Reflections on clinical trials – the distance from results to action.

Adrian Boswood

I have been fortunate in my career to have been involved in a number of clinical trials. The conclusion of a clinical trial and the publication of a trial's results represent a great achievement for all those who participated, but it is only part of the process involved in ensuring the adoption of a new therapy. One reason for this is summarised well by the following quote –

> "If we prove a new intervention works but no one receives it, we have achieved only a sterile intellectual milestone that has little impact on …. health"(Packer 2002)

The challenge after demonstration of a favourable outcome in a trial is to ensure the results are appropriately interpreted and disseminated so that those patients that will benefit from therapy receive the treatment, and by inference those that will not, do not.

The ease with which the results of a trial can be applied is dependent in part on the way the trial was designed and therefore the ease with which the results can be extended from the specific population of animals recruited to the study to the more general population of animals seen by veterinarians. The design of most prospective clinical trials can be reduced to four key components. These can be summarised with the acronym PICO (Haynes 2006). **P** stands for the "**P**atients" or "**P**opulation" being studied, **I** for the "**I**ntervention" that is being evaluated, **C** for the "**C**omparative intervention" to which the intervention will be compared and O for the "**O**utcome" of interest that may be influenced by the intervention. If we take as an example the recently published EPIC study (Boswood and others 2016); the patients we recruited were dogs with stage B2 mitral valve disease, the intervention was pimobendan, the comparative intervention was a placebo and the outcome of interest was the development of heart failure or the death of a patient attributable to their heart disease.

As well as having a shared basic structure, clinical trials share some implicit assumptions. Perhaps the most important assumption – and one which underlies most of evidence based (veterinary) medicine (EBM) – is that the population enrolled in a clinical trial are representative of a larger population of similar patients. This means that if the clinical trial shows a significant effect of the intervention being evaluated in the population enrolled in the trial, we can conclude not just that it ***had*** an effect in the population in the study, but that it ***will have*** a similar effect when administered to a population of patients similar to those enrolled in the study. This extrapolation of the results of the study to other patients is sometimes referred to as the "generalisability" or "external validity" of the study.

In order to be enrolled to a study, patients usually need to meet a number of specific criteria; these are often referred to as inclusion and exclusion criteria. The purpose of these criteria is to establish that the enrolled patients definitely have the disease of interest at an appropriate stage. In order to show that they meet these criteria patients need to undergo a series of diagnostic tests prior to inclusion in the trial. The population of interest is therefore defined, in part, on the basis of these test results. Unfortunately the more inclusion and exclusion criteria that need to be met, the more tests one will subsequently need to perform in order to identify patients that will benefit from the therapy. This restricts the population to which the results can be generalised.

To illustrate this point with a current example – in order to ensure that the patients recruited to the EPIC study (Boswood and others 2016) definitely had primary mitral valve disease *and* enlargement of their heart they underwent both echocardiography and radiography prior to study entry. If I am now asked by a practitioner – "should I treat this dog, which I believe has preclinical mitral valve disease, with pimobendan?" – my response will typically be "to be certain that therapy is indicated the dog will need to undergo echocardiography and ideally also radiography". Had the dogs in the study not undergone those tests our results would have been (justifiably) doubted by our peers because the severity of the recruited dogs' heart disease would not have been known, but because the dogs in the study *did* undergo these tests our ability to confidently apply the results to similar patients in practice is restricted to the subpopulation of dogs that undergoes similar diagnostic tests. Even the most cocooned cardiology specialist must realise that this will prevent some, and probably quite a lot of, dogs from receiving potentially beneficial therapy because of restricted access to, or ability to afford, the necessary diagnostic tests.

The more stringently one defines the population recruited to a study the more one can be certain of the severity and type of disease with which that population is afflicted, but paradoxically the less easily one can generalise the results to a less well defined population. This illustrates the tension between what are sometimes called studies of an intervention's "efficacy" and knowledge of its "effectiveness".

Efficacy and effectiveness have been defined and distinguished from each other as follows – **efficacy** refers to the beneficial and harmful effects of an intervention when "applied under ideal circumstances" whereas **effectiveness** refers to the beneficial and harmful effects of an intervention when "applied under the usual circumstances that apply in healthcare" (Sackett 2006). Under "ideal circumstances" (in studies of efficacy) all patients that receive a medication have carefully characterised disease, they are known to be free of certain concurrent disease processes and in order to be sure of these things they have undergone extensive diagnostic testing. Patients have their health carefully monitored throughout the duration of the trial. Those patients (or the owners of patients) participating in studies are highly motivated and likely to diligently comply with instructions regarding the administration of medication and typically that compliance is in some way quantified. The reality of the "usual circumstances that apply in primary healthcare", both in veterinary and human healthcare, is somewhat different. Patients will have comorbidities, do not undergo diagnostic testing or monitoring as comprehensive as that undergone by patients recruited to clinical trials and may less diligently comply with medication instructions. These differences in patient populations may lead to some differences in the observable beneficial and harmful effects.

So how can this gap between the patients in clinical trials and patients in general practice be bridged, or at least narrowed? There are several ways in which this can be done and these should be borne in mind by those designing and conducting clinical trials. One way is to minimise the number of restrictions to entry to a clinical trial – which is something we tried to do when designing the EPIC study (for instance more than 40% of dogs recruited to the study had a reported comorbidity)(Boswood and others 2016). Perform the bare minimum number of diagnostic tests necessary to establish the presence, and stage, of a patient's disease and try not to have inclusion criteria that rely too heavily on tests to which general practitioners have only limited access (e.g. 3D-Echocardiography, MRI and CT scans). Another way is to design clinical trials with very simple entry criteria that recruit patients in ways that are readily applicable in a general practice setting; for

instance on the basis of history or physical examination findings, blood test results or a combination of these things. What would be better still would be to conduct simple, but large, trials in a general practice setting – so called “practical” or “pragmatic” clinical trials in human patients have taken this approach (Tunis and others 2003).

An important conclusion from the foregoing discussion is that the further away any clinical research is from the “usual circumstances that apply in healthcare” the more difficult it will be for that research to have an immediate impact on the majority of patients that might benefit from its findings; because the majority of patients will always be seen in primary care practice. There is also a risk that externally generated recommendations for “optimal” treatment based on data obtained under “ideal circumstances” will be impossible to implement or unaffordable for the majority of patients. This can contribute to the perception of EB(V)M as an abstract unrealistic method, externally imposed on practitioners, that robs them of their autonomy while setting a standard they consistently cannot reach. This extreme, but sometimes espoused, view of EBM is, in my view, a result of the way some evidence is generated rather than an inevitable consequence of EBM.

Where I hope the EPIC study struck the right balance was in performing sufficient tests to convince a sceptical readership of the nature and severity of the enrolled patients’ disease, while using relatively straightforward and accessible tests to do so (2-D and M-mode echocardiography with plain radiography). This will mean that many, but admittedly not all, patients in primary care practice will now be able to access the improved standard of care that we have demonstrated. Whether there is an even simpler way of identifying the same, or a similar population of patients that will benefit from therapy is not yet known, but if we are to broaden still further the applicability of our results we should strive to see if we can find one.

If research is conducted in such a way that it can only make those who are already practising at a specialist level better it may not benefit a large number of patients. If research can be conducted in such a way as to translate readily into improvement in the standard of care offered in general practice, the immediate impact on the health of a large number of animals will be greater. Advances in the practice of veterinary medicine are likely to rely on a combination of both approaches. We need the former to show us what can be achieved by specialists under optimal conditions, but we should not lose sight of the fact that we also need the latter to help find ways of easily disseminating demonstrated improvements in outcome to the majority of animals who might benefit from them.

Boswood, A., Haggstrom, J., Gordon, S. G., Wess, G., Stepien, R. L., Oyama, M. A., Keene, B. W., Bonagura, J., MacDonald, K. A., Patteson, M., Smith, S., Fox, P. R., Sanderson, K., Woolley, R., Szatmari, V., Menaut, P., Church, W. M., O'Sullivan, M. L., Jaudon, J. P., Kresken, J. G., Rush, J., Barrett, K. A., Rosenthal, S. L., Saunders, A. B., Ljungvall, I., Deinert, M., Bomassi, E., Estrada, A. H., Fernandez Del Palacio, M. J., Moise, N. S., Abbott, J. A., Fujii, Y., Spier, A., Luethy, M. W., Santilli, R. A., Uechi, M., Tidholm, A. & Watson, P. (2016) Effect of Pimobendan in Dogs with Preclinical Myxomatous Mitral Valve Disease and Cardiomegaly: The EPIC Study-A Randomized Clinical Trial. *J Vet Intern Med* **30**, 1765-1779

Haynes, R. B. (2006) Forming research questions. In: Clinical epidemiology: how to do clinical practice research, 3rd edn. Eds R. B. Haynes, D. L. Sackett, G. H. Guyatt and P. Tugwell. Lippincott Williams and Wilkins, Philadelphia. pp 3-14

Packer, M. (2002) The impossible task of developing a new treatment for heart failure. *J Card Fail* **8**, 193-196

Sackett, D. (2006) An introduction to performing therapeutic trials. In: Clinical epidemiology: how to do clinical practice research, 3rd edn. Eds R. B. Haynes, D. L. Sackett, G. H. Guyatt and P. Tugwell. Lippincott Williams and Wilkins, Philadelphia. pp 59-65

Tunis, S. R., Stryer, D. B. & Clancy, C. M. (2003) Practical clinical trials: increasing the value of clinical research for decision making in clinical and health policy. *Jama* **290**, 1624-1632