problem is framed can lead to different options for addressing it.	Make sure you are considering the correct diagnosis or problem, and appropriate options for addressing it.
3.2 Evidence should be relevant.			
a) Attention should focus on all important effects of treatments, and not surrogate outcomes.	*What outcomes matter to you?*	A fair comparison may not include all outcomes - short and long-term - that are important to you. Patients, professionals and researchers may have different views about which outcomes are important. For example, studies often measure outcomes, such as heart rhythm irregularities, as surrogates for important outcomes, like death after heart attack. However, the effects of treatments on surrogate outcomes often do not provide a reliable indication of the effects on outcomes that are important. Similarly, short-term effects may not reflect long-term effects.	Always consider the possibility that outcomes that are important to you may not have been addressed in fair comparisons. Do not be misled by surrogate outcomes.
b) Fair comparisons of treatments in animals or highly selected groups of people may not be relevant	*Are the people (or animals) very different from you?*	Systematic reviews of studies that only include animals or a selected minority of people may not provide results that are relevant to most people.	Results of systematic reviews of studies in animals or highly selected groups of people may be misleading.
c) The treatments compared should be similar to those of interest.	*Are the treatments different from those available to you?*	A fair comparison of the effects of a surgical procedure done in a specialised hospital or delivered by an experienced practitioner may not provide a reliable estimate of its effects and safety in other settings, or in the hands of less experienced practitioners. Similarly, comparing a new drug to a drug or dose that is not commonly used (and which may be less effective or safe than those in common use) would not provide a relevant estimate of how the new drug compares to what is commonly done.	Be aware that treatments available to you may be sufficiently different from those in the research studies that the results may not apply to you.
d) There should not be important differences between the circumstances in which the treatments were compared and those of interest.	*Are the circumstances different from yours?*	Some treatment comparisons are designed to find out if a treatment can work under ideal circumstances, for example, with people who are most likely to benefit, and most likely to comply, and with highly trained practitioners who deliver the treatment exactly as intended. These comparisons, which are sometimes called 'explanatory' or 'efficacy' studies, may not reflect what happens under usual circumstances.	Be aware that the results of studies with the aim of finding out if a treatment can work may overestimate the benefits of a treatment under more usual circumstances.
3.3 Expected advantages should outweigh expected disadvantages			
a) Weigh the benefits and savings against the harms and costs of acting or not.	Do the advantages outweigh the disadvantages?	Decisions about whether to use a treatment should be informed by the balance between the potential benefits and the potential harms, costs and other advantages and disadvantages of the treatment. When a decision affects many people, it is important to consider the distribution of the advantages and disadvantages; i.e. who will benefit, who will be harmed, who will achieve savings, and who will bear the costs.	Always consider the balance between advantages and disadvantages of treatments,

Concepts	*Short titles**	Explanations	Implications
b) Consider the baseline risk or the severity of the symptoms when estimating the size of expected effects.	*What is your baseline risk?*	The balance between the benefits and harms of treatments often depends on the baseline risk (the likelihood of an individual experiencing an undesirable event), or on the severity of the symptoms. The balance between the advantages and disadvantages of a treatment is more likely to favour the use of a treatment by people with a higher baseline risk, or more severe symptoms.	Consider the baseline risk or the severity of symptoms when making a decision.
c) Consider how important each advantage and disadvantage is when weighing the pros and cons.	*How important are the advantages and disadvantages to you?*	The balance between the benefits and harms depends on how much people value (how much weight they give to) the treatment advantages and disadvantages. Different people may value outcomes differently and sometimes make different decisions because of this. In addition, people usually place more value on things that happen soon than things that happen years into the future. In other words, the further into the future something is (for example, reducing the chance of heart disease or cancer after many years) the more people tend to "discount" its value or importance. The balance between the advantages and disadvantages of treatments may also depend on how much costs and events in the future are discounted.	Consider how important each advantage and disadvantage is when making a decision.
d) Consider how certain you can be about each advantage and disadvantage.	*How sure are you?*	The certainty of the evidence (the extent to which the research provides a good indication of the likely effects of treatments) can affect the treatment decisions people make. For example, someone might decide not to use or to pay for a treatment if the certainty of the evidence is low or very low. How certain the evidence is depends on the fairness of the comparisons, the risk of being misled by the play of chance, and how directly relevant the evidence is. Systematic reviews provide the best basis for these judgements and should report an assessment of the certainty of the evidence based on these judgements. Unexplained inconsistencies in effect estimates from different studies can also affect the certainty of the evidence.	Consider the certainty of the evidence when making a decision.
e) Important uncertainties about the effects of treatments should be addressed in further fair comparisons.	*Are there important uncertainties that should be reduced by further fair comparisons?*	There is always some uncertainty about the effects of treatments. If that uncertainty affects decisions that are important to people, the uncertainty should be reduced by further fair comparisons whenever possible. Individuals should consider participating in those fair comparisons when they are uncertain about which alternative to choose because of uncertainty about the effects of the alternatives. Participating in a fair comparison is a good hedging strategy when there is important uncertainty about effects. Moreover, people in fair comparisons often fare better than people receiving the same treatments outside of fair comparisons. In addition, the results of fair comparisons can help to generate reliable information on which to base future decisions.	Consider advocating for and participating in fair comparisons of treatments when there are important uncertainties about the effects of the treatments.

Glossary

Absolute effects	Absolute effects are differences between outcomes in the groups being compared. For example, if 10% (10 per 100) experience an outcome in one of the treatment comparison groups and 5% (5 per 100) experience that outcome in the other group, the absolute effect is 10% - 5% = a 5% difference.
Allocation	Allocation is the assignment of participants in comparisons of treatments to the different treatments (groups) being compared.
Association or correlation	Association or correlation is a relationship between two attributes, such as using a treatment and experiencing an outcome.
Average difference	The average difference is used to express treatment differences for continuous outcomes, such as weight, blood pressure or pain assessed using a scale. It is the difference between the average value for an outcome measure (for example kilograms) in one group and that in a comparison group.
Baseline risk	Baseline risk is an estimate of the likelihood that an individual or group will experience a health problem before a treatment is used.
Bias	A systematic error that may affect the results of a study because of weaknesses in its design, analysis or reporting.
Certainty of the evidence	The certainty of the evidence is an assessment of how good an indication a systematic review provides of the likely effect of a treatment; i.e. the likelihood that the effect will be substantially different from what the studies found (different enough that it might affect a decision). Judgements about the certainty of the evidence are based on factors that reduce the certainty (risk of bias, inconsistency, indirectness, imprecision and publication bias) and factors that increase the certainty.
Chance	In the context of comparisons of treatments, chance is the occurrence of differences between comparison groups that are not due to treatment effects or bias. The play of chance (random error) can lead to incorrect conclusions about treatment effects if too few outcomes occur in studies.
Confidence interval	A confidence interval is a statistical measure of a range within which there is a high probability (usually 95%) that the actual value lies. Wide intervals indicate lower confidence; narrow intervals greater confidence.
Confounders	In treatment comparisons, confounders are any factors other than the treatments being compared which may affect the health outcomes being measured.
Contamination	Contamination is the inadvertent application of a treatment allocated to one comparison group to people in another comparison group in treatment comparisons.
Data	Information gathered in studies to help address research questions, such as assessing treatment effects
Eligibility criteria	Characteristics used to decide whether people are eligible to participate in a study and should be invited to participate.
Evidence	Facts (actual or asserted) intended for use in support of a conclusion.
Explanatory study	An explanatory study (sometimes called an 'efficacy' study) is designed to assess the effects of a treatment given in ideal circumstances, in contrast to a 'pragmatic' study.
Fair comparison	Fair comparisons of treatments are comparisons designed to minimize the risk of systematic errors (biases) and random errors (resulting from the play of chance).
Indirect comparison	A direct comparison is a head-to-head comparison of treatments within a study. If there are no direct comparisons of the treatments of interest, indirect comparisons - comparisons across studies.
Model	A representation of the relationship between components of a system. Causal models represent causal relationships in a system or population.
Nocebo effect	An undesirable effect that is or could be caused by an inactive treatment, presumed to act psychologically through suggestion.
Outcome	An outcome is a potential benefit or harm of a treatment assessed in a treatment comparison. An outcome measure is how the outcome is assessed in a study.
P-value	A p-value is the probability of observing a result, as extreme or more extreme than the actual result, simply by chance, if in reality there is no treatment difference.
Placebo	A placebo is a treatment that does not contain active ingredients, which has been designed to be indistinguishable from the active treatment being compared with it.
Placebo effect	A measurable, observable, or felt improvement in health or behaviour not attributable to the treatment administered.
Pragmatic study	A pragmatic study (sometimes called an 'effectiveness' study) is designed to assess the effects of a treatment given in the circumstances of everyday practice.
Precision	The extent to which errors resulting from the play of chance affect the results of a study are likely to have occurred.
Probability	Probability is the chance or risk of something, such as an outcome, occurring. See Risk.
Protocol	A document providing detailed plans for a study
Regression to the mean	The tendency of unusually large or small measurements of something that fluctuates, such as pain, to return to a more usual or average level.

Relative effects	Relative effects are ratios. For example, if the probability of an outcome in the treatment group is 10% (10 per 100) and the probability of that outcome in a comparison group is 5% (5 per 100), the relative effect is 5/10 = 0.50.
Reliable	The reliability of a claim or evidence about a treatment effect is the extent to which it is dependable or can be trusted. It should be noted that reliability often has a different meaning in the context of research, which is the degree to which results obtained by a measurement procedure can be replicated.
Risk	Risk is the probability of an outcome occurring. See Probability.
Scale	A scale is a means for measuring or rating an outcome with a potentially infinite number of possible values within a given range, such as weight, blood pressure, pain or depression.
Statistical significance	Statistical significance is a difference that is unlikely (below a specified level of confidence – typically 5%) to be explained by the play of chance.
Study	A study is an investigation that uses specified methods to evaluate something. Different types of studies can be used to evaluate the effects of treatments. Some are more reliable than others.
Subgroup	A subgroup is a subdivision of a group of people; a distinct group within a group. For example, in studies or systematic reviews of treatment effects, questions are often asked about whether there are different effects for different subgroups of people in the studies, such as women and men, or people of different ages.
Surrogate outcomes	Surrogate outcomes are outcome measures that are not of direct practical importance but are believed to reflect outcomes that are important. For example, blood pressure is not directly important to patients but it is often used as an outcome in studies because it is a risk factor for stroke and heart attacks.
Systematic review	A systematic review is a summary of research evidence (studies) that uses systematic and explicit methods to summarise the research. It addresses a clearly formulated question using a structured approach to identify, select, and critically appraise relevant studies, and to collect and analyse data from the studies that are included in the review.
Theory	A theory is a supposition or a system of ideas intended to explain something.
Treatment	A treatment is any intervention (action) intended to improve health, including preventive, therapeutic and rehabilitative interventions, and public health or health system interventions.
Treatment comparison	Treatment comparisons are studies comparing the effects of treatments.