Topotecan versus paclitaxel in second-line therapy: a lost opportunity for quality-of-life evaluation?

Ten Bokkel Huinink et al., on behalf of the International Topotecan Group, have recently published a paper on the long-term outcomes of 226 patients with advanced epithelial ovarian cancer who failed one prior platinum-based regimen and were enrolled in a multicentre randomised phase III study comparing topotecan versus paclitaxel as second-line therapy [1].

This study is an update of a study whose main results were previously published in 1997 and 2001 [2, 3]. In the first paper, activity (response rate, duration of response, time to progression), safety (quantitative and qualitative data by common toxicity criteria) and efficacy (survival) appeared to be statistically equivalent (i.e. no significant statistical differences were observed at the usual 0.05 P value level) [2]. The second paper reported, together with an analysis of the non-cross resistance of the two drugs, an update of (progression-free and overall) survival that confirmed the substantial comparability of topotecan when compared with paclitaxel [3].

In the present paper, a further update of survival is presented using follow-up data that have now been collected for more than 4 years [1]. In addition, quality-of-life data are also reported. Overall, topotecan and paclitaxel continue to show comparable efficacy as the differences observed in terms of median survival between the two arms (63 versus 53 weeks) did not reach statistical significance (P = 0.44). As to the quality of life, ten Bokkel Huinink and colleagues reported that the findings from the EORTC QOL-C30 questionnaire were ‘similar’ for the two groups, although no quantitative data were shown. These findings, together with supportive data regarding haematological and non-haematological toxicity, are the basis of their conclusions on the topotecan value: ‘...the gain in survival was not achieved at the expense of quality of life, as measured by the EORTC QOL-C30 questionnaire, which showed quality of life to be consistent throughout therapy’ [1].

We believe that this statement is an over-interpretation of the evidence shown in the paper and, according to our understanding, it is not actually based on solid findings. As a matter of fact, what ten Bokkel Huinink et al. have presented here is not a formal quality-of-life assessment but an evaluation of patient-reported outcomes based on simple, although standardised, cancer-generic symptom scores, the less reliable section of the EORTC QOL-C30 questionnaire [4]. The fact that changes in symptom scores from baseline to the end of best response in the evaluable patients (75 and 85% of patients in the two arms) were not statistically different does not mean that a difference in quality of life does not exist, but simply means that with this approach and with these measures, a statistically significant difference was not observed. The interpretation of ten Bokkel Huinink et al. (i.e. no statistically significant differences in self-reported symptoms = no differences in quality of life = gain in survival for topotecan not hampered by quality-of-life worsening) is just an opinion that is not supported by data, as alternative explanations are possible.

To support our point, we would like to raise three issues. The probability that a difference will be found when there is one relies on several factors that usually are prospectively taken into account with the study design, including the size of the sample (that is estimated a priori), and the measurement precision: in general, with larger differences (effect size) and better measures (reliability), differences will be detected more easily. In the present paper, with a sample size estimated to detect a difference between groups in terms of response rate or survival, it is not clear (because it is not explicitly reported or discussed) what was the power of the study to detect differences in self-reported symptoms using measures that have a relatively poor reliability [4].

In addition, information about how quality-of-life assessment has been carried out has not been given either in the previous reports or in the present paper, e.g. hypothesis testing, rationale to use given measures, the mode and timing of the questionnaire administration, handling of missing data, etc.

Finally, it is quite unusual that the measures actually used were those from the EORTC QOL-C30 questionnaire pertaining to the symptoms and not the more valid and reliable multi-item scales exploring more sensitive and relevant domains that have, a priori, a higher probability to detect a difference in quality of life, if present.

Quality-of-life assessment is a complex task, either from the conceptual or pragmatic point of view and, although there are recommendations and standards published to assure quality in this field [5], previous reviews on randomised controlled trials including quality-of-life evaluation in different cancer sites have shown, overall, a number of methodological shortcomings [6] that hamper the validity and interpretation of results. This setting was indeed a very good opportunity to check whether, in the presence of equivalence of activity and efficacy and with a different pattern in toxicity, quality-of-life...

© 2004 European Society for Medical Oncology
evaluation would give additional and valuable information for clinical decision making. The paper by ten Bokkel Huink et al. is another example of a lost opportunity to corroborate good quality-of-life data with opinions not fully supported by traditional clinical outcomes.

G. Apolone* & A. Buda

Department of Oncology, Istituto di Ricerche Farmacologiche ‘Mario Negri’, Via Eritrea 62, 20157 Milan, Italy
(*E-mail: apolone@marionegri.it)

References


doi:10.1093/annonc/mdh357


Apolone and Buda have provided very thoughtful observations regarding our statements on quality of life (QoL) in our update on the long-term outcomes in patients with advanced epithelial ovarian cancer. They rightly point out several issues that are endemic to most studies when addressing QoL evaluations, such as power of statistical detection, analysis methodology issues, and proper use of validated instruments. While Apolone and Buda may be justified in challenging the statistical merit of our statement, we stand by the clinical merit of the statement.

Quality-of-life and symptom measurements provide useful qualitative assessments of patient benefits which are complementary to safety and efficacy evaluations. The Huink trial in question was the first in its indication—large, multinational, randomized trial in patients with relapsed ovarian cancer. Before the commencement of this study, there were no prospectively-defined randomized trials in this relapsed setting.

With respect to QoL, the trial design was state-of-the-art during its period of conduct, 1993–1994, as it included the relatively new QoL parameters as captured by the EORTC QLQ C30 questionnaire. Although not stated in the original publication, both individual QoL subscales and symptom questions were formally evaluated as secondary endpoints.

The importance and demand for psychometrically validated quality-of-life assessments have progressed over the past 10 years. Studies, like the Huink study, offered evidence that quality of life and well-being data provide useful insight into patient benefit beyond safety and efficacy alone. Apolone and Buda correctly point out that today, the most accepted analyses of the EORTC QLQ C30 data focus on individual subscales. At the time of the original study reporting, however, symptoms were of the greatest interest and acceptance to clinical oncologists. Given the complicated nature of the QoL evaluations as highlighted by Apolone and Buda, symptoms are still the focus for many, if not most, clinical oncologists even to this day. In addition, Apolone and Buda correctly reiterate the importance of properly designed and developed quality-of-life instruments and knowledge of their reliability and validity, thankfully not issues with the EORTC survey.

While retrospective analysis of this original data with newly established standards may be of exploratory interest, such an analysis was beyond the intent of this long-term efficacy update. While we understand and agree with the authors’ criticism about the weakness of the statistical establishment of no compromises in QoL, we are confident from the data in the Huink study that underlying differences would not be so great as to produce clinical concern in the patient population at hand.

G. A. Ross* & J. A. Kettels

GlaxoSmithKline, New Frontiers Science Park, Third Avenue, Harlow CM19 5AW, Essex, UK (*E-mail: graham.a.ross@gsk.com)

doi:10.1093/annonc/mdh358

Cyclooxygenase-2 inhibitor and cisplatin combination as a radiosensitizer in the treatment of head and neck cancer patients

We read with interest the study by Nix et al. [1], in which they attempted to investigate the possible relationship between...