Name: Dr Huw Llewelyn, MD FRCP
Job title: Hon Departmental Fellow
Organisation / Institution: Aberystwyth University
Is this input submitted as an organisational or individual response? Individual
Are you happy for your response to be published by the Academy? Yes

Submitted using the online form, now formatted into a PDF.

---

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?

The evidence required to judge the risks and benefits of medicinal products is currently incomplete. Randomised clinical trials (RCTS) are central to the process but they are insufficient on their own and too much reliance is placed on them at present. However well the RCT is done, an attempt should be made to replicate some important aspects of its findings in various clinical settings, without necessarily repeating the entire RCT in these different settings.

RCTs only provide evidence of efficacy in that the medicine CAN benefit some patients under special circumstances. Other equally important studies need to be done, usually on larger numbers of patients, to show the extent of probable benefit and harm in subgroups of patients, especially those within the spectrum of mild, moderate and severe disease. It is also important to assess the ability of different symptoms, signs and test results to provide accurate probabilities of such benefits and harms.

At present, information about harms is gathered in unstructured post-marketing surveys and ad-hoc reporting of suspected adverse events. There is little attempt to assess using prospectively planned structured studies, the probable benefit and harms in different sub-groups. This needs to be assessed in those with different degrees of illness as assessed using different symptoms, signs and test results. The current generation of students and young doctors are introduced to these ideas in the Oxford Handbook of Clinical Diagnosis 3e pp 615 to 642.

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

RCTs, and meta-analyses performed on them, provide the strongest assessment of a medicine's efficacy by making all causes of bias and a change in a target outcome due to some other cause unlikely. The weakness of RCTs is that they are expensive, require highly skilled operators, they result in patients being treated differently and therefore cause problems with issues of perceived fairness and consent.

The weakness of case reports, observational studies or large data bases is that they are affected easily by bias. However meticulous the reasoning process used to try to explain away bias, it will be very difficult to convince a critical listener that an outcome is solely due to the medicine being tested and nothing else. If it is not possible to perform an RCT, this reasoning will have to be done by showing that all known biases are unlikely so that by a process of elimination, it only the

medicine under study that is the probable cause of the outcome. The strength of non-RCT assessments is that they are cheaper to do, disrupt the lives of patients less than RCTs and considered to be the only option remaining if it is not possible to do an RCT.

AN EVOLVING AND NOVEL TRIAL DESIGN -THE CUT-OFF STUDY

This novel approach is to conduct a controlled study during day to day care without using randomisation. This is not designed to replace RCTs but to supplement the data from a RCT. It is done by examining the original RCT and identifying the absolute probability of benefit from the medicine at different severities of the disease being treated. This is done by subtracting the estimated frequency of a beneficial outcome on placebo from the estimated frequency of the beneficial outcome on treatment.

A cut-off point is then placed at the point where the difference between the frequency of an outcome on treatment and control is great enough to consider treating the patient (see page 633 to 635 of the Oxford Handbook of Clinical Diagnosis 3e).This cut-off point can be of any measure such as test result (e.g. an albumin excretion rate), the calculated baseline probability of some surrogate end-point (e.g. of a vascular event within one year) or the calculated baseline probability of a degree of well-being index without treatment.

Patients are then treated if they are one side of the cut-off point and not treated if they are on the other side. The frequency of each outcome measured at each degree of severity of the measure is plotted against the estimated outcome frequency based on the original RCT. If the estimated frequencies from the RCT match those of the cut-off controlled study, then this aspect of the RCT result would have been replicated.

If not, then the plot of the estimated outcome frequency from the RCT and the observed outcome frequency from the cut-off study (COS) could be used as a calibration curve.

If there was a difference between treatment and placebo at the cut-off point in the COS study, then there would be a discontinuity of the plotted curve on each side of the cut-off point. The displacement should be equal to the proportion benefiting in the RCT. If this were the case, then the proportion benefiting in the trial would have been replicated in the COS. A COS study would also allow the ability of different tests to predict who would benefit from a treatment (and who would not benefit) to be compared. In other words, a COS study could be used to assess which tests stratified patients better in order to identify those more likely to respond.

The advantage of this approach is that it would be cheaper, and easier to do than a very large additional RCT, and could be done therefore on a much larger group of patients during day to day care with minimal disruption. The larger numbers would also mean that less common adverse effects could be detected. If desirable, the patients given and not given active treatment because they were on the relevant sides of the cut-off could be blinded to this fact. However, all patients above or below a cut-oof point would be treated consistently. The disadvantage is that there would not be a placebo treated limb to compare with those treated. However, if the COS cast doubt on the original RCT it could be repeated. However the knowledge that a COS would be done on all RCTs showing a positive result would ensure that such RCTS were carried out carefully and honestly.

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

RCTs provide the best evidence that a medicine can cause an improvement in outcome compared to a control but their limitations are that they are difficult to organise, expensive and they mainly assess efficacy. A limitation is that their cost means that very large trials, perhaps in different clinical settings are impracticable.

Data mining in unstructured data sets collected for clinical purposes might help in screening for adverse effects but they would not provide information of value for assessing efficacy or clinical effectiveness.

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

The following study can be used to plan a cut-off study as described in Section 2 above:
Llewelyn D E H, Garcia-Puig, J.(2004) How different urinary albumin excretion rates can predict progression to nephropathy and the effect of treatment in hypertensive diabetics. JRAAS, 5; 141-5.
(see
http://jra.sagepub.com/content/5/3/141?ijkey=d78a5738be3c72cc97d9caffdf15915216a6b96c&keytype2=tf_ipsecsha

The way in which a cut-off study might be set up and be interpreted is explained in the following reference: Llewelyn H, Ang AH, Lewis K, Abdullah A. The Oxford Handbook of Clinical Diagnosis 3rd edition. Oxford University Press, Oxford, 2104, pp 633 to 635

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

The over-riding principle is that the results of all RCTs (which are usually conducted in specialised research situations) should be checked in day to day clinical settings (e.g. by conducting a cut-off study) to see if the result can be replicated. This would also mean that all RCTs in research situations would be conducted with utmost care so that the probability of subsequent replication in day to day practice would be high. Also post-licensing studies should not be conducted retrospectively on data (however big) that were stored for clinical or other purposes. Instead, the studies should be planned carefully and interpreted in a planned prospective way (e.g. as a cut-off study (see Section 2 above).

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?

Conflicts of interest of different degrees may always be a problem for industry funded studies (and non-industry funded studies too). One solution would be for commercial organisations to out-source their scientific studies to independent laboratories or clinical investigators who would not be biased in favour or against the medicine to be assessed (but this would not guarantee a solution in view of human nature).
Another problem is that RCTs are always conducted in specialised research situations that could be misleading. It is important that such studies are checked by replicating aspects of the RCT result as explained in Section 2 and also supplementing the RCT by continuing the cut-off study on large numbers of patients to get further information about clinical effectiveness and adverse effects.

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.

My current initiative is writing the Oxford Handbook of Clinical Diagnosis to teach students and young doctors that the effective use medicals products depends not only on their efficacy (as shown in RCTs) but also on an intelligent use of diagnostic information (symptoms. signs and test results). I am also planning to teach patients how to ask doctors the question to allow them to understand how the choice of such a treatment was made wisely.

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

I think that the most effective way of communicating evidence would be to first teach a wide range of stakeholders, including patients, the public, health care professionals (general practitioners, nurses, pharmacists, clinicians, etc),and the media about how shared degrees of certainty is driven by prospectively collected data (not retrospectively data-dredging big data for example) and that a planned RCT is one example of this.