Assessing claims about treatment effects: Key concepts that people need to understand

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 fields have also found them relevant; for example, for assessing claims about the effects of educational interventions or environmental measures. Work to adapt these concepts to other fields is ongoing.

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.[1] They are also the basis for a database of multiple-choice questions that can be used for assessing people's ability to apply the IHC Key Concepts.[2]

We started to develop this list of concepts in 2013. We published the first version of the list in 2015 (original version), with 33 concepts in six groups. We published a revised list with 34 concepts in three groups in October 2016. The current list has 36 concepts in the same three groups.

The IHC Key Concepts serve as standards for judgment, or principles for evaluating the trustworthiness of treatment claims, comparisons, and choices. The list is intended to be universally relevant. The concepts can help people to:

1. Recognise **claims** about the effects of treatments which have an unreliable basis
2. Understand whether **comparisons** of treatments are fair and reliable
3. Make informed **choices** about treatments

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.[3] 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. That is something we plan to do in the future.

---

[1] 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. Evid Based Med, in press.

[2] Austvoll-Dahlgren A, Semakula D, Nsangi A, et al. The development of the "Claim Evaluation Tools": assessing critical thinking about effects. BMJ Open 2017; 7(5):e013184.

[3] Austvoll-Dahlgren A, Oxman AD, Chalmers I, et al. Key concepts that people need to understand to assess claims about treatment effects. J Evid Based Med 2015; 8:112-25.

Although we have written the concepts and explanations in plain language, some of them may be unfamiliar and difficult to understand. However, 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 to understand and apply the concepts.

When will the list of concepts next be updated?

The list is a "living" document allowing modification, additions and deletions, and it is subject to yearly review by a working group at the Centre for Informed Health Choices in Oslo. Astrid Austvoll-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 different types of expertise. The next update will be made available in October 2018. Please send any comments or suggestions that you have 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 webpages.

Suggested citation: Austvoll-Dahlgren A, Chalmers I, Oxman AD, and the Informed Health Choices Group. Assessing claims about treatment effects: Key concepts that people need to understand. Version: October 2017. IHC Working Paper; 2017. Available at: http://www.informedhealthchoices.org/wp-content/uploads/2016/08/Key-Concepts-2017-edition.pdf

A cautionary note: Treatments are not always 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. For these reasons, knowledge of the natural course of a health problem should be the starting point for making informed decisions about treatments.

Overview of the IHC Key Concepts (short titles)

Recognising an unreliable basis for treatment claims

1.1 Treatments can harm
1.2 Anecdotes are unreliable evidence
1.3 Association is not the same as causation
1.4 Common practice is not always evidence-based
1.5 Newer is not necessarily better
1.6 Expert opinion is not always right
1.7 Beware of conflicting interests
1.8 More is not necessarily better
1.9 Earlier is not necessarily better
1.10 Hope may lead to unrealistic expectations
1.11 Explanations about how treatments work can be wrong
1.12 Dramatic treatment effects are rare

Understanding whether comparisons are fair and reliable

2.1 Comparisons are needed to identify treatment effects
2.2 Comparison groups should be similar
2.3 Peoples' outcomes should be analyzed in their original groups
2.4 Comparison groups should be treated equally
2.5 People should not know which treatment they get
2.6 Peoples' outcomes should be assessed similarly
2.7 All should be followed up
2.8 Consider all the relevant fair comparisons
2.9 Reviews of fair comparisons should be systematic
2.10 Peer-review and publication does not guarantee reliable information
2.11 All fair comparisons and outcomes should be reported
2.12 Subgroup analyses may be misleading
2.13 Relative measures of effects can be misleading
2.14 Average measures of effects can be misleading
2.15 Fair comparisons with few people or outcome events can be misleading
2.16 Confidence intervals should be reported
2.17 Don't confuse "statistical significance" with "importance"
2.18 Don't confuse "no evidence of a difference" with "evidence of no difference"

Making informed choices about treatments

3.1 Do the outcomes measured matter to you?
3.2 Are you very different from the people studied?
3.3 Are the treatments practical in your setting?
3.4 Do treatment comparisons reflect your circumstances?
3.5 How certain is the evidence?
3.6 Do the advantages outweigh the disadvantages?

Recognising claims about the effects of treatments that have an unreliable basis

Not all claims about the effects of treatments are reliable. Well-informed treatment decisions require reliable information.

Concepts	Short titles	Explanations	Implications
1.1 Treatments may be harmful	Treatments can harm	People often exaggerate the benefits of treatments and ignore or downplay potential harms. However, few effective treatments are 100% safe.	Always consider the possibility that a treatment may have harmful effects.
1.2 Personal experiences or anecdotes (stories) are an unreliable basis for assessing the effects of most treatments	Anecdotes are unreliable evidence	People often believe that improvements in a health problem (e.g. recovery from a disease) was due to 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 got better after receiving a treatment does not mean that the treatment caused the improvement, or that others receiving the same treatment will also improve. The improvement (or undesirable health outcome) might have occurred even without treatment.	If an individual got better after receiving a treatment it does not necessarily mean that the treatment caused the improvement, or that others receiving the same treatment will also improve.
1.3 An 'outcome' may be associated with a treatment, but not caused by the treatment	Association is not the same as causation	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 of their being healthier and having 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.
1.4 Widely used treatments or treatments that have been used for a long time are not necessarily beneficial or safe	Common practice is not always evidence-based	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.
1.5 New, brand-named, or more expensive treatments may not be better than available alternatives	Newer is not necessarily better	New treatments are often assumed to be better simply because they are new or because they are more expensive. 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 it may not be possible to know whether they will appear without long term follow-up.	A treatment should not be assumed to be beneficial and safe simply because it is new, brand-named or expensive.
1.6 Opinions of experts or authorities do not alone provide a reliable basis for deciding on the benefits and harms of treatments	Expert opinion is not always right	Doctors, researchers, and patients – like anyone else - often disagree about the effects of treatments. This may be because their opinions are not always based on systematic reviews of fair comparisons of treatments.	Do not rely on the opinions of experts or other authorities about the effects of treatments, unless they clearly base their opinions on the findings of systematic reviews of fair comparisons of treatments.
1.7 Conflicting interests may result in misleading claims about the effects of treatments	Beware of conflicting interests	People with an interest in promoting a treatment (in addition to wanting to help people), such as making money, may promote treatments by exaggerating benefits and ignoring potential harmful effects. 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 have conflicting interests, be careful not to be misled by their claims about the effects of treatments.

Concepts	Short titles	Explanations	Implications
1.8 Increasing the amount of a treatment does not necessarily increase the benefits of a treatment and may cause harm	More is not necessarily 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.
1.9 Earlier detection of disease is not necessarily better	Earlier is not necessarily better	People often assume that early detection of disease leads to better outcomes. However, screening people to detect 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 tests can be inaccurate (e.g. misclassifying people who do not have disease as having disease). Screening 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.
1.10 Hope or fear can lead to unrealistic expectations about the effects of treatments	Hope may lead to unrealistic expectations	Hope can be a good thing, but sometimes people in need or desperation hope that treatments will work and assume they cannot do any harm. Similarly, fear can lead people to use treatments that may not work and can cause harm. As a result, they may waste time and money on treatments that have never been shown to be useful, or may actually cause harm.	Do not assume that a treatment is beneficial or safe, or that it is worth whatever it costs, simply because you hope that it might help.
1.11 Beliefs about how treatments work are not reliable predictors of the actual effects of treatments	Explanations about how treatments work can be wrong	Treatments that should work in theory often do not work in practice, or may turn out to be harmful. An explanation of how or why a treatment might work does not prove that it works or that it is safe.	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.
1.12 Large, dramatic effects of treatments are rare	Dramatic treatment effects are rare	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.	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.

Understanding whether comparisons of treatments are fair and reliable

Well-informed treatment decisions require systematic reviews of fair comparisons of treatments, i.e. comparisons designed to minimise the risk of systematic and random errors. Not all comparisons of treatments are fair comparisons and unsystematic summaries can be misleading.

Concepts	Short titles	Explanations	Implications
2.1 Identifying effects of treatments depends on making comparisons	Comparisons are needed to identify treatment effects	Unless a treatment is compared to something else, it is not possible to know what would happen without the treatment, so it is difficult to attribute outcomes to the treatment.	Always ask what the comparisons are when considering claims about the effects of treatments. Claims that are not based on comparisons are not reliable.
2.2 Apart from the treatments being compared, the comparison groups need to be similar at the beginning of a comparison (i.e. 'like needs to be compared with like')	Comparison groups should be similar	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.	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 were 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.
2.3 People's outcomes should be counted in the group to which they were allocated	Peoples' outcomes should be analysed in their original groups	Randomized allocation helps to ensure that the comparison groups have similar characteristics. However, people sometimes do not receive or take the allocated treatments. The characteristics of such people often differ from those who do take the treatments as allocated. Excluding from the analysis people who did not receive the allocated treatment may mean that like is no longer being compared with like. "Contamination" may lead to an underestimate of effect relative to what would have happened if everyone had received what was intended.	Be cautious about relying on the results of treatment comparisons if patients' outcomes are not 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.
2.4 People in the groups being compared need to be cared for similarly (apart from the treatments being compared)	Comparison groups should be treated equally	Apart from the treatments being compared, people in the treatment comparison groups should otherwise receive similar care. If, for example, people in one group receive more attention and care than people in the comparison group, differences in outcomes could be due to differences in the amount of attention each group received rather than due to the treatments that are being compared. One way of preventing this is to keep 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
2.5 If possible, people should *not* know which of the treatments being compared they are receiving	People should not know which treatment they get	People in a treatment group may experience improvements (for example, less pain) because they *believe* they are receiving a better treatment, even if the treatment is not actually better (this is called a placebo effect), or because they behave differently (due to knowing which treatment they received, compared to how they otherwise would have behaved). If individuals know that they are receiving (they are not "blinded" to) a treatment that they believe is better, some or all of the apparent effects of the treatment may be due either to a placebo effect or because the recipients behaved differently.	Be cautious about relying on the results of treatment comparisons if the participants knew which treatment they were receiving because this may have affected their expectations or behaviour. The results of such comparisons can be misleading.
2.6 Outcomes should be measured in the same way (fairly) in the treatment groups being compared	Peoples' outcomes should be assessed similarly	If an outcome is measured differently in two comparison groups, differences in that outcome may be due to *how* the outcome was measured 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 observe 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 measured in the same way in the different treatment comparison groups. The results of such comparisons can be misleading.
2.7 It is important to measure outcomes in *everyone* who was included in the treatment comparison groups	All should be followed up	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, 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 percentages of people lost to follow-up. The results of such comparisons can be misleading.
2.8 The results of single comparisons of treatments can be misleading	Consider all of the relevant fair comparisons	A single comparison of treatments rarely provides conclusive evidence and results are often available from other comparisons of the same treatments. These other comparisons may have different results or may help to provide more reliable and precise estimates of the effects of treatments.	The results of single comparisons of treatments can be misleading. Consider all of the relevant fair comparisons.
2.9 Reviews of treatment comparisons that do not use systematic methods can be misleading	Reviews of fair comparisons should be systematic	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. Even reviews that purport to be systematic may not be.	Whenever possible, use systematic reviews of fair comparisons rather than non-systematic reviews of fair comparisons of treatments to inform decisions.
2.10 Peer-reviewed and published treatment comparisons may not be fair comparisons	Peer-review and publication does not guarantee reliable information	Even though a comparison of treatments 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.	Always consider whether a comparison of the effects of treatments is fair and whether the results are reliable. Peer-review is a poor indicator of reliability.

Concepts	Short titles	Explanations	Implications
2.11 Unpublished results of fair comparisons may result in biased estimates of treatment effects	All fair comparisons and outcomes should be reported	Many fair comparisons never get published, and outcomes are sometimes left out. Those that do get published are more likely to report favourable results. As a consequence, reliance on published reports 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.	Be aware of the possibility of biased underreporting of fair comparisons, and whether or not the authors of systematic reviews have addressed this risk
2.12 Results for a selected group of people within a systematic review of fair comparisons of treatments can be misleading	Subgroup analyses may be misleading	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 differences.	Findings based on results for subgroups of people within a treatment comparison may be misleading.
2.13 Relative effects of treatments alone can be misleading	Relative measures of effects can be misleading	Relative effects (e.g. 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 larger 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 is likely to be worthwhile. If, however, the risk of getting the illness is 2 in 10,000, then receiving the treatment is unlikely to be worthwhile even though the *relative* effect is the same. The absolute effect of a treatment is likely to vary for people with different baseline risks.	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.
2.14 Average differences between treatments can be misleading	Average measures of effects can be misleading	For outcomes that are measured on a scale (e.g. weight, or pain) the difference between the average in one treatment group and the average in a comparison group may not make it clear how many people experienced a big enough change (e.g. in weight or pain) for them to notice it, or that they would regard as important.	When outcomes are measured on a scale, it cannot be assumed that everyone has experienced the average effect of a treatment.
2.15 Small studies in which few outcome events occur are usually not informative and the results may be misleading	Fair comparisons with few people or outcome events can be misleading	When there are only 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 between the treatments.	Be cautious about relying on the results of treatment comparisons with few outcome events. The results of such comparisons can be misleading.
2.16 The use of p-values to indicate the probability of something having occurred by chance may be misleading; confidence intervals are more informative	Confidence intervals should be reported	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. 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 an estimated treatment effect. Whenever possible, consider confidence intervals when assessing estimates of treatment effects. Do not be misled by p-values.

Concepts	Short titles	Explanations	Implications
2.17 Saying that a difference is statistically significant or that it is not statistically significant can be misleading	Don't confuse "statistical significance" with "importance"	"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.
2.18 Lack of evidence of a difference is not the same as evidence of "no difference"	Don't confuse "no evidence of a difference" with "evidence of no difference"	Systematic reviews sometimes conclude that there is "no evidence of a difference" when there is uncertainty about the difference between two 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.

Making informed choices about treatments

Well-informed treatment decisions require judgements about relevance, importance and the certainty of relevant evidence. The results of fair comparisons may not be relevant.

Concepts	Short titles	Explanations	Implications
3.1 A systematic review of fair comparisons of treatments should measure outcomes that are important	Do the outcomes measured matter to you?	A fair comparison may not include all outcomes that are relevant to treatments. 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.	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.
3.2 A systematic review of fair comparisons of treatments in animals or highly selected groups of people may not be relevant	Are you very different from the people studied?	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.
3.3 The treatments evaluated in fair comparisons may not be relevant or applicable	Are the treatments practical in your setting?	A fair comparison of the effects of a surgical procedure done in a specialised hospital may not provide a reliable estimate of its effects and safety in other settings. 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 good 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.
3.4 Comparisons designed to evaluate whether a treatment can work under ideal circumstances may not reflect what you can expect under usual circumstances.	Do treatment comparisons reflect your circumstances?	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.
3.5 Well done systematic reviews often reveal a lack of relevant evidence, but they provide the best basis for making judgements about the certainty of the evidence	How certain is the evidence?	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.	When using the findings of systematic reviews to inform your decisions, always consider the certainty of the evidence.

3.6 Decisions about treatments should not be based on considering only their benefits	Do the advantages outweigh the disadvantages?	Decisions about whether or not 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. This balance often depends on the baseline risk (i.e. 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 taking a treatment for people with a higher baseline risk or more severe symptoms. The balance also depends on how much people value (how much weight they give to) the advantages and disadvantages. Different people may value outcomes differently and sometimes make different decisions because of this.	Always consider the balance between advantages and disadvantages of treatments, taking into consideration the baseline risk or the severity of the symptoms.

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 measured on 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.
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.
Contamination	Contamination is the inadvertent application of a treatment allocated to one comparison group to people in another comparison group in treatment comparisons.
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).
Outcome	An outcome is a potential benefit or harm of a treatment measured in a treatment comparison. An outcome measure is how the outcome is measured in a study.
P-value	A p-value is the probability (ranging from zero to one) that the results observed in a study (or results more extreme) could have occurred by chance if in reality there were no treatment differences.
Placebo	A placebo is a treatment that does not contain active ingredients, which has been designed to be indistinguishable from the active treatment being assessed.
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.
Probability	Probability is the chance or risk of something, such as an outcome, occurring. See Risk
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 an instrument 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 of the effects of treatments.