Name: David Tovey
Job title: Editor in Chief
Organisation / Institution: Cochrane
Is this input submitted as an organisational or individual response? Individual
Are you happy for your response to be published by the Academy? Yes

---

1. The overarching aim of the workstream is to better understand how society uses evidence to judge the risks and benefits of medicinal products. In your view, what are the key factors underpinning this process that the Academy should consider?

In the past 30 years the way healthcare decisions are made has changed dramatically and for the better. Patients and carers rightly expect to have a far greater say in the decisions that affect them, but in many cases both they and indeed the professionals that are guiding them do not have access to the clear and unbiased evidence they need, or that evidence is not sufficient to support a decision. The classic evidence-based triad - best current evidence, the professionalism and expertise of the clinician, and patients' values and preferences - is a necessary element of shared decision making. However, it does not seem to be consistently applied or even widely understood.

Cochrane has been at the centre of the evidence-based medicine community for 22 years. Our mission is to inform decision making by the production and maintenance of high quality systematic reviews (SRs). Reflecting this, in this submission I will focus on the successes and challenges of Cochrane (and the evidence community in general) in providing the best current evidence to guide decision making. However, the other elements of the triad are also important and are arguably less well understood. In particular, in the context of this project, the issue of patients' values and preferences is pivotal. Thirty years ago, the norm would be for the patient to be told by his or her clinician what to do and to be expected to comply. In the context of both statins and Tamiflu, patients' values and preferences are an essential part of the decision: put simply, reasonable people could consider the evidence and come to different conclusions on action. This is progress, and a natural consequence of shared decision making. There are many challenges relating to identifying and reporting the evidence though, and hence this project is a good opportunity to reflect on progress and consider how organizations like Cochrane can do better in future.

Trust is an important feature in determining the well being of health systems. Professor Per Fugelli identified a number of factors that increase the trust patients have in their doctor.[1] These include "sharing of power" and realism. He found that where trust was limited patients were more demanding, and when trust was strong they were less so. In the UK, as health care becomes increasingly "consumerized", and with controversies around NICE and perceived "rationing", patients and their carers may have diminished trust, leading them to feel that they have to fight harder to ensure that they are not denied treatment on cost grounds only. Patients may also have become more sceptical and questioning about the exaggerated claims made for drugs over several decades, and if so, that also seems mainly a good thing, although the downside is that some patients will decline to take treatments that might be effective.

---

[1] Fugelli, P. (2001). James Mackenzie Lecture. Trust--in general practice. *The British Journal of General Practice*, *51*(468), 575–579.

There is some evidence that both health professionals and patients/carers frequently have an insufficient understanding of risk and how to interpret evidence. The evidence community needs to address this by ensuring that it incorporates those measures that are evidence based to improve understanding. Accessibility is an issue and encompasses both availability and comprehension.

The failure to communicate accurately the results of trials and SRs has led to the situation where health professionals and patients/carers frequently hold overly optimistic and simplistic assumptions about the benefits and harms of treatments, as seen, for example in the enthusiastic adoption in some areas of homeopathy, which evidence has shown consistently to be ineffective. Even for traditional treatments, there is frequently uncertainty, modest effects and a narrow balance of benefit and harm. This is far away from the media assumption that drugs either 'work' in everyone ("wonder drugs") or no one. Such media accounts are frequently accompanied by moving personal stories that may be more impactful than evidence based on hundreds of thousands of patients.

It is also important that the methodological research conducted over the past three decades has identified weaknesses in the existing evidence base. This has delivered better understanding and stimulated better research, but in the short term it may have led to increased scepticism amongst the population. Citizens are now more likely to understand the limitations of clinical research, both in conduct and reporting through the work of organizations such as AllTrials[2] and CONSORT[3], and individuals such as Ben Goldacre. And the challenges are important. They include the following:

- inappropriate dependency on drug manufacturers to sponsor trials
- reporting of selected proxy and surrogate outcomes in preference to ones that are most important to decision makers
- publication and selective outcomes reporting bias
- lack of transparency or incomplete reporting of randomized trials
- unacknowledged conflict of interest

In addition, over the past few decades we have seen in many conditions a reduced threshold for diagnosing disease or identification of 'pre-disease' across a range of conditions including hypercholesterolaemia, cancer of the prostate, hypertension, osteoporosis, as well as the more recent phenomenon of companies essentially creating conditions for their drugs rather than the other way around.[4] This inevitably leads to people at lower absolute risk of poor outcomes being treated, with consequent effect on the benefits / harms calculation in these patients.

Whilst guidelines bodies such as NICE have made increasing use of SRs, it is not clear that this has spread to health professionals or patients and carers. This lack of awareness of the role of high quality evidence synthesis in general and SRs in particular, and the additional value they provide compared to individual studies, has been an obstacle to progress.

Another challenge is caused by scientific disagreement, as when, for example, researchers view the same data and draw markedly different conclusions. In one notable instance, salt reduction for preventing stroke, Cochrane reviews of low salt diets agree that salt reduction leads to a modest

[2] http://www.alltrials.net/

[3] http://www.consort-statement.org/

[4] Over-diagnosed: Making people sick in the pursuit of health. Welch G, Schwartz L, Woloshin S. Beacon Press EPUB ISBN 978-0-8070-2201-6 2011

reduction in blood pressure. [5] [6] However, researchers who have an individualistic perspective remark that the diet is a serious imposition for a relatively minor benefit; at the same time, more public health orientated researchers note the global number of deaths from stroke that could ensure and therefore come to a very different conclusion from the same data.

A final important factor is the increasing complexity of evidence. This takes many forms, with different methods addressing different types of question, for example risk modelling, or use of qualitative data within mixed methods research. However, even within traditional forms of research, the sources of evidence used will influence the complexity of the results. As research increasingly demonstrates the limitations of published reports in journals, future evidence will depend increasingly on data from Clinical Study Reports and the submissions to regulatory bodies. These influences will present even greater challenges to researchers to provide meaningful and comprehensible conclusions.

There is considerable work going on relating to patient understanding of risk e.g. “Reckoning with risk” by Gerd Gigerenzer, multiple publications by Lisa Schwartz and Steven Woloshin, and over diagnosis and over-treatment, including the *BMJ* series, and “Overdiagnosed: making people sick in the pursuit of health” by Gilbert Welch

2. When evaluating the risks and benefits of medicinal products, what are the strengths of evidence that originates from different sources?

At Cochrane, we believe that SRs can provide the most reliable source of evidence for evaluation because of their:

- rigorous methods, including a comprehensive search of all relevant studies
- ability to pool the results of clinical trials so it is more likely to detect small but clinically relevant effects that may not be seen from one trial
- ability to assess bias and take it into account in interpreting the results
- ability to replicate and review findings because of explicitly stated methods
- ability to show gaps in existing research
- potential to resolve controversy between conflicting randomized controlled trials (RCTs) on the same health intervention

We recognize that different questions in health care require evidence from different sources. Since Cochrane has historically been dominated by questions of effectiveness, for the reasons given below, we have become associated with SRs of RCTs. However, this is only part of the picture.

When evaluating evidence on healthcare interventions, SRs of RCTs represent the highest quality of evidence for benefits. In general treatment benefits are modest compared with previous best

[5] He FJ, Li J, MacGregor GA. Effect of longer-term modest salt reduction on blood pressure. Cochrane Database of Systematic Reviews 2013, Issue 4. Art. No.: CD004937. DOI: 10.1002/14651858.CD004937.pub2.

[6] Graudal NA, Hubeck-Graudal T, Jurgens G. Effects of low sodium diet versus high sodium diet on blood pressure, renin, aldosterone, catecholamines, cholesterol, and triglyceride. Cochrane Database of Systematic Reviews 2011, Issue 11. Art. No.: CD004022. DOI: 10.1002/14651858.CD004022.pub3.

treatment and even in many cases placebo.[7] Randomization remains the best way to achieve study populations that are balanced and limit the risk of bias in such circumstances. There are multiple examples (HRT, vitamins, etc.) of interventions appearing to be of benefit in non-randomized studies, only to be shown to be ineffective or even harmful in subsequent RCTs.

In addition, evidence is changing: questions and interventions are becoming more complex. These changes mean that in future SRs will need to incorporate different methods and sources of evidence in order to provide the sophistication and complexity demanded by decision makers.

3. When evaluating the risks and benefits of medicinal products, what are the limitations of evidence that originates from different sources?

The last 10 years have provided compelling evidence that in addition to publication bias (non-publication of reports from a study), published reports of studies are also at risk of selective outcome reporting. This has led to the use of the more comprehensive data submitted to licensing regulators and clinical study reports, as was the case for the Cochrane Tamiflu review.

In addition, in evaluating many public health or health systems questions, RCTs are not always appropriate or feasible. As an example, a Cochrane review evaluated the effects on public health of slum clearance. It would seem inappropriate to randomize populations to slum clearance or non-clearance, so different forms of evidence were considered. The Cochrane Effective Practice and Organisation of Care Group has, since its inception in the mid-1990s, also evaluated quasi RCTs, controlled before and after studies, and interrupted time series in its reviews.

For common harms that occur within the time span of an RCT (often short), SRs of RCTs can provide high quality evidence, but for rare or delayed harms they are frequently inadequate. In such cases the best current evidence can come from well conducted observational studies. Even case series can be compelling if the outcome in question is not seen otherwise in the population, e.g. mandibular necrosis in people taking bisphosphonates. Similarly, scientific misconduct in the RCT reports of suicidality while using SSRIs have been recognized at least in part after concerned citizens and grieving relatives reached out to one another.

Limitations of SRs:

- SRs rely on high quality RCTs being conducted and accurately reported. The best estimate is that about 50% of all clinical trials have never been published so SRs are only able to provide a partial answer. As founding members of the AllTrials campaign, we are working to ensure that all clinical trials are registered and their results reported. In future, clinical trials will be reported, but we still need the vast numbers of clinical trials already conducted to be reported.
- There may be no clinical trials about the specific question to conduct the SR or trials reports may be of such poor quality that the review can't be conducted.

[7] Djulbegovic B, Kumar A, Glasziou PP, Perera R, Reljic T, Dent L, Raftery J, Johansen M, Di Tanna GL, Miladinovic B, Soares HP, Vist GE, Chalmers I. New treatments compared to established treatments in randomized trials. Cochrane Database of Systematic Reviews 2012, Issue 10. Art. No.: MR000024. DOI: 10.1002/14651858.MR000024.pub3.

- The study population may be non-representative of the general population, and results may not be applicable.
- Results may prove inconclusive.

4. Please provide details of any further examples or case studies that it would be useful for the project to consider.

One of the classic examples of the value of a SR is the forest plot that forms the centre of the Cochrane logo. Each horizontal line represents the results of one study, while the diamond represents the combined result, our best estimate of whether the treatment is effective or harmful. The diamond sits clearly to the left of the vertical line representing "no difference"; therefore the evidence indicates that the treatment is beneficial. We call this representation a "forest plot". This forest plot within our logo illustrates an example of the potential for SRs to improve health care. It shows that corticosteroids given to women who are about to give birth prematurely can save the life of the newborn child.
Despite several trials showing the benefit of corticosteroids, adoption of the treatment among obstetricians was slow. The Cochrane Review (originally published by Crowley et al. and subsequently updated) was influential in increasing use of this treatment. This simple intervention has probably saved thousands of premature babies.

There are many example scenarios where evidence from one source was misleading and additional sources provided intelligence of importance to decision makers. For example:

- Hormone Therapy (HT) for women for menopausal symptoms: prescribed widely on the basis of promising evidence from observational evidence on both benefit and harm evidence which was found to be overly optimistic when tested in a large, independent RCT.
- Cisparide for neonatal gastro-oesophageal reflux: found to be beneficial in RCTs, but a possible association with arrthymias identified in observational studies led to the drug being withdrawn in most countries for this indication.

In the future, we will also have "big data" from a variety of sources that will need to be considered and synthesized. It is very likely to become widely used (and valuable) as a source of evidence but the causes, extent, and direction of bias are not yet fully understood.
Another case example that might be considered relates to Glucosamine for osteoarthritis. In this example, small, early studies appeared to demonstrate a benefit, but later a Cochrane Review found that when only studies at low risk of bias were considered the magnitude of benefit was marginal or non-existent.[8]

[8] Towheed T, Maxwell L, Anastassiades TP, Shea B, Houpt J, Welch V, Hochberg MC, Wells GA. Glucosamine therapy for treating osteoarthritis. Cochrane Database of Systematic Reviews 2005, Issue 2. Art. No.: CD002946. DOI: http://onlinelibrary.wiley.com/doi/10.1002/14651858.CD002946.pub2/abstract

5. Please highlight any broadly applicable principles that should govern the presentation, interpretation and weighting of evidence about medicinal products.

A "good data for decision making" Charter might insist on:

1. Benefits AND risks
   All clinical trials and SRs should be expected to actively seek and report evidence on benefits and harms, to present any limitations of their studies (e.g. duration of studies) and also to highlight other sources of evidence that might address any uncertainties that remain.

2. Absolute values
   All studies and SRs should present both relative and absolute effect data. Whilst relative effect sizes may be more stable across different at-risk groups, they are prone to lead to a perception of an exaggerated effect (for example a 50% reduction in 2-year recurrence of breast cancer in women taking Trastuzumab equates to about eight women per 100 benefiting from treatment). In presenting absolute effects, an evidence-based approach should be taken as to the way the calculation is presented (e.g. natural frequencies versus percentages etc).[9]

3. Some measure of the quality of the evidence and the certainty of the results
   It is still common to see reports from clinical studies or SRs that do not evaluate or present any information on the risk of bias, or the quality of the evidence (often expressed as the degree of certainty of the result). It is crucial for decision makers to know what level of confidence they should place in the reported effect estimates.

4. Outcomes and comparisons that matter to decision makers
   Studies that evaluate novel treatments should compare the new treatment against current best practice (as opposed to placebo or outdated therapies) so that the reader can understand the comparative effectiveness. Similarly, studies and SRs should look for and report outcomes that are most important to decision makers, rather than focussing on short term or proxy indicators that do not relate directly to clinical outcomes.

5. Avoiding over-reliance on arbitrary definitions of statistical significance and incorporating pre-determined estimates of minimal important difference
   Whilst scientists understand that no evidence of effect is not equivalent to evidence of no effect, it is still common to see this distinction blurred in reports from trials and in SRs, due to over-reliance on 'statistical significance'. Non-statistically significant results often denote uncertainty in the context of a possible effect, but are reported as demonstrating equivalence. At the other end of the scale, statistically significant results can be so small in magnitude that they are not important to decision makers. Using a pre-determined

[9] Akl EA, Oxman AD, Herrin J, Vist GE, Terrenato I, Sperati F, Costiniuk C, Blank D, Schünemann H. Using alternative statistical formats for presenting risks and risk reductions. Cochrane Database of Systematic Reviews 2011, Issue 3. Art. No.: CD006776. DOI: http://onlinelibrary.wiley.com/doi/10.1002/14651858.CD006776.pub2/abstract

assessment of the magnitude of a difference is important in guiding a decision is a much more intuitive and valid way to describe results.[10]

The GRADE system is now the world leader for judging our certainty in the evidence. Used initially to communicate the results of SRs to guidelines panels, the judgement incorporates a number of characteristics that influence our confidence that the estimated effect estimate is close to the actual effect. These include type of study, risk of bias of the studies, "directness" or applicability of the studies, variation in the effects found in the studies, the likelihood of reporting bias, and the precision of the effect estimate. GRADE has the additional advantage that it rates outcomes across studies, and can also be used to rate the quality of the evidence from observational studies.

6. Concerns have been raised about how industry funding impacts on the validity, or the perception of validity, of evidence. For example, the ability of academic researchers funded by industry to remain impartial when evaluating evidence has come into question. How should conflicts of interest be addressed? How important is industry funding in generating and analysing evidence? Other than industry sponsorship, what are other potential sources of conflicts of interest?

Clinical research affecting how doctors practice medicine is increasingly sponsored by companies that make drugs and medical devices.

A Cochrane Review published in 2012 found that industry sponsored drug and device studies more often had favourable efficacy results, (risk ratio (RR): 1.32, 95% confidence interval (CI): 1.21 to 1.44), harms results (RR: 1.87, 95% CI: 1.54 to 2.27) and overall conclusions (RR: 1.31, 95% CI: 1.20 to 1.44), compared with non-industry sponsored drug and device studies. Among the standard factors that may increase risk of bias, Industry sponsored studies more frequently reported a satisfactory approach to describe blinding. Nevertheless, sponsors' products' favourable results could not be completely explained by the standard risk of bias factors assessed in Cochrane reviews.[11]

It is very important to note that Industry sponsorship is not the only source of bias: in many cases researchers have a strong financial or non-financial commitment to an intervention - e.g. researchers who provide a service such as CBT, or who may have trademarked a particular intervention, may have a strong prior belief that is effective that can influence their research. SR authors who are also trialists may equally favour the results of their study unless there are

[10]Effective Practice and Organisation of Care (EPOC). Results should not be reported as statistically significant or statistically non-significant. EPOC Resources for review authors. Oslo: Norwegian Knowledge Centre for the Health Services; 2013. Available at: http://epoc.cochrane.org/sites/epoc.cochrane.org/files/uploads/22%20Interpreting%20statistical%20significance%202013%2008%2012_2.pdf

[11] Lundh A, Sismondo S, Lexchin J, Busuioc OA, Bero L. Industry sponsorship and research outcome. Cochrane Database of Systematic Reviews 2012, Issue 12. Art. No.: MR000033. DOI: 10.1002/14651858.MR000033.pub2.

barriers in place to avoid this, as there are in Cochrane Reviews.

7. Please outline any past, current or planned initiatives to examine how patients, citizens and healthcare professionals (and those who seek to inform them) evaluate scientific evidence about medicinal products.

The EU Decide project has been influential in testing how patients, citizens, and healthcare professionals understand scientific evidence, and has led to a number of initiatives aimed at improving the potential for shared decision making. This builds on the work of the GRADE working group, which has developed guidance for communicating the results of SRs for policy makers and guidelines bodies.

8. What are the most effective ways of communicating evidence to various stakeholders and engaging with them about such evidence?

SRs have been conducted to look at how to communicate research findings to clinicians as well as to policy makers. General principles that have some effect on knowledge transfer include:

- who is communicating what, to whom, in what medium; how the evidence was generated in the first place; things pertaining to formatting (such as layering content from simple-to-complex, etc.) .
- The scope of evidence goes beyond communication in isolation and includes findings about processes that can also affect communication (e.g. evidence created in collaboration with stakeholders) and the need to involve end users from the beginning to end of the process.

Sir Iain Chalmers and the James Lind Alliance[12] have been highly influential in developing and implementing methods for ensuring that the priorities of patients, carers, and health professionals are incorporated into research planning. This also includes a widely disseminated book and website, attempting to explain research methods and implications to the public.[13] Cochrane has also been active in developing a vibrant consumer network. Members of the network have been able to influence the nature of the SRs (by indicating the outcomes and comparisons that they consider most important) and by providing input into the plain language summaries aimed at communicating the findings of the reviews to non-technical audiences. However, it is unreasonable to expect that individual patients can represent diverse communities and populations and accurately reflect their preferences and values. Thus, it is important that approaches are based on evidence drawn from studies of the preferences of the populations of interest where possible.

[12] http://www.jla.nihr.ac.uk/

[13] http://www.testingtreatments.org/