published so far focusing on the alternative schedule of sunitinib administered 2 weeks on/1 week off (2/1) to 208 patients who had developed significant adverse events with the standard schedule of 4 weeks followed by a pause of 2 weeks (4/2).

As the authors pointed out, safety was the primary objective of this analysis, and the 2/1 schedule demonstrated a reduced incidence of overall grade 3–4 toxicities (8.2%) compared with the first period of 4/1 schedule administered to the same patients (45.7%) as well as to the overall incidence of grade 3–4 toxicities registered in a French control group (29.4%). Apart from the limitations already raised by the authors and the comments previously expressed by Iacovelli et al. [2], we would add the following considerations:

(i) The 2/1 schedule was applied to patients who had already received a median of 4.3 months of the standard 4/2 schedule, who had likely undergone the first radiological restaging. The same happened in other similar studies with the 2/1 schedule [3]. This means that refractory patients (21% in registrative studies [4]) have not been switched to the 2/1 schedule, as well as those who had unacceptable toxicities or dropped out for treatment-unrelated reasons. This fact clearly represents a strong selection bias and may explain, at least in part, the striking improved median progression-free (PFS) and overall survival (OS) of the 2/1 cohort. It would be useful to know the reasons for which 41 patients of the total 4/2 group were not switched to the 2/1 schedule, and how fared those patients who were able to continue the 4/2 sunitinib schedule afterward.

(ii) The second cohort of 41 patients starting 2/1 sunitinib on the basis of ‘non-optimal general conditions’ according to the local treating oncologists shows only a moderate and insignificant increase of poor risk features compared with the first cohort, such as median age, ECOG PS of 2 or more, bone or CNS involvement. Indeed, they had a PFS of 10.4 months and a 36-month survival rate of 39.4%, very close to the 9.7 months and 39.5% of the unselected French cohort. In fact, patients with unfavorable clinical conditions may be expected to be given temsirolimus or alternative tyrosine kinase inhibitors such as sorafenib and more recently pazopanib, which are perceived as less toxic than sunitinib. This second group does not appear to us a significantly disadvantaged cohort, and the higher toxicity of 2/1 sunitinib administered since the first cycle (26.8% versus 8.2%) may reflect the absence of the selection bias applied to the first cohort. As the authors correctly point out in the discussion, there are still not enough data to justify the start of 2/1 sunitinib upfront.

(iii) Eighty-two of 188 patients starting sunitinib at 50 mg/daily were not able to maintain the full dose of sunitinib even after switching (43.6%), as a demonstration that the 2/1 schedule may still frequently require dose reductions to ameliorate tolerability.

(iv) Registration of toxicity by physicians may not intercept the actual incidence and severity of treatment toxicities, as we have recently learnt from the PISCES trial in which patient-reported outcomes were assessed and compared with toxicities of sunitinib and pazopanib reported by doctors [5]. This may be particularly true when the physician claims that a modification of treatment will improve tolerability (observer bias).

In conclusion, we compliment the authors for carrying out and reporting strengths and limits of the Rainbow analysis, which adds further information on feasibility and tolerability of the 2/1 schedule of sunitinib administered to patients who cannot tolerate the standard 4/1 schedule [3]. We agree with them that a randomized, controlled trial is warranted, for which both the standard 4/2 sunitinib or the continuous pazopanib could be suitable control arms.

U. Basso* & M. Maruzzo

Medical Oncology 1 Unit, Department of Clinical and Experimental Oncology, Istituto Oncologico Veneto IOV–IRCCS, Padova, Italy

(*E-mail: umberto.basso@ioveneto.it)

disclosure

The authors have declared no conflicts of interest.

references


doi: 10.1093/annonc/mdv493
Published online 23 October 2015

Reply to the letter to the editor

‘The ESMO Magnitude of Clinical Benefit Scaling Tool: from theory to practice’ by Hartmann and the letter ‘Comment on ESMO Magnitude of Clinical Benefit Scale’ by Muhonen et al.

Muhonen et al. [1] propose that the point measure of the hazard ratio (HR) may be a more robust measure of comparing the relative efficacy of treatments than the low end of the confidence interval (CI). We disagree. Members of the task force had substantial discussion regarding the use of HR data in the ESMO-MCBS as a measure of relative efficacy alongside a measure of absolute benefit. The approach that we have adopted is based on a nuanced understanding of the relationship between the point estimate and CIs. Simulations supported this approach and the application of the scale on existing trials has so far been in concordance with the expert opinion in the oncology field.
HR is typically described in comparative research studies using a point estimate as well as a 95% CI. The point estimate represents a crude one-number summary of data, and the 95% CI describes a range of reasonable uncertainty between the optimistic estimate (at the lower end of the CI) and a most skeptical estimate (at the upper limit of the CI).

The one-number summary of data, the point estimate, is not appropriate for threshold evaluations because it invariably has the potential for both under- or overestimation of effect depending on the threshold being evaluated. Consequently, threshold evaluations are best performed taking into consideration the 95% CI. Thus, by convention, hypothesis testing of superiority in a comparative trial (i.e. HR <1) is determined by the upper end of the 95% CI. This precautionary approach essentially ensures that if one cannot be reasonably certain that there is benefit, it should not be ascribed.

Crediting the magnitude of benefit is also a threshold evaluation, but at the other end of the scale. The choice of the low limit of the 95% CI essentially says that if the benefit may be as great as this, it ought to be accredited so as to preclude the potential for under-evaluating the benefit by using the point estimate.

With regard to concern that this approach favors small studies, it is noted that only studies leading to significantly statistical results are evaluated by the MCBS scale. Specifying a lower limit while the upper limit needs to be <1 to reach a significant result, moves the wider CI to be centered around a lower value (to be precise the center of the interval in the log scale). Thus, the observed relative benefit needs to be considerably large for the rule to be satisfied. Indeed, in an appropriately powered study for such a benefit (power 80%–90%), the lower CI limit rule outperforms the point estimate threshold criterion. For the particular example referred to by Muhonen et al., with an observed HR value of 0.76, the 95% CI from 0.65 to 0.89 corresponds to a power of 95% for a design HR of the same magnitude (already an excessive power).

Dr Hartmann [2] suggests that the credibility of the ESMO-MCBS would be augmented by external validation, comparing its results with those of health technology assessments carried out in Europe. This form of correlative validation has been published evaluations of these same agents using the ESMO-MCBS and that this methodology is still in a state of refinement and that this process may be augmented by more widespread adoption of uniform processes that are now being encouraged by the framework of the European Network for Health Technology Assessment (EUnetHTA). The methodology incorporated into the ESMO-MCBS is remarkably consistent with EUnetHTA guidelines for outcome measures [3], surrogate outcomes [4], health-related quality of life [5] and the application of these outcomes in relative effectiveness evaluations [6].

It is important to remember, however, that HTA processes, such as the UK’s NICE and the Israeli basket of services, make value judgments based on countries’ willingness-to-pay thresholds which differ from member state to member state. Thus, the HTA system diverges from ESMO-MCBS in that the former is intimately linked with willingness-to-pay thresholds and pricing policy whereas the ESMO-MCBS is focused solely on the level of clinical benefit.

ESMO looks forward to feedback from HTA bodies and the constructive dialog that may be generated from the analysis of diverging assessments if they arise.

The ESMO-MCBS is designed to be an evolving tool. Suggested improvements are always welcomed and will be adopted if successful in the repeated rigorous testing performed on the current rules.


1Department of Medical Oncology, Cancer Pain and Palliative Medicine Service, Shaare Zedek Medical Center, Jerusalem, Israel; 2Kings Health Partners Integrated Cancer Centre, King’s College London, Institute of Cancer Policy, London, UK; 3Division of Public Health, Department of Nursing, Laboratory of Biostatistics, University of Athens, Athens, Greece; 4Department of Medical Oncology, Netherlands Cancer Institute, Amsterdam, The Netherlands; 5Department of Medical Oncology, Ospedale San Martino, Genova, Italy; 6Division of Oncology, Medical University Vienna, Vienna, Austria; 7Department of Medical Oncology, University Medical Center Groningen, Groningen, The Netherlands; 8Department of Medicine, Université Libre de Bruxelles, Jules Bordet Institute, Brussels, Belgium

(*E-mail: chernyn@netvision.net.il)

disclosure

The authors have declared the following: UD: lecture fees—Amgen. EGEV: research grants from Roche/Genentech, Amgen, Novartis are made available to my hospital the UMCG in Groningen. MJP: board member—PharmaMar; Consultant honoraria)—Amgen, Astellas, AstraZeneca, Eli Lilly, GSK, Invivis, MSD, Novartis, Pfizer, Roche/Genentech, Sanofi Aventis, Symphogen, Synthon, Verastem; Research grants to Institute: most companies; Speakers bureau/stock ownership: none. AS: advisory Board and Symposia Satellite with; Amgen, Bayer, Celgene, Merck, Roche and Sanofi. CZ: advisory Boards—Roche, Celgene, Bristol Myers-Squibb; lecture fees—Amgen, Bristol Myers-Squibb.

references

How much evidence isn’t in evidence-based guidelines?

We were pleased to see that the often neglected, poorly assessed and sub-optimally managed symptom of breathlessness in cancer has been now addressed by a European-wide guideline [1]. However, we have significant concerns about the basis of some of the recommendations which are not consistent with available best evidence.

choice of presented evidence

The guidelines do not claim to be based on a systematic review and the methodology is unclear. It is therefore difficult to see the rationale for the choice of source literature. This is a particular concern for the recommendations involving the use of drugs outside their license where drugs and regimens are recommended without a corresponding hierarchy of evidence. For example, similar weight has been afforded to findings from phase II studies and observational cohorts as for adequately powered phase III randomised, controlled trials (RCTs). Further, limitations are encountered when the generalisability of findings is not critiqued in the clinical and research contexts of the original work [2].

Specifically, the recommended dose and titration regimen for opioids are not clearly supported by the referenced papers. Immediate release dosing is recommended yet the only phase III effectiveness RCT against placebo which included people with cancer deliberately chose to use a sustained release morphine formulation [3]. Furthermore, a recent systematic literature review and meta-analysis of opioids for breathlessness in people with chronic obstructive pulmonary disease (COPD) demonstrated greater relief from breathlessness in steady-state trials compared with trials using ‘as required medication’ [4]. Why therefore, have sustained release preparations not been mentioned at all in either titration or steady state? Although most trials have been conducted in people with non-cancer conditions, a pooled data analysis from trials which included participants with a variety of conditions (COPD, cancer, heart failure) showed that the underlying disease did not predict response to opioids suggesting that this body of evidence is useful for people with cancer [5].

With regard to the use of non-pharmacological interventions, it was stated that there is little evidence to support their use other than clinical consensus. In distinct contrast to this statement, the Cochrane review [6] cited to support this assertion shows evidence supporting a variety of interventions. Even if the guideline committee wish to limit source evidence to research in cancer patients, the seminal phase III RCT of a complex breathlessness intervention in people with cancer published in 1999 is omitted [7]. This landmark trial formed the basis for subsequent breathlessness intervention services, including the service evaluated in some of the referenced papers. We must avoid the systematic bias against implementation of non-pharmacological interventions in the relief of breathlessness given their pivotal role in patient-centred care. It is challenging to ensure the uptake of these evidence-based interventions if guidelines fail to recognise their contribution.

M. J. Johnson1* & D. C. Currow2

1 Palliative Medicine, Hull York Medical School, University of Hull, Hull, UK;
2 Discipline, Palliative and Supportive Services, Flinders University, Adelaide, Australia
(*E-mail: miriam.johnson@hyms.ac.uk)

disclosure

The authors have declared no conflicts of interest.

references


doi: 10.1093/annonc/mdv502
Published online 30 October 2015