Well-conducted meta-analyses are powerful and useful instruments to produce synthesis of all the evidence available on a specific clinical issue. In recent years, like in other fields of oncology, the number of meta-analyses conducted in the setting of advanced non-small-cell lung cancer has consistently grown. Results of meta-analyses are considered as strong evidence in all the most important guidelines. However, heterogeneity of results among clinical trials included in a meta-analysis should be investigated and discussed, because it might significantly affect the interpretation of results and their application in clinical practice. After demonstration of a statistically significant advantage, the magnitude of the observed benefit should always be weighed against risks and costs of treatments.

Keywords: advanced NSCLC • afatinib • bevacizumab • cetuximab • clinical trial • crizotinib • erlotinib • gefitinib • individual patient data • meta-analysis

Meta-analyses are a precious instrument that allow production of synthesis of all the evidence available on a specific clinical issue [1]. Of course, a meta-analysis should be performed only after the conduction of an accurate systematic review of all the evidence available on that specific topic [2]. A systematic review is a process that uses systematic and explicit methods to identify, select and critically appraise relevant evidence available about a specific question, and to collect and analyze data from the studies that are included in the review. A meta-analysis is based on the use of statistical procedures, when appropriate, to integrate the results of studies previously identified in a systematic review. In recent years, the number of meta-analyses conducted in all the fields of oncology, and also in non-small-cell lung cancer (NSCLC), has rapidly grown (Figure 1). This is also due to the availability and diffusion of dedicated software that make it relatively easy and quick to perform meta-analyses based on data extracted from the literature, so it is not difficult to explain the ‘mushrooming’ number of published meta-analysis based on abstracted data.

Although the decisions of regulatory agencies about the approval of a specific drug are often based on a single (or a few) registrative trial, meta-analyses are a useful instrument to increase the precision in the estimate of treatment effect. According to an evidence-based approach, meta-analysis results are considered a strong level of evidence in all the most important guidelines. In 1995, when there was still considerable pessimism and skepticism about the role of chemotherapy in the treatment of patients with advanced NSCLC, the publication of a meta-analysis demonstrating a significant benefit in overall survival with the addition of chemotherapy to best supportive care represented a milestone in the knowledge about the best management for these patients [3]. Nearly 20 years later, many meta-analyses have been published on a high number of clinical issues that are relevant for the management of patients with advanced NSCLC. The quality of these meta-analyses is high or very high in...
some cases, but much lower in others. In addition, their relevance for clinical practice is very high in some cases, but questionable in others.

**Individual patient data or summary data?**

Conduction of meta-analyses based on individual patient data (IPD), implies the collection of original data from all the existing trials [1]. Without doubt, if compared with meta-analyses based on data extracted from the publications, IPD meta-analyses are time-consuming and more difficult to perform. Based on the authors experience [4–6], I would estimate that performing an IPD meta-analysis – in a setting where the number of existing trials is not enormous – can take, more or less, 2 years, from protocol writing to the submission of final results for publication. The process can be particularly difficult if some of the trials have not been conducted recently, because those data sets could not be easily retrieved.

One might be skeptical about the real utility of performing IPD meta-analyses if an informative result can be easily and more quickly obtained with abstracted data. Furthermore, when meta-analyses based on abstracted data are conducted accurately, they will probably produce results that are similar to the results obtained with individual patient data [7]. So, what is the reason, if any, to make such a relevant effort and base the meta-analysis on individual patient data? First, the availability of the database will allow analysis to reproduce and verify the results presented in the original study publication, and this increases the quality of the analysis. Second, the IPD approach allows the use of updated data, compared with the data use for each single publication. I recognize that this could be not particularly relevant in the case of trials dedicated to advanced disease, where the number of events is already high and the follow-up period short, but an update of follow-up could be particularly relevant in the setting of adjuvant treatments. If the follow up is updated several years after the original publication, this could produce a significantly higher number of events for the analysis, thus obtaining a more precise estimate of treatment effects compared with the original publications. Third, using individual data of the patients allows the preplanning of exploratory analyses to identify groups of patients who may benefit more from the treatment [1]. Advanced NSCLC is a very heterogeneous disease, from both a clinical and molecular point of view, and the average effect of treatments that can be precisely estimated in a meta-analysis is not the only information of interest, since knowledge of predictive factors, associated with a greater or smaller efficacy of treatments, is particularly relevant. In the context of meta-analysis based on data abstracted from the publications, the potential impact of patients’ characteristics on treatment effect (if subgroup analysis was not already performed in each single trial and available in the publication) is only ‘roughly’ possible with the application of meta-regression techniques. On the contrary, availability of IPD will allow the conduction of subgroup analyses, such as in a single, large randomized trial, with the obvious advantage of a greater sample size. Of course, all the limitations of subgroup analysis are
The role of meta-analysis in defining clinical practice in advanced NSCLC

Clinical Trial Perspective

The heterogeneity of clinical trials considered in a meta-analysis

To be rationally combined in a meta-analysis, different clinical trials should address the same clinical question. However, some level of heterogeneity among the trials cannot be avoided [8,9]. When discussing the heterogeneity of trials included in a meta-analysis, clinical heterogeneity should not be confounded with statistical heterogeneity. Clinical heterogeneity is related to variability in patients’ characteristics, type of interventions and outcomes studied in the trials considered in a meta-analysis. On the other hand, statistical heterogeneity is present when results of trials included in a meta-analysis are not consistent with each other; for example, when experimental treatment shows a benefit in some trials and harm in other trials, or the magnitude of benefit or harm is not consistent among trials. Statistical heterogeneity can be assessed only after performing the analysis: this heterogeneity can be measured with quantitative methods [8]. When the difference is too large to be consistent with chance, this represents a substantial problem for the validity and for the interpretation of the meta-analysis. On the contrary, clinical heterogeneity should be analyzed before performing the analysis and its evaluation is qualitative. For instance, the trials can be heterogeneous in terms of patients’ inclusion and exclusion criteria, in terms of treatment administered and in terms of exams and procedures used for patient monitoring and follow-up. These differences can be more or less relevant for the interpretation of the meta-analysis results. For instance, when we decided to perform a meta-analysis of trials comparing weekly versus every 3-week administration of docetaxel as a second-line treatment of advanced NSCLC, we pooled together the results of five randomized trials [4]. Schedules used in the experimental arms of the five trials were different in terms of weekly dose (with docetaxel dose ranging from 33 to 40 mg/m²) and in terms of number of weeks of administration; however, these differences in drug administration were not substantial, and the pooled results (no significant difference in survival between the every 3-week and the weekly schedule) were easily interpreted and commented upon. However, what if the clinical heterogeneity among the trials pooled together in a meta-analysis is substantially higher? If a meta-analysis demonstrates the efficacy of a treatment strategy, but the trials included are different in terms of drugs (or combinations of drugs) used, which regimen should be recommended in clinical practice? Does the absence of statistical heterogeneity in the results of the meta-analysis allow the conclusion that all the treatments used in the different trials are acceptable alternatives for clinical practice? Of course, interpretation of results should be prudent. If the drug, A, was tested in a single trial (without proving a statistically significant benefit versus the control arm) and a subsequent meta-analysis demonstrates a significant benefit, substantially driven by other trials conducted with different drugs (for instance, B and C), without statistical heterogeneity, are we able to claim that the same benefit is associated with each of the three drugs A, B and C? This would be a dangerous conclusion. The absence of statistical heterogeneity allows us to conclude that the results observed in the different trials are compatible with the hypothesis that the different drugs have the same efficacy, but this does not mean that the same efficacy is proven.

In recent years, many trials have been conducted to test the efficacy of maintenance treatment for patients without progression at the completion of first-line chemotherapy for advanced NSCLC [10]. The various trials are different, not only for the drugs administered as a first-line treatment, but also for the maintenance strategy adopted: ‘switch maintenance’ (i.e., the use of different drugs from those previously administered as first-line) or ‘continuation maintenance’ (i.e., the prolonged administration of one or more drugs already used as part of first-line treatment). Several meta-analyses have recently been conducted on this ‘hot topic’ [11–13]. In my opinion, however, when put together in a meta-analysis, the results are quite difficult to interpret and to translate into clinical decisions. For instance, in the meta-analysis published by Behera et al. with single-agent maintenance therapy, 12 studies (for a total of 13 comparisons) were included, considering randomized trials [13]. In detail, five comparisons addressed the issue of continuation maintenance, and eight comparisons addressed the issue of switch maintenance. Among the former, continuation maintenance was based on the administration of different drugs (gemcitabine or paclitaxel or pemetrexed [Eli Lilly, USA]). Similarly, among the latter, switch maintenance was based on the administration of different cytotoxic or targeted agents (erlotinib [Roche, Switzerland],
The risk of indirect comparisons

When several randomized clinical trials are available that compare two or more different experimental treatments versus the same control in the absence of a direct, head-to-head comparison, one might ask for a meta-analysis providing indirect comparison of the efficacy of the different experimental arms, thus helping clinical decisions. This is exactly the case of EGF receptor (EGFR) tyrosine kinase inhibitors in the treatment of patients with EGFR-mutation positive advanced NSCLC. Currently, we have the results of trials comparing gefitinib versus platinum-based chemotherapy [15,16], erlotinib versus platinum-based chemotherapy [17,18] and afatinib (Boehringer-Ingelheim, Germany) versus platinum-based chemotherapy [19]. Each inhibitor has been convincingly proven superior to chemotherapy in terms of progression-free survival, but we still lack randomized trials directly comparing one inhibitor versus the other. Can a meta-analysis help clinicians in deciding which drug is better? In 2011, Bria et al., published a meta-analysis based on summary data, and they showed that EGFR tyrosine kinase inhibitors (at that time, gefitinib or erlotinib) significantly prolonged progression-free survival over standard first-line chemotherapy (HR: 0.45; 95% CI: 0.36–0.58; p < 0.0001) [20]. A significant heterogeneity among trials in progression-free survival was described, and a significant interaction between that outcome and the specific drug used (erlotinib vs gefitinib) was found (p < 0.0001). However, authors prudently did not emphasize this finding in the discussion of the paper because solid conclusions on the relative efficacy of one drug versus the other based on indirect comparisons are potentially misleading.

As a general rule, in the absence of direct comparison between competing treatments, one might perform adjusted indirect comparisons by using data from published meta-analyses of randomized trials [21]. The adjusted indirect comparison may provide useful or supplementary information on the relative efficacy of competing interventions. However, adjusted indirect comparisons usually (but not always) agree with the results of head-to-head randomized trials, and the validity of the adjusted indirect comparisons is strictly dependent on appropriate search and selection of relevant trials, on internal validity and on similarity (both clinical and methodological) of the included trials [22]. In my opinion, adjusted indirect comparisons can be useful, but their results should be interpreted with caution, and they should not substitute head-to-head comparison between different options.

Too much’ power: the difference between statistically significant & clinically relevant

Of course, we can think to perform a useful meta-analysis when at least two clinical trials are available on
the topic of interest. Consequently, the meta-analysis will be based on a higher number of patients than each single trial, and this larger sample size will allow us to obtain higher statistical power to detect differences between treatments. Several authors have correctly emphasized that this can be considered an advantage of prospective meta-analyses, and that, for instance, while each single trial may use surrogate end points (tumor response or progression-free survival) as the main end point, due to larger expected benefit allowing smaller sample size, a meta-analysis of several of these trials might have enough statistical power to allow a meaningful study of the strongest end point, overall survival [23]. However, demonstration of a statistically significant benefit in overall survival does not necessarily mean that this advantage is also clinically relevant. With a great number of patients, also a very small, even negligible difference in overall survival between treatment arms can be statistically significant, but it could be considered not clinically relevant because of the small magnitude. This could happen in a randomized clinical trial, but the risk with meta-analyses is even higher, given the higher number of patients and the higher statistical power that could identify, as statistically significant, even small differences between treatment arms. When interpreting the results of a meta-analysis, and discussing a positive result, the reader should always pay attention to the magnitude of benefit demonstrated by the experimental treatment.

For example, bevacizumab (Roche, Switzerland) is approved by regulatory agencies for the treatment of patients with advanced NSCLC, following the results of two Phase III randomized trials [24,25]. In the first of these trials, the addition of bevacizumab was associated with a significant prolongation of overall survival [24], which, on the contrary, was not demonstrated in the second trial [25,26]. A meta-analysis based on summary data of randomized trials comparing first-line bevacizumab plus platinum-based chemotherapy with chemotherapy alone for inoperable locally advanced, recurrent or metastatic NSCLC was published in 2013 [27]. The meta-analyses pooled together the data of four randomized Phase II and Phase III trials. The authors conclude that bevacizumab significantly prolonged overall survival and progression-free survival when added to first-line platinum-based chemotherapy in patients with advanced NSCLC. No doubt that this conclusion is formally correct, because the difference in overall survival (HR: 0.90; 95% CI: 0.81–0.99; p = 0.03) and progression-free survival (HR: 0.72; 95% CI: 0.66–0.79; p < 0.001) is statistically significant. However, their data show a 10% reduction in the risk of death (corresponding to HR of 0.90) and a 4% increase in the proportion of patients alive 1 year after randomization from 51 to 55%. Would this magnitude of benefit be considered worthwhile when planning a clinical trial? Probably not. For instance, the first randomized Phase III trial was planned to detect a HR of 0.80 (20% reduction in the risk of death), with the addition of bevacizumab to carboplatin plus paclitaxel, and the observed HR was 0.79 (with a 2-month prolongation of median overall survival) [24]. Interestingly, although the absence of individual patient data did not allow a deep exploration of predictive factors of efficacy, authors of the meta-analysis, thanks to the availability of subgroup analysis in three of the four pooled trials, try to explore interaction between patients characteristics and treatment effect. According to these subgroup analyses, bevacizumab showed a significantly greater treatment effect on overall survival in patients with adenocarcinoma than with other histological types (i.e., large cell and other types; p = 0.03).

Another example is the meta-analysis of four trials comparing the efficacy and toxicities of chemotherapy plus cetuximab (Merck Serono, Germany) versus chemotherapy alone in patients with previously untreated advanced NSCLC. When summary data were used for the meta-analysis, the pooled HR for overall survival in favor of addition of cetuximab to chemotherapy was 0.87 (95% CI: 0.79–0.96; p = 0.004) [28]. A meta-analysis based on the individual patient data of the same four trials gave very similar results, with a significant advantage in overall survival (HR: 0.88, corresponding to a 12% reduction in the risk of death [29]). Is this benefit clinically relevant?

In both cases of bevacizumab and cetuximab, the statistically significant prolongation in overall survival in unselected patients was obtained with the addition (until progression) of another drug to first-line chemotherapy, with all the implications in terms of toxicity and costs. This should always be considered when interpreting the results [30]. From my point of view, this attention to risk–benefit and cost–benefit ratio should always be a priority, and this is true in the interpretation of a meta-analysis as well as each single clinical trial.

Conclusion

When well conducted, meta-analyses are a powerful and useful instrument to produce synthesis of the evidence about a specific topic. However, a meta-analysis cannot overcome many of the limitations of the clinical trials it is based on. Similar to a randomized trial, interpretation of the results of a meta-analysis should take into account all these limitations. In particular, heterogeneity of results among clinical trials should be investigated and discussed, and, after demonstration of a statistically significant advantage, the magnitude of
Clinical Trial Perspective

Di Maio

Well-conducted, meta-analyses are a powerful and useful instrument to produce synthesis of the evidence. According to compared to single trials, meta-analyses are associated with an increased statistical power. After demonstration of a statistically significant advantage favoring the experimental treatment \[31,32\]. This change is evident, for instance, if we consider the trials demonstrating the efficacy of EGFR tyrosine kinase inhibitors in the treatment of EGFR-mutation positive advanced NSCLC. In all the randomized Phase III trials comparing gefitinib \[15,16\] or erlotinib \[17,18\] versus platinum-based chemotherapy published to date, a statistically significant and clinically relevant benefit was demonstrated enrolling about 200 patients or fewer, which is far fewer than the sample size that is usually applied in the trials comparing different chemotherapy schedules. This will probably be the same with the vast majority of targeted agents that are in clinical development, because it is reasonable that, when predictive factors are well known before the conduction of the trial, a large advantage can be anticipated and a small sample size should be enough to demonstrate benefit. Furthermore, evidence will often be limited to the results obtained in the representative trial and there will be a lower amount of data to meta-analyze. In addition, regulatory agencies will probably base the majority of their decisions regarding drug approval on the results of a few number of small trials, and meta-analysis will probably have a small role in this process. In this scenario, the room for meta-analyses seems to be reduced, because each single trial can produce a clear result in itself and a few clinicians, after considering the data of each trial alone, would wait for further, combined evidence in order to estimate treatment efficacy and change their mind about treatment decisions. After the positive results of the Phase III trial comparing crizotinib (Pfizer, USA) versus chemotherapy as a second-line treatment of patients with advanced NSCLC selected for ALK translocation \[33\], would anyone need another trial on the same topic to be convinced of the efficacy of crizotinib for that molecular subgroup of patients? However, small sample size of each clinical trial will be associated with wider confidence intervals and, when more than one trial is available on the same issue, a meta-analysis can be useful to obtain a more precise estimation of treatment.

Future perspective

What will be the role of meta-analyses in the near future, in the era of targeted agents? When predictive factors are well known and prospectively applied as eligibility criteria at the time of patients’ enrollment, trials with a small sample size can be sufficient to demonstrate a significant advantage favoring the experimental treatment \[31,32\]. This change is evident, for instance, if we consider the trials demonstrating the efficacy of EGFR tyrosine kinase inhibitors in the treatment of EGFR-mutation positive advanced NSCLC. In all the randomized Phase III trials comparing gefitinib \[15,16\] or erlotinib \[17,18\] versus platinum-based chemotherapy published to date, a statistically significant and clinically relevant benefit was demonstrated enrolling about 200 patients or fewer, which is far fewer than the sample size that is usually applied in the trials comparing different chemotherapy schedules. This will probably be the same with the vast majority of targeted agents that are in clinical development, because it is reasonable that, when predictive factors are well known before the conduction of the trial, a large advantage can be anticipated and a small sample size should be enough to demonstrate benefit. Furthermore, evidence will often be limited to the results obtained in the representative trial and there will be a lower amount of data to meta-analyze. In addition, regulatory agencies will probably base the majority of their decisions regarding drug approval on the results of a few number of small trials, and meta-analysis will probably have a small role in this process. In this scenario, the room for meta-analyses seems to be reduced, because each single trial can produce a clear result in itself and a few clinicians, after considering the data of each trial alone, would wait for further, combined evidence in order to estimate treatment efficacy and change their mind about treatment decisions. After the positive results of the Phase III trial comparing crizotinib (Pfizer, USA) versus chemotherapy as a second-line treatment of patients with advanced NSCLC selected for ALK translocation \[33\], would anyone need another trial on the same topic to be convinced of the efficacy of crizotinib for that molecular subgroup of patients? However, small sample size of each clinical trial will be associated with wider confidence intervals and, when more than one trial is available on the same issue, a meta-analysis can be useful to obtain a more precise estimation of treatment.

Financial & competing interests disclosure
The author has no relevant affiliations or financial involvement with any organization or entity with a financial interest in or financial conflict with the subject matter or materials discussed in the manuscript. This includes employment, consultancies, honoraria, stock ownership or options, expert testimony, grants or patents received or pending, or royalties. No writing assistance was utilized in the production of this manuscript.

Executive summary

- Well-conducted, meta-analyses are a powerful and useful instrument to produce synthesis of the evidence. According to an evidence-based approach, meta-analysis results are considered a strong level of evidence in all the most important guidelines about the treatment of advanced non-small-cell lung cancer.

Individual patient data or summary data?
- Compared with meta-analyses based on data extracted from the literature, individual patient data meta-analyses are more difficult and longer to perform, but they allow deeper verification of data quality, allow the use of updated data, permit more powerful subgroup analyses and enhance the cooperation among participating investigators.

The heterogeneity of clinical trials considered in a meta-analysis
- If the clinical heterogeneity among the included studies is high (different inclusion criteria, different treatments), it will be difficult to translate the results of a meta-analysis into clinical decisions.

The risk of indirect comparisons
- A meta-analysis can provide indirect comparison of the efficacy of different experimental arms. Adjusted indirect comparisons can be useful, but their results should be interpreted with caution, and they should not substitute the need for head-to-head comparison between different options.

‘Too much’ power: the difference between statistically significant & clinically relevant
- Compared to single trials, meta-analyses are associated with an increased statistical power. After demonstration of a statistically significant advantage, the magnitude of the observed benefit should always be weighed against risks and costs of treatments.
The role of meta-analysis in defining clinical practice in advanced NSCLC

Clinical Trial Perspective

References

Papers of special note have been highlighted as:

- of interest
- of considerable interest


- A seminal paper describing principles, methods and limits of meta-analyses, with a specific focus on cancer.


- An important, easy-to-read paper, emphasizing the importance of clinical and statistical heterogeneity in conducting and interpreting meta-analyses.


- An interesting review discussing the usefulness of meta-analyses as compared with large scale trials.


Clinical Trial Perspective

Di Maio


33 Shaw AT, Kim DW, Nakagawa K et al. Phase 3 randomized study of crizotinib versus pemetrexed or docetaxel chemotherapy in advanced, ALK-positive NSCLC (PROFILE 1007). Ann. Oncol. 23 (Suppl. 9), ixc1–ixc30 (2012).

Website