ELSEVIER

European Journal of Cancer 41 (2005) 1501–1502

European Journal of Cancer

www.ejconline.com

Editorial Comment

Confirmation of response in cancer clinical trials: A meaningless exercise?

P. Therasse [a,*], E. Eisenhauer [b]

[a] *EORTC Data Center, Av. E. Mounier, 83/11, 1200 Brussels, Belgium*
[b] *NCIC Clinical Trials Group, Kingston, Canada*

Received 15 April 2005; accepted 15 April 2005
Available online 22 June 2005

In this issue of EJC, Perez-Gracia and colleagues question the need for confirmation of response reported in cancer clinical trials. To illustrate this problem, the authors have selected several phase II trials testing the activity of a new anticancer agent in different tumour types. The authors have compared the response rates obtained after the first response evaluation with those obtained after confirmation. The correlation between response rates (assessed with Cohen's Kappa coefficient) was excellent for the pooled analysis of the trials and also within each study. On the basis of the results of their study and in absence of an established scientific rationale to justify confirmation of response the authors concluded that the need to confirm response is questionable and other similar studies should be undertaken to confirm their findings.

As mentioned by the authors, the requirement to confirm the responses observed in cancer clinical trials is referred to in all published response evaluation guidelines and inherited from the principles established to evaluate response in breast cancer by Hayward *et al.* [1] in the late 1970s. These were the early days of modern chemotherapy and tumour response was usually interpreted as having true clinical benefit for the patient, hence a minimum duration of response or disease stabilisation was required to assure the response was of sufficient length to be meaningful. Thus, the rationale for repeat evaluation was not really to identify possible measurement error but rather to document a minimum duration of response.

Twenty years later, the RECIST working group took on the task of developing a revised set of response criteria [2]. Amongst the many topics debated during that process, one was whether there was a need to confirm partial and complete response. Since the early days of response assessment in the 1970s, we have learned that objective tumour response is not a guarantee of true clinical benefit but it clearly is an excellent marker of the potential for anticancer activity. Given the central importance of response as defined over three decades in making appropriate decisions about further development of new agents, the RECIST working group considered carefully whether to eliminate the requirement for confirmation and was reluctant to do so for a variety of reasons as summarised here:

(a) Experience (in the form of blinded response review) has demonstrated that there is a great deal of variability in measuring lesion size – because of the imprecision inherent to the methodology used (technical and human factors confounded) – thus having a response "confirmed" by a second set of measurements helps assure it was not simply error that led to the response designation.
(b) The response rate in phase II trials is not only used as an indicator of biological anticancer activity, but is also often critical for the future development of the drug. In many instances, drug marketing authorisations have been granted on the basis of promising response rates from phase II trials [3,4], thus reasonable efforts to assure its veracity by confirmation seem justifiable.
(c) The sample size of phase II trials is small and consequently even very few errors could have a substantial impact on the final outcome (in terms of declared response rate) of a trial.
(d) The overall response rate is generally comprised of partial responses sometimes together with a few complete responses. Partial responses are those at

* Corresponding author. Tel.: +32 2 774 16 14; fax: +32 2 772 61 97.
E-mail address: patrick.therasse@eortc.be (P. Therasse).

0959-8049/$ - see front matter
doi:10.1016/j.ejca.2005.04.021

greatest risk of erroneous designation because of measurement errors or small physiological tumour volume variations that could impact on the overall assessment. Thus, confirmation serves to reduce this source of error.

(e) Response duration is not reported consistently in clinical trials – external review of responses is not systematic: confirming responses within those trials that do not have external review is an added measure of certainty that the responses were not based on flawed technique or measurement.

(f) It is worth noting that the RECIST working group indicated that for phase III randomised trials, where response rate is not the primary endpoint, the justifications mentioned above are much less relevant hence confirmation of response not mandated.

What can we learn from the data presented by Perez-Gracia and colleagues? Certainly the authors should be commended for undertaking such a study. Any attempt to supply evidence that can lead to simplification of response evaluation in clinical trials is important, since the scientific community is eagerly awaiting improvements in methodology. The approach used by the authors to try to demonstrate that response confirmation does not add value to the overall response evaluation process was useful but probably incomplete. Although the kappa coefficient is very good for the pooled analysis of trials and also within each study, this is not surprising: a confirmed response must be preceded by an unconfirmed response by definition, so correlation is expected. However, the authors did not clearly discuss the possible impact of the small but real differences between response rates (confirmed versus unconfirmed) on the final interpretation of each study – would the trials have been interpreted differently and would different development plans have been made? Phase II trials with objective response as the primary endpoint are usually set-up so that a certain threshold in response rate is defined in the protocol to determine the sample size and to have a decision of go or no go at the end of the study. In the present study, the overall response rates rages from 9% to 36% and small variations in response rate could theoretically impact substantially on the final interpretation of the individual trials.

It is also disappointing to see that out of 16 patients who did not have response confirmed, 15 patients did so simply because they were not re-evaluated. Thus, it can be argued that this is not really a failure to confirm in the traditional sense, but failure to evaluate. This represents not less than 14% of the responders.

This being said, the arguments put forward by the authors to abandon the requirement for confirmation of response are worth considering. The complexity and the costs of confirmation, when the actual impact on response rates is not high, argue against the practice. In fact it is probable, as pointed out, that response rate variations between studies of the same agent in the same disease are greater than the impact within a study of eliminating the requirement for confirmation. Further, it is now commonplace to undertake independent review of responses in phase II trials, decreasing the value of confirmation to reduce measurement error (errors will be identified by the independent reviewer(s)).

However, it is worth noting that for trials where response is the primary endpoint, patients are usually carefully followed until progression. This means that repeated examinations are being performed in any case to determine response duration or to track progression objectively. Thus, the ''real'' gain in convenience of eliminating the confirmation requirement might be relatively modest.

The study by Perez-Gracia and colleagues is interesting and stimulates once again the debate about how complex or simple response criteria should be. Further, this raises the interesting question of how changing the definition, and thus the likelihood of declaring response, could (or should) impact on the interpretation of trials that have a primary response endpoint. Before changing response criteria to eliminate confirmation, more data from other trials and agents should be evaluated. This will enable researchers to understand the magnitude of the effect on response rates in the event that confirmation is no longer required and whether such a change will also require modifying the thresholds for decision-making when tumour response is the primary endpoint of the study.

Conflict of interest statement

None declared.

References

1. Hayward JL, Carbone PP, Heuson JC, et al. Assessment of response to therapy in advanced breast cancer: a project of the programme on clinical oncology of the international union against cancer, Geneva, Switzerland. *Cancer* 1977, **39**(3), 1289–1294.
2. Therasse P, Arbuck SG, Eisenhauer EA, et al. New guidelines to evaluate the response to treatment in solid tumours. European Organization for Research and Treatment of Cancer, National Cancer Institute of the United States, National Cancer Institute of Canada. *J Natl Cancer Inst* 2000, **92**(3), 205–216.
3. Pignatti F, Aronsson B, Vamvakas S, et al. Clinical trials for registration in the European Union: the EMEA 5-year experience in oncology. *Crit Rev Oncol Hematol* 2002, **42**(2), 123–135.
4. Dagher R, Johnson J, Williams G, et al. Accelerated approval of oncology products: a decade of experience. *J Natl Cancer Inst* 2004, **96**(20), 1500–1509.