

ANALYSIS

Evidence of comparative efficacy should have a formal role in European drug approvals

Despite methodological concerns, comparative efficacy evidence should be required at the time of drug approval, **says Corinna Sorenson and colleagues**, to allow patients, clinicians, and other healthcare decision makers to determine whether a new drug is superior, equivalent, or inferior to its existing alternatives

Corinna Sorenson *research fellow*¹, Huseyin Naci *PhD candidate*¹, Jonathan Cylus *technical officer*^{1,2}, Elias Mossialos *professor and director*^{1,2}

¹LSE Health and Social Care, London School of Economics and Political Science, London WC2A 2AE, UK ; ²European Observatory on Health Systems and Policies, London, UK

Manufacturers of new drugs need to demonstrate that their products are efficacious and safe for a defined group of patients to obtain market approval. However, demonstrating these outcomes relative to existing therapies is required by regulators only when use of placebo is deemed unethical.^{1,2} Regulators, clinicians, patients, and payers therefore often lack the necessary information to distinguish between available medicines in terms of their comparative therapeutic value and safety.

Comparative efficacy evidence at the time of drug approval is important, and there are methodological tools available to generate such information. When one or more treatment alternatives are available, demonstrating lack of inferiority through comparative assessment should be a formal requirement, and there are ways to support this objective in European drug licensing.

Need for comparative efficacy evidence

When a drug comes to market, evidence on the comparative risks and benefits is needed to help regulatory authorities to safeguard public health from inferior and unsafe treatments, to ensure that health technology assessment agencies and payers make funding decisions based on the best available evidence of different treatments, and to aid clinicians' and patients' understanding of what therapies work best and their appropriate position in the treatment pathway.³ However, comparative assessment (box 1) is often conducted or made available only once a therapy is already on the market. This is partly because pre-marketing comparative efficacy studies entail potential uncertainty and risk for manufacturers, as failure to demonstrate a therapeutic advantage over older, and less costly, alternatives may affect drug sales or result in a drug not being approved.² This lack of comparative assessment allows manufacturers to

differentiate their products from competitors on factors unrelated to demonstrated relative efficacy and safety.³

Although post-market assessment, which largely focuses on comparative effectiveness as opposed to comparative efficacy, is important, reliance on this approach alone to determine the relative therapeutic value of medicines can be problematic. For instance, such studies are infrequently fulfilled by industry when required by regulators,⁴ and reports communicating comparative evidence may be biased.⁵ Even when objective evidence is available, it may be difficult to shift prescribing behaviour once a therapy is used in practice and patterns of care are established.⁶ Furthermore, effectiveness data are not usually available at the time of initial assessment by regulators (and, often, payers), resulting in important evidence gaps when new treatments are introduced into the market. Consequently, earlier information to ascertain the relative benefit of new drugs is needed.

A lack of early comparative efficacy evidence can result in the widespread use of potentially less efficacious and unsafe drugs, as highlighted by the recent case of the diabetes drug rosiglitazone. The relative effect of rosiglitazone against pioglitazone emerged after years of widespread use,⁷ where rosiglitazone was shown to increase the risk of myocardial infarction and cardiovascular death.^{8,9} A number of widely publicised studies have also questioned the true added value offered by new (and often more expensive) drugs compared with existing treatments.^{10,11,12} The Antihypertensive and Lipid-Lowering Treatment to Prevent Heart Attack Trial (ALLHAT), for example, demonstrated that coronary heart disease risk was not reduced for any of the three newer and more costly drug classes compared with older and cheaper thiazide based diuretics.¹³

Box 1 Frequently used terms in comparative drug assessment

- **Efficacy**—the extent to which an intervention does more good than harm under ideal circumstances
- **Effectiveness**—the extent to which an intervention does more good than harm when provided under the usual circumstances of healthcare practice
- **Comparative efficacy**—the extent to which an intervention does more good than harm, under ideal circumstances, compared with one or more intervention alternatives for achieving the desired results
- **Comparative effectiveness**—the extent to which an intervention does more good than harm compared with one or more intervention alternatives for achieving the desired results when provided under the usual circumstances of healthcare practice

From: High Level Pharmaceutical Forum of the European Commission

Current requirements for and availability of comparative efficacy evidence at the time of drug approval

The European Medicines Agency has long encouraged that, when possible, pre-market studies should be undertaken to establish comparative efficacy and risk,¹ but has yet to set comparative assessments as the default evidentiary standard for market approval; rather, requirements for comparative studies are made on a case by case basis.^{1 14 15}

Existing estimates from Europe and the United States suggest that comparative efficacy data are available for 50–70% of new molecular entities at the time of market authorisation.^{16 17} While these estimates appear encouraging, the situation is more complex.

Firstly, the availability of such evidence varies across therapeutic areas. For example, comparative efficacy data are available for 89% of new molecular entities for diabetes mellitus, versus 50% and 44% for neurological and gastrointestinal tract conditions, respectively.¹⁷

Also, available comparative evidence is too infrequently accessible at the time of market authorisation: only about a quarter of studies are published, with positive findings more often published than negative findings.¹⁴ In the case of rosiglitazone, it was only following a legal settlement requiring its manufacturer to release results of all its clinical trials that researchers could carry out a systematic analysis of available trial data.¹⁸

Methodological considerations for generating comparative efficacy evidence

The need for comparative efficacy assessment brings into focus the methodological approaches available to generate such evidence. A number of clinical study designs are available to assess comparative efficacy, each with advantages and disadvantages (box 2).

Direct (head to head) comparisons of one or more drug therapies in an RCT have long been considered the gold standard for assessing comparative efficacy. For new drugs, this is best achieved using an active comparator trial (RaCT), where patients are randomised to receive either the experimental drug or an established treatment (reference or standard). Well designed RaCTs can demonstrate comparative superiority (improved efficacy of new medicine against existing therapies), equivalence (absence of a difference), or non-inferiority (that a new medicine is no worse than alternatives).²

However, active trials are associated with a number of methodological issues, especially with regards to demonstrating superiority. RaCTs must be large enough and lengthy enough to assess clinical versus surrogate outcomes, and to detect

meaningful differences in outcomes to ascertain superiority versus equivalence. This makes them more expensive and longer to conduct than placebo trials.¹⁶ In addition, selecting an appropriate comparator is not always straightforward or feasible, especially when a large number of similar comparators exist or when trials are lengthy (for example, a new standard of treatment may be adopted during the study).^{19 20}

RaCTs also possess many of the same disadvantages inherent to any randomised trial—namely, that they follow strict protocols using carefully selected patient groups that may not be representative of the wider patient population, thereby hindering their generalisability to routine clinical practice. One recent example is dabigatran, used in atrial fibrillation. Its primary trial (the randomised evaluation of long term anticoagulant therapy trial, or RE-LY), an active comparator trial, demonstrated that dabigatran was similar to (and possibly better than) warfarin, based on a number of clinical outcomes. However, as RE-LY excluded patients with relevant comorbidities such as renal insufficiency, superiority claims were complicated by concerns over the generalisability of findings. For instance, trial results may not be applicable to settings where patients receive warfarin with excellent international normalised ratio control.²¹

In light of some of these challenges, other research approaches have been proposed to generate comparative efficacy evidence, offering more flexibility than traditional RCT designs in terms of their design, conduct, and implementation. These include, but are not limited to, adaptive and pragmatic trial designs, prospective network meta-analyses, and observational studies.

RCTs can be designed with adaptive features that allow for ongoing changes in trial design in response to emerging external information (for example, evidence from systematic reviews or observational studies).²² Such features provide opportunities to include (or exclude) relevant comparators and to reduce the sample size, time, or cost requirements of studies, improving both the efficiency and relevance of trial results. However, there is concern that following adaptations the actual patient population could deviate from the original target patient population, leading to erroneous claims of efficacy. Major adaptations of trials may also result in a trial that is unable to address the questions it intends to answer.

Pragmatic clinical trials impose fewer restrictions on patient populations and practice settings than RCTs by conducting studies in routine clinical settings, and generally allow for inclusion of a wider range of outcomes (quality of life, longer term effect, for example).²³ Consequently, they are better able to parallel clinical practice, capturing both comparative efficacy and effectiveness information. However, as a subset of RCTs, pragmatic clinical trials share some of the similar methodological and ethical challenges. Moreover, they often require higher sample sizes than traditional clinical trials given their less tightly controlled design.

Box 2 Advantages and disadvantages of study design options for assessing comparative efficacy*Traditional active comparator randomised controlled trials (RaCTs)*

Patients are randomised to receive an experimental or standard treatment (with or without an additional group receiving placebo).

Advantages: Gold standard for determining whether a drug works—high internal validity; most informative if three armed design, including the experimental drug, standard treatment, and placebo. Randomisation ensures that there are minimal systematic differences between patient groups before treatment. Carefully controlled research environments in traditional designs ensure minimal systematic differences in how groups receive treatment.

Disadvantages: Demonstrating superiority, equivalence, or non-inferiority is challenging as the number of active comparators increases. Strict selection criteria in traditional designs may limit generalisability of findings to broader patient populations in routine clinical practice. Large sample sizes usually required. Costly and take a long time to complete. Owing to short durations, important side effects may go undetected. Address comparative efficacy only.

Adaptive clinical trials

A type of randomised controlled trial designed to change or adapt in response to information generated during the trial.

Advantages: Allow trial design to be changed during the course of the study based on new data. May reduce the sample size, time, or cost requirements of studies, thereby enhancing efficiency. Ability to include or exclude relevant comparators can enhance clinical relevance of trial results.

Disadvantages: Adaptations may result in difficulty addressing original research questions. Changes in patient population may deviate from original target population.

Pragmatic clinical trials

A type of randomised controlled trial (RCT) of viable treatments specifically designed to address practical questions about the risks, benefits, and costs of an intervention as they would occur in routine clinical practice.

Advantages: High external validity. Impose only minimal restrictions in patient population and research environment to parallel routine clinical practice. Can address both comparative efficacy and effectiveness.

Disadvantages: Subset of RCTs; therefore similar methodological and ethical challenges. May mask small true differences between treatments.

Network meta-analyses

A form of meta-analysis comparing multiple treatments using data from direct and indirect comparisons.

Advantages: Maintain randomisation within individual trials. Assumptions are generalisations of those in established pair wise meta-analyses. Multiple drugs can be simultaneously compared in an internally coherent analysis by combining evidence from head to head trials and indirect comparisons. Less expensive than randomised trials. When designed prospectively, analyses can deliberately introduce heterogeneity, enhancing the generalisability of trial findings. Can address both comparative efficacy and effectiveness.

Disadvantages: Indirect evidence may not be consistent with direct evidence. Share the same limitations as randomised trials. Although randomisation is maintained within trials, the relation across trials is observational. Clinical and statistical comparability of trials (homogeneity, similarity, and consistency of evidence) may require expert judgment.

Observational studies

A type of study where patients receiving a particular treatment are observed rather than being assigned to treatment randomly.

Advantages: Representative of routine clinical practice by observing actual patient and prescriber practices. Possibility of evaluating a large number of comparators at relatively low cost and high speed. Useful for studies of rare conditions without any known treatments and for examining multiple treatment paradigms simultaneously.

Disadvantages: Prone to confounding: measurable or immeasurable factors (such as underlying differences in patient groups before treatment, differences in patient selection, and adherence to treatment) may affect the outcomes of interest. Methodological tools to minimise confounding generally cannot remove all bias. Information normally not available at the time of market approval.

Network meta-analyses, an extension of meta-analyses, assess the comparative efficacy of treatments when they have not been compared directly in a randomised trial, but have each been compared with other treatments.²⁴ If designed prospectively, these techniques allow estimation of heterogeneity in the effect of a drug, enhancing the generalisability of trial findings. These approaches are particularly valuable when there are a large number of existing comparators. However, synthesis across trials cannot always provide sufficient assurance that underlying differences in patient characteristics are identical or evenly distributed across trials.²⁵

In the absence of RCTs, observational studies allow for retrospective analysis of data collected in routine clinical practice. Observational methods offer the possibility of

evaluating a large number of comparators at relatively low cost and high speed. Moreover, they provide information in a routine care setting, including critical patient groups that are often omitted from RCTs.²⁵ However, owing to their non-randomised design, they are susceptible to potential confounding. Given that observational data are not normally available at the time of market authorisation, this approach alone would not be sufficient to assess pre-market comparative efficacy.

In view of these opportunities and challenges, it is clear that comparative efficacy assessment is not necessarily a panacea. No particular study design alone is ideal for assessing comparative efficacy and there is the inherent challenge of ensuring studies adequately address the most relevant clinical and policy questions to support safe and effective use of new

therapies. Some approaches may be more appropriate under certain conditions (such as with complex patient populations) and with different goals in mind (safety assessment, generalisability, for example), and the accurate interpretation of resulting evidence is often dependent on robust design and effective execution. Rather, available study designs should be considered as potential complementary methodological tools. For example, the EMA might recommend that submissions be supported by one traditional and one pragmatic RCT and, where possible, collect cost and quality of life data. This would help ensure that clinically relevant evidence on the benefits and risks of new drugs is collected and considered, and would facilitate meeting the downstream information needs of health technology assessment agencies and payers. Similarly, observational studies or network meta-analyses accompanying submission of clinical trials could strengthen pre-market comparative assessments.

Future directions toward formalising the role of comparative efficacy in drug licensing

Comparative efficacy assessment at the time of drug approval serves as an important tool to help ensure that the most beneficial and safest treatments reach patients and that limited healthcare resources are invested wisely. Therefore, despite existing methodological complexities, comparative efficacy evidence should have a formal role in drug licensing decisions; a position increasingly supported by the EMA and European Commission.^{26 27} The EMA, in particular, has recommended RaCTs when a new drug might be associated with safety or inferiority concerns.²⁶ However, we propose that non-inferiority be recommended for all conditions where alternative treatment options exist, which would help address the current variability in availability of such evidence across therapeutic areas (see table 1). While evidence of superiority is ideal, regulators would argue that prescribers need multiple options available on the market to individualise therapy for patients.²⁸ It will, of course, be important to further substantiate such scenarios and transparently outline study requirements (for example, the potential requirement of an RaCT and PCT) and possible exceptions. If study criteria or evidence standards are set unrealistically high or are unclear, manufacturers may prematurely terminate development programmes for potentially valuable drugs, or authorisation may be delayed. In attempts to arrive at a feasible approach, regulatory authorities could consider use of so called progressive licensing in certain contexts. For example, a drug could be initially approved for a defined population based on short term trials collecting surrogate outcomes, but later extended to a broader population following availability of post-market data with clinically meaningful outcomes. The EMA has expressed interest in this approach.²⁹

Moving toward these ends will also require strategies to enhance the performance and accessibility of comparative efficacy studies. An important first step is to achieve open dialogue and agreement between the EMA, manufacturers, payers, and governments on the outstanding methodological questions and hurdles associated with these types of studies. An independent EMA advisory board could be established to lead such discussions and ultimately provide guidance on: what type of study design(s) can be accepted as fit for purpose for generating comparative efficacy evidence; standards for comparator selection and assessed endpoints; dosing schemes; sample size requirements to demonstrate the margin of superiority, equivalence or non-inferiority; and options to make the conduct of comparative studies more efficient. In addition, the board

should collaborate with national health technology assessment agencies (such as the National Institute for Health and Clinical Excellence) to achieve better congruence between licensing and reimbursement requirements.

Finally, an important limitation of present regulatory systems is the suboptimal publication of and public access to comparative information on the efficacy, effectiveness, and safety of new drugs. Regulators should therefore require submission of full data, including trial protocols and raw datasets, and adopt and enforce a policy to publicly share available data.^{18 30 31} More effective data exchange could entail housing protocols and datasets within trial registries, enhancing the quality and transparency of the European Public Assessment Reports, reporting comparative evidence in product labels and marketing materials, and through journals adopting similar data submission requirements.^{3 18 32} These efforts would help ensure that researchers have access to detailed enough data to allow independent re-analysis of trials and potentially contribute to improving post-market comparative effectiveness assessment. Numerous promising medicines have been developed and many more are on the way to initial clinical trials. With this success comes an equally important additional need—to develop a systematic approach to evaluate the risks and benefits of these new therapies in the context of existing alternatives. An important initial step is to support a formal role for comparative efficacy evidence in drug licensing.

Contributors and Sources: CS, HN, and EM have written several articles on pharmaceutical policy and regulation and comparative efficacy and effectiveness research in the EU and the US. CS and HN have experience in comparative efficacy and comparative effectiveness research methodological approaches. JC and EM have extensive experience of European health systems through their involvement with the European Observatory on Health Systems and Policies. CS, HN, JC, and EM devised the article. CS wrote the first draft. CS, HN, JC, and EM contributed to subsequent drafts. All have read and agreed to the final version. CS is guarantor.

Competing interests: All authors have completed the ICMJE uniform disclosure form at www.icmje.org/coi_disclosure.pdf (available on request from the corresponding author) and declare: no support from any organisation for the submitted work; no financial relationships with any organisations that might have an interest in the submitted work in the previous three years, no other relationships or activities that could appear to have influenced the submitted work.

Provenance and peer review: Commissioned; externally peer reviewed.

- 1 The International Conference on Harmonization/Committee for Human Medicinal Products guidelines (CPMP/ICH/364/96) define this as “a pharmacological treatment that is shown to be lifesaving or to prevent irreversible mortality”. In this case, non-inferiority comparative trials are required. European Medicines Agency. EU standard of medicinal product registration: clinical evaluation of risk/benefit—the role of comparator studies. EMEA/119319/04. 2004. www.ema.europa.eu/pdfs/human/press/pos/11931904en.pdf.
- 2 van Luijn JC, Gribnau FW, Leufkens HG. Superior efficacy of new medicines? *Eur J Clin Pharmacol* 2010;66:445-48.
- 3 Stafford RS, Wagner TH, Lavori PW. New, but not improved? Incorporating comparative-effectiveness information into FDA labelling. *N Engl J Med* 2009;361:1230-3.
- 4 Garattini S, Bertele V. How can we regulate medicines better? *BMJ* 2007;335:803-5.
- 5 Smith R. Medical journals are an extension of the marketing arm of pharmaceutical companies. *PLoS Med* 2005;2:e138.
- 6 Maio V, Gagne JJ. Impact of ALLHAT publication on antihypertensive prescribing patterns in Regione Emilia-Romagna, Italy. *J Clin Pharm Ther* 2010;35:55-61.
- 7 Juurlink DN. Rosiglitazone and the case for safety over certainty. *JAMA* 2010;304:469-71.
- 8 Loke YK, Kwok CS, Singh S. Comparative cardiovascular effects of thiazolidinediones: a systematic review and meta-analysis of observational studies. *BMJ* 2011;342:d1309.
- 9 Nissen S. Rosiglitazone: a critical overview. 2010. www.fda.gov/downloads/AdvisoryCommittees/CommitteeMeetingMaterials/Drugs/EndocrinologicalandMetabolicDrugsAdvisoryCommittee/UCM218483.pdf.
- 10 Eichler HG, Bloechl-Daum B, Abadie E, Barnett D, Konig F, Pearson S. Relative efficacy of drugs: an emerging issue between regulatory agencies and third-party payers. *Nat Rev Drug Discov* 2010;9:277-291.
- 11 Garattini S, Chalmers I. Patients and the public deserve big changes in the evaluation of drugs. *BMJ* 2009;338:b1025.

- 12 Motola D, De Ponti F, Poluzzi E, Martini N, Rossi P, Silvani MC, et al. An update on the first decade of the European centralized procedure: how many innovative drugs? *Br J Clin Pharmacol* 2006;62:610-6.
- 13 ALLHAT Officers and Coordinators for the ALLHAT Collaborative Research Group. Major outcomes in high risk hypertensive patients randomized to angiotensin-converting enzyme inhibitor or calcium channel blocker vs. diuretic: The Antihypertensive and Lipid-Lowering Treatment to Prevent Heart Attack Trial (ALLHAT). *JAMA* 2002;288:2981-97.
- 14 van Luijn JC, Stolk P, Gribnau FW, Leufkens HG. Gap in publication of comparative information on new medicines. *Br J Clin Pharmacol* 2007;65:716-22.
- 15 European Social Insurance Platform. Response to the European Commission public consultation on the future of pharmaceuticals for human use in Europe. 2007. www.esip.org/publications/pb139.pdf.
- 16 van Luijn JC, Gribnau FW, Leufkens HG. Availability of comparative trials for the assessment of new medicines in the European Union at the moment of market authorization. *Br J Clin Pharmacol* 2006;63:159-62.
- 17 Goldberg NH, Schneeweiss S, Kowal MK, Gagne JJ. Availability of comparative efficacy data at the time of drug approval in the United States. *JAMA* 2011;305:1786-9.
- 18 Jefferson T, Doshi P, Thompson M, Heneghan C, Cochrane Acute Respiratory Infections Group. Ensuring safe and effective drugs: who can do what it takes? *BMJ* 2011;342:c7258.
- 19 van Luijn JC, van Loenen AC, Gribnau FW, Leufkens HG. Choice of comparator in active control trials of new drugs. *Ann Pharmacother* 2008;42:1605-12.
- 20 Chokshi DA, Avorn J, Kesselheim AS. Designing comparative effectiveness research on prescription drugs: lessons from the clinical trial literature. *Health Aff (Millwood)* 2010;29:1842-8.
- 21 Gage BF. Can we rely on RE-LY? *N Engl Med J* 2009;361:1200-2.
- 22 Luce BR, Kramer JM, Goodman SN, Connor JT, Tunis S, Whicher D, et al. Rethinking randomized clinical trials for comparative effectiveness research: the need for transformational change. *Ann Intern Med* 2009;151:206-9.
- 23 Tunis SR, Stryer DB, Clancy CM. Practical clinical trials. Increasing the value of clinical research for decision making in clinical and health policy. *JAMA* 2003;290:1624-32.
- 24 Song F, Loke YK, Walsh T, Glenny AM, Eastwood AJ, Altman DG. Methodological problems in the use of indirect comparisons for evaluating healthcare interventions: survey of published systematic reviews. *BMJ* 2009;338:b1147.
- 25 Fleurence RL, Naci H, Jansen JP. The critical role of observational evidence in comparative effectiveness research. *Health Aff (Millwood)* 2010;29:1826-33.
- 26 European Medicines Agency. Reflection paper on the need for active control in therapeutic areas where use of placebo is deemed ethical and one or more are available. EMA/759784/2010. November 2010.
- 27 European Commission. Final conclusions and recommendations of the High Level Pharmaceutical Forum. 2008. http://ec.europa.eu/pharmforum/docs/final_conclusions_en.pdf.
- 28 Gagne JJ, Choudhry NK. How many "me-too" drugs is too many? *JAMA* 2011;305:711-2.
- 29 European Medicines Agency. Road map to 2015: The European Medicines Agency's contribution to science, medicines and health. 2010.
- 30 Garattini S, Bertele V. Europe's opportunity to open up drug regulation. *BMJ* 2010;340:c1578.
- 31 Chan A. Access to clinical trial data. *BMJ* 2011;342:d80.
- 32 Barbui C, Baschiroto C, Cipriani A. EMA must improve the quality of its clinical trial reports. *BMJ* 2011;342:d2291.

Accepted: 22 July 2011

Cite this as: *BMJ* 2011;343:d4849

© BMJ Publishing Group Ltd 2011

Table

Table 1 | Potential scenarios for current versus proposed comparative efficacy evidence requirements in drug approvals

Potential scenarios	Current requirements	Proposed requirements
New medicine in a therapeutic area where no pharmacological treatment is available	Comparative efficacy cannot be required	Comparative efficacy cannot be required
New medicine in a therapeutic area where placebo is deemed unethical and active control exists	Comparative efficacy to active control is required	Comparative efficacy to active control is required
New medicine in a therapeutic area where placebo is deemed ethical and one or more established medicines are available	Comparative efficacy to active control is recommended if: (1) the new medicine might be associated with safety concerns that influence mortality and morbidity, markedly impair quality of life, or cause active treatment to be discontinued or delayed leading to significant, long term, or irreversible harm; (2) treatment with a medicine of inferior efficacy might conceivably lead to significant, long term, or irreversible harm for the patient.	Comparative efficacy to active control is required for all conditions where alternative medicines exist

Sources: European Medicines Agency. Reflection paper on the need for active control in therapeutic areas where use of placebo is deemed ethical and one or more established medicines are available. EMA/759784/10. Nov 2010. European Medicines Agency. EU standard of medicinal product registration: clinical evaluation of risk/benefit—the role of comparator studies. EMA/119319/04. Oct 2004.