Real Medical Benefit Assessed by Indirect Comparison

Bruno Falissard[1,2,3], Myriam Zylberman[4], Michel Cucherat[5], Valérie Izard[6], François Meyer[6]
and the participants in Round Table N°5, Giens XXIV[*]

1 INSERM, U669, Paris, France
2 Université Paris-Sud et Université Paris Descartes, UMR-S0669, Paris, France
3 AP-HP, Hôpital Paul Brousse, Département de Santé Publique, Villejuif, France
4 LILLY France, Suresnes, France
5 Faculté de Médecine Laënnec, EA 3736, Lyon, France
6 Haute Autorité de Santé (HAS), Saint-Denis La Plaine, France

Abstract – Frequently, in data packages submitted for Marketing Approval to the CHMP, there is a lack of relevant head-to-head comparisons of medicinal products that could enable national authorities responsible for the approval of reimbursement to assess the Added Therapeutic Value (ASMR) of new clinical entities or line extensions of existing therapies.

Indirect or mixed treatment comparisons (MTC) are methods stemming from the field of meta-analysis that have been designed to tackle this problem. Adjusted indirect comparisons, meta-regressions, mixed models, Bayesian network analyses pool results of randomised controlled trials (RCTs), enabling a quantitative synthesis.

The REAL procedure, recently developed by the HAS (French National Authority for Health), is a mixture of an MTC and effect model based on expert opinions. It is intended to translate the efficacy observed in the trials into effectiveness expected in day-to-day clinical practice in France.

1. Introduction

In France, the assessment of the potential effectiveness of new pharmaceutical entities is carried out with assistance from a large body of experts by the Transparency Committee of the Haute Autorité de Santé (HAS, French National Authority for Health). The 20 appointed voting members rate the added value (ASMR for Amélioration du Service Médical Rendu, Improvement in Actual Benefit) of the medicinal product from 1 (major improvement) to 5 (no improvement). This rating is then used straightaway by the French Ministry of Health to help determine the price and reimbursement rate.

In practice, pharmaceutical companies must submit a standard report of all their efficacy and safety data on the medicinal product to the Transparency Committee; these data are obtained in most cases from randomised controlled trials (RCTs). Unfortunately, in many situations there is a lack of head-to-head trials comparing the product under evaluation with the most relevant competing medicinal products. [1]

To deal with this problem, it is often suggested to integrate data from a series of trials to produce an indirect or a mixed treatment comparison (MTC). [2,3] For example, results of RCTs comparing A to a placebo and those comparing B to a placebo can be pooled to obtain, indirectly, a comparison of A *versus* B.

The objective of the present paper is to determine whether such methodologies can be used routinely for the assessment of the potential effectiveness of new medicinal products based on the marketing authorization package at launch. In a first part, we shall detail the statistical methods that are presently available. Since the data are most frequently obtained from phase III RCTs, the evaluation will focus on efficacy ("does the drug work under optimal circumstances?"), [4] rather than effectiveness ("does the drug

* For a list of participants, see the end of the article.

work in routine care?").[4] In a second part, we shall thus tackle the question of the indirect comparison of effectiveness, which is central in the context of the comparative effectiveness assessment of new medicinal products with existing therapies.

2. Indirect and mixed treatment comparisons

Basically, it may be tempting to deal with a lack of direct comparisons by qualitatively comparing the level of the effect of competing medications. For example, experts often examine whether the confidence intervals of different treatment effects overlap or not. This is definitely not recommended.[5]

The next approach that naturally comes to mind is referred to as "unadjusted indirect comparisons" and consists in collecting all the RCTs involving treatment A and all the RCTs involving treatment B. The average improvement of patients receiving A is then compared to the average improvement of patients receiving B. This approach nevertheless has a major drawback: it negates the randomised nature of each individual RCT and thus should also be avoided.[5]

Adjusted indirect comparisons deal with this problem and, for this reason, only consider RCTs comparing, for example, A and Placebo and RCTs comparing B and Placebo. The differences in the improvement of patients receiving A and those receiving P are then averaged and compared to the differences in the improvement of patients receiving B and those receiving P. The methodology is thus formally similar to that used in meta-analyses so that, as it is the case in this last area, the question of heterogeneity becomes crucial. Indeed, the different RCTs included in the analysis are performed at different times, in various populations, possibly with different dosages. This heterogeneity can lead to possible bias. For example, if all A-Placebo trials are conducted on severe patients (so that the difference between A and the placebo is possibly great) and if all the B-placebo trials are conducted on moderately ill patients (so that the difference between B and the placebo will be smaller), then the A-B comparisons from the indirect comparison are likely to be inaccurate.

The concept of meta-regression[5] has been typically developed to deal with this drawback: confounders (severity of patients in the example above) can be added to the model so that some known possible sources of bias can be taken into account. In the same way, mixed-models have been proposed: a variable "study" is considered in this case as a random effect, which can help in the generalisation of the results.

Finally, a more sophisticated approach called network meta-analysis has recently been developed.[2] Generally based on Bayesian models, it can handle a wide variety of situations: several medicinal products can be compared indirectly without a single common comparator having to be available.

An important question about indirect comparisons and MCT relates to their validity. A paper[6] has compared results obtained from direct comparisons and results obtained from adjusted indirect comparisons. Among the 44 situations considered (comparison of 2 treatments for a given pathology), there was an inconsistency between both approaches in only 3 cases.

3. Generalisation of MTC findings to real-life practices

The improvement of actual benefit as assessed by the Transparency Committee cannot rely solely on results of phase III RCTs designed to obtain FDA (Food and Drug Administration) or EMEA (European Medicines Agency) approval. Indeed, these RCTs are generally conducted in countries with specific health care systems; patients enrolled must verify specific enrolment criteria; moreover, the follow-up and the physicians involved in these trials are possibly not representative of day-to-day practice.

Ideally, real-life RCTs or observational studies should be conducted in each country where a new medication is available. In practice, at the present time, these studies are the exception. To deal with this problem, the HAS has recently proposed a specific methodology referred to as REAL (Relative Effectiveness Assessment of pharmaceuticals).[7] REAL consists in a three-phase procedure:

1/ First step: collection of RCT data. A list of trials is currently submitted by the firm in a standard report. Other trials can be added by the Commission according to a systematic formal review of the literature or according to expert advice.

2/ Second step: collection of data on the suitability for generalisation and transposition of the results of the RCTs selected in step 1. This step is based on an elementary effect model.[8] More precisely, it is postulated that there is a linear relationship between the relative efficacy of A *versus* B as given by an RCT [this relative efficacy is expressed in this case by a log odds-ratio denoted "log(ORRCT)"] and the expected relative effectiveness of A *versus* B estimated ideally in a real-life study [*i.e.* log(ORRL)]. Let:

$$E[\log(ORRL)] = a + E[\log(ORRCT)] \qquad (1)$$

and

$$Var[\log(ORRL)] = b \times Var[\log(ORRCT)](b \geqslant 1) \qquad (2)$$

If a<0, then the effectiveness will be less than the efficacy observed in the trial. If a>0, the converse is true. The constant b reflects a lack of confidence in the estimation of log(ORRL) as compared to the confidence available for the estimation of

log(ORRCT). The constants a and b will be obtained from the answer of the 20 experts of the Transparency Committee to 4 questions, obtained using visual analogue scales:

Question 1: Tick the segment below that best reflects your lack of confidence in the relative efficacy (Re) of the medicinal product due to the potential methodological flaws of the trial.

Question 2: Tick the segment below that best reflects your lack of confidence in the relative effectiveness (Rr) of the medicinal product under real-life conditions due to the problems involved in transposing trial results to real life.

Question 3: Tick the segment below that best reflects your perception of the relationship between the relative efficacy (Re) of the medicinal product observed in the trial and the relative effectiveness (Rr) that might be observed under real-life conditions after short-term administration (a few months, a few years). Base your answer on your clinical experience and knowledge of clinical trials.

Question 4: In light of your clinical experience, do you think that the relative effectiveness of A and B might change significantly in the long term: yes/no. If yes, tick the segment below that best reflects the change you would anticipate in the relative effectiveness Rr of the two medicinal products in the long term (patient's life expectancy) in real life.

If b1 and b2 denote the answers to questions 1 and 2, and a1 denotes the answer to question 3, and if the answer to question 4 is "no", then a=log(a1) and b=1/(b1×b2), and the relative effectiveness log(ORRL) and its confidence can be calculated using equations (1) and (2). If the answer to question 4 is "yes" and is denoted a2, then a=log(a1×a2) and b=1/(b1×b2).

3/ Third step: pooling of log(ORRL), sensitivity analysis. Pooling is performed using unidimensional scaling,[9] the sensitivity analysis using a bootstrap.

Example: an antidepressant drug A is proposed in the treatment of generalised anxiety disorder. Four trials are available. The first compares A (75 mg per day) to placebo in 160+168 patients. The second trial also compares A (75 mg per day) to placebo in 160+168 patients. The third trial compares A (75 mg per day), A (150 mg per day), B (40 mg per day) to placebo in 157+158+160+154 patients. The fourth trial also compares A (75 mg per day), A (150 mg per day), B (40 mg per day) to placebo in 156+160+151+157 patients. As presented above, the effect model is obtained from the experts' ratings, the means of a1, a2, b1 and b2 for all the experts and all trial arms are respectively 1.13, 0.94, 0.88 and 0.69. It is noteworthy that a1×a2 is greater than 1. In other words, it is suggested by the experts that the expected relative effectivenesses of A and B are, on average, slightly superior to the relative efficacies observed in the trials. A possible interpretation that could explain this apparently unexpected result is that patients in the placebo arm of a phase III RCT
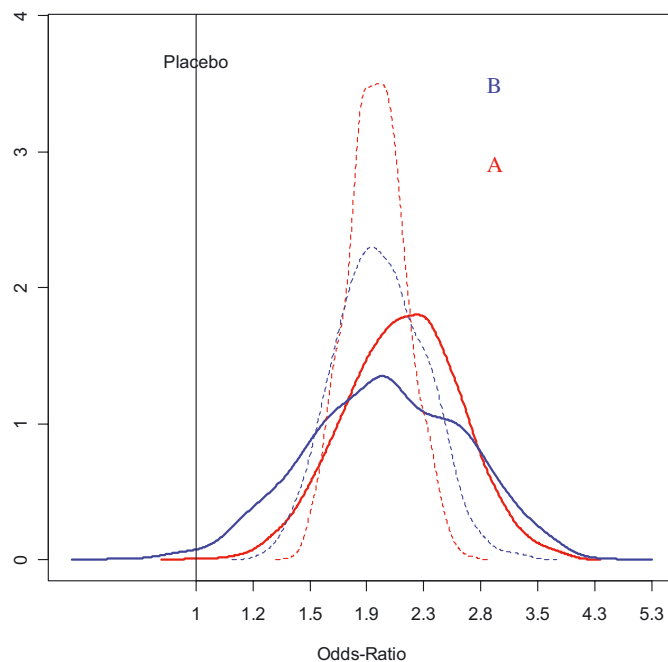

Fig. 1. Relative effectiveness (solid lines) and relative efficacy (dashed lines) *versus* placebo of 2 treatments (A and B) in advanced generalised anxiety disorder.

in GAD (Generalized Anxiety Disorder), are likely to receive, *de facto*, an effective non-pharmacological treatment. Such a thorough follow-up is, however, infrequent in day-to-day practice (at least in France) so that the relative effectiveness (in a real-life setting) of treatment A or B against no medication is likely to be superior to the relative efficacy of A or B against placebo.

The results of the REAL procedure applied to these trials are presented in figure 1. In terms of effectiveness, treatment A appears to be superior to placebo, with an effect that is comparable to treatment B. Relative efficacies are also shown (for estimates obtained where a=0 and b=1) in order to visualize the influence of the experts' ratings on the distributions. As expected, there is an increase in the variance of the relative effectiveness compared to the relative efficacy; this increase in variance is due to the uncertainty of the generalisation and transposition of the results of the RCTs to a real-life setting. There is also a slight shift between the relative effectivenesses and the relative efficacy; this is consistent with the remarks exposed above.

4. Conclusion

There is no doubt that indirect or mixed treatment comparisons require a rigorous methodology: a thorough formal review of the literature and the use of appropriate statistical models. Sensitivity analyses need to be conducted to assess the robustness of

findings. There are many limitations, the main one being related to the heterogeneity of the RCTs included in the analysis, and another to the minimum number of trials that should be available to use sophisticated methods like mixed models for MTCs. A minimum as high as 4 trials per comparison of treatments has been proposed.[10]

In practice, it is suggested that indirect comparisons or MTC should be performed when no direct comparisons are available or when the direct comparisons relate to a non-optimal comparator or RCTs with methodological shortcomings.

As regards the transposition and generalisation of a quantitative synthesis of RCT results, the HAS is currently developing the REAL procedure. This procedure consists in a mixture of MTC with an effect model. This effect model is obtained from expert opinions and is intended to translate the efficacy observed in trials into effectiveness expected in day-to-day clinical practice in France. The distributions of the effectiveness obtained in this way are then used by the Transparency Committee for a more reliable review of the Improvement of Actual Benefit. At the present time, the REAL procedure is in an experimental phase and should be used routinely in the near future. Since REAL is partly based on the subjective opinions of experts, it will need to be evaluated following a trial phase.

Participants

Alexandre Barna (HAS), Michel Cucherat (Hôpital Laënnec, Lyon), Frédéric De Bels (HAS), Rima de Sahb (Laboratoires Roche), Jean Delonca (Laboratoire Bristol Myers Squibb), Valérie Izard (HAS), Pascale Jolliet (CHU Nantes), Jean-Michel Joubert (UCB Pharma), Robert Launois (Réseau Évaluation en Économie de la Santé), Claire Le Jeunne (Service de Médecine Interne, Hôtel Dieu), Michel Lièvre (Faculté de Médecine, Lyon), Florence Mahlberg Gaudin (Laboratoires Servier), Aurélie Moser (Laboratoire Roche), Patricia Perles (Laboratoire Sanofi Aventis), Martine Pigeon (CNAMTS), Cécile Rey Coquais (Laboratoire Wyeth Pharmaceutical), Nadine Weisslinger (Laboratoire GlaxoSmithKline), Bertrand Xerri (HAS).

References

1. Ioannidis JP. Indirect comparisons: the mesh and mess of clinical trials. Lancet; 2006: (368): 1470-2
2. Lu G, Ades AE. Combination of direct and indirect evidence in mixed treatment comparisons. Star Med 2004; 20: 3105-24
3. Cucherat M. Les comparaisons indirectes : méthodes et validité. 2009, Haute Autorité de Santé - Service évaluation des médicaments: Paris. p. 66
4. Bombardier C, Maetzel A. Pharmacoeconomic evaluation of new treatments: efficacy *versus* effectiveness studies? Ann Rheum Dis 1999; 58 Suppl 1: I82-5
5. Gartlehner G, Moore CG. Direct *versus* indirect comparisons: a summary of the evidence. International Journal of Technology Assessment in Health Care 2008; 24(2): 170–7
6. Song F, Altman DG, Glenny AM, *et al*. Validity of indirect comparison for estimating efficacy of competing interventions: empirical evidence from published meta-analyses. BMJ 2003; 326 (7387): 472 [Letter]
7. Falissard B, Izard V, Xerri B, *et al*. Relative Effectiveness Assessment of Listed Drugs (REAL): a new method for an early comparison of the effectiveness of approved health technologies. International Journal of Technology Assessment in Health Care, 2009. in press
8. Boissel JP, Collet JP, Lievre M, *et al*. An effect model for the assessment of drug benefit: example of antiarrhythmic drugs in postmyocardial infarction patients. J Cardiovasc Pharmacol 1993; 22(3): 356-63
9. De Leeuw J. Unidimensional scaling, in encyclopedia of statistics in behavioral science, BS Everitt and D Howell, editors. 2005, John Wiley & Sons
10. Glenny AM, Altman DG, Song F, *et al*. International stroke trial collaborative group: indirect comparisons of competing interventions. Health Technol Assess 2005; 9(26): 1-134

Correspondence and offprints: *Bruno Falissard*, INSERM U 669, Maison de Solenn, 97 boulevard de Port Royal, 75679 Paris Cedex 14, France
E-mail: falissard_b@wanadoo.fr