# Questions to consider when using the COVID-END inventory

David Tovey, Senior advisor, COVID-END

The following is a list of questions to consider when you are using the COVID-END inventory.

## 1. Is the review addressing the question you are interested in?

This may be less straightforward to judge than one would imagine. For assessing the effects of interventions, epidemiologists often refer to the PICO framework, where PICO stands for population, intervention, comparison and outcomes. Variations of this framework exist that are applicable to different forms of evidence. Each of the four elements of PICO may represent areas that differ from the context and question that are important to you.

Perhaps the population studied is subtly different? This can refer to geography or the availability of resources, or to the setting, stage, duration or severity of illness. You may be interested in specific groups, such as the elderly, women, or vulnerable communities. Some reviews that seek evidence on COVID-19 will also capture reviews that address other situations that are judged to be similar (e.g., different viruses, pneumonia from a different cause, problems related to displaced communities). Such a broader focus can provide additional data or information that may be useful, particularly when data from COVID-19 are absent or sparse. You may need to make a judgement on the extent to which this additional data is applicable to your own context.

The intervention and comparison may also be different in important ways from your own context. For example, it is clear that the standard of care has changed in several aspects since the pandemic was first recognized, and thus the issue of co-interventions, for example, might influence how useful a piece of research published early on in the pandemic is to current person, practice or policy decisions.

Finally, it is important to check that the outcomes sought or reported by the review authors are a match with those you consider important for the decisions you are seeking to make. One consequence of the poor coordination of the global research response to COVID-19 has been the significant variation in outcomes studied. This has been a major challenge for groups undertaking evidence synthesis and for decision-makers. The [World Health Organization](#) and [COMET](#) have initiated programmes of work to identify key outcomes and outcome measures for the prevention and treatment of COVID-19, and these may be a useful reference.

## 2. Is the review up to date?

This is a critical issue for most clinical questions, but may be equally important for those in other domains. There is a vast amount of clinical and non-clinical research being undertaken as researchers have pivoted from their usual work to address questions raised by the pandemic. All reviews should report the dates on which they conducted their last search, and COVID-END reports these dates in the inventory. If you are aware of major studies that have been reported since that date, they are not likely to have been included in the review. This may substantially affect the usefulness of results and the validity of conclusions presented.

Some reviews are described as 'living'. In theory this should mean that they are updated promptly to incorporate emerging relevant research and data. However, there is no universally agreed definition of 'living', and in any case there may be delays in updating the review in the light of new evidence – hence caution and careful scrutiny are still advisable.

## 3. Is the review of high quality?

All of the reviews included in the COVID-END inventory have been judged to be above a minimum threshold for quality using the AMSTAR 1 tool, which can be used to appraise all review types. However, there will inevitably be important differences in the quality of the reviews presented and some scepticism remains important. Key markers for high quality reviews are:

- a publicly accessible protocol that can be used to determine whether the review authors conducted and reported the review as they originally intended, or if they did not, why they changed the plan
- conflicts of interests transparently declared and any support for the review clearly described
- transparency of the methods for conducting key elements of the review
- assessing and reporting on the risk of bias for the studies included
- assessing and reporting on the quality or certainty of the body of evidence – the likelihood that an effect estimate is close to the 'real' effect (see below)
- absence of spin: conclusions should match the results and data presented
- sound reporting: for reviews of effects of interventions it is always important to assess and report harmful effects, and to present results in relative and absolute terms.

## 4. Assessing certainty and the use of GRADE

Evidence is always subject to a degree of uncertainty. The use of the GRADE methodology to judge the degree of certainty that an estimation of the effects of any intervention is closely aligned to the 'real world' effect has become the favoured approach internationally for a number of reasons. One of the main benefits of GRADE compared with alternative measures is that it is outcome, not study based. In addition, the terminology used to describe degrees of certainty (high, moderate, low, very low) are easy to understand and can readily be translated into recommendations or guidelines. For any body of evidence, the GRADE method uses consistent and transparent criteria across five domains to assess the degree of certainty. These domains include the internal validity (or risk of bias), directness (or applicability), consistency (or variability), precision of the effect estimate, and the likelihood of missing data or reporting bias. GRADE also encourages the principles of sound reporting as described above, and discourages reliance on the presence or absence of 'statistical significance' to denote an effect. The emphasis is on determining the likelihood that a specified intervention is associated with an important difference compared with an alternative intervention.

Reviews that do not systematically assess the degree of certainty are prone to making over-confident assertions about the nature and magnitude of reported effects – usually amounting to an overly optimistic assessment of the studied intervention.

## 5. Absence of evidence is NOT evidence of absence

This is a basic principle of evidence-informed healthcare but one that researchers and decision-makers all too frequently ignore. In the context of a novel condition, it is inevitable that for many questions there will be a high degree of uncertainty. We know that a significant number of interventions and possible associations have been studied since the COVID-19 pandemic started. Some have now been largely abandoned because there has been clear evidence of no effect, but for the large majority the evidence is currently insufficient to make a

judgement on effects – whether of benefit of harm. This does not mean either that an intervention is safe, or that it is ineffective. We simply don't know.

## 6. Evidence isn't the only determinant for a personal, clinical or policy decision

Since the origin of evidence-informed health care, researchers have stressed that evidence itself is insufficient to inform a decision. Other factors play important roles: at the individual level this may be the patient's values and priorities or the healthcare professional's expertise and experience, at the community level it may focus around cost, availability of services or resources, or known community beliefs and preferences, amongst many other factors. All of this is particularly important when the evidence is uncertain or contested. During the COVID-19 pandemic an example of this has been the use of face masks to prevent spread, on which official advice has turned on its head as the pandemic has developed. In this case, for many people, the 'precautionary principle' described by Greenhalgh and colleagues (it probably doesn't cause harm and might be effective) has become more influential than the absence of certainty of the underlying evidence.

## 7. Be prepared to look beyond the headline

Over the past 2-3 decades, systematic reviews have become much more complex, largely to increase the utility of the evidence to decision-makers. Complexity comes in many forms, and at many levels. For instance, decision makers may wish to safeguard the health of a particularly vulnerable population, or to understand the differential effects of different modes of delivery of a multi-faceted intervention, or may wish to understand the accuracy of diagnostic tests, the effects of risk factors or the likely preferences and priorities of a community. All of these are likely to be impossible to determine from a single sentence in a declarative title. COVID-END is working to identify ways to signpost some of these factors, so that policy makers might identify the relevant content within the text of a review more easily.

## 8. It is inevitable that reviews will differ in their conclusions: be prepared to explore why

High quality reviews addressing a similar subject may differ in their conclusions for a number of valid reasons. The factors described may provide important clues (e.g., PICO question, use of GRADE, date of search, methodological approach), there may still be differences. For example, two high-quality reviews of the use of Remdesivir differed in their conclusions at least in part because their selection of outcomes did not match. Even when these factors are similar, there may be differences due to interpretation of the evidence. For example, someone who takes an individualistic approach to a decision may judge a small effect as unimportant, whereas someone who is more public health focussed may be influenced by the number of people worldwide who might be harmed or benefit from a particular decision.

## 9. Be skeptical and be prepared to adapt to changing evidence

If there is one message to be taken from this list of questions it is to encourage a degree of curious skepticism on the part of the reader. Systematic reviews are usually the best vehicle we have to judge the effects of interventions but they remain fallible for all the reasons described above. Even a review that meets all the criteria for high quality, and includes a GRADE assessment, can be misleading through error or unconscious bias on the part of the authors or because the review is not sufficiently applicable to the circumstances of the reader. In real life there are few 'wonder treatments', and surprisingly few certainties. In many cases a delicate balance exists between potential benefits and harms and reasonable people could make different choices.

Throughout the COVID-19 pandemic, policymakers have been required to make decisions based on their understanding of the best current evidence, even when this is sparse or flawed. The COVID-END inventory seeks specifically to provide this evidence in four key areas of concern: public-health measures, clinical management, health-system arrangements, and economic and social responses to the pandemic. It is inevitable that the evidence in all of these areas will change and strengthen over time, and the inventory will keep pace with these changes.