ESMO GOOD SCIENCE BETTER MEDICINE BEST PRACTICE

ESMO OPEN CANCER HORIZONS

REVIEW

11 ESMO 2021 breakthroughs: practicing oncologist's perceptions on data presentation

H. K. van Halteren[1*], A. Tan[2], B. Pellegrino[3,4], B. Brasiuniene[5], J. Bennouna[6], A. J. Cunquero-Tomás[7] & M. Strijbos[8], on behalf of the ESMO Practicing Oncologists Working Group

[1]Department of Medical Oncology, Adrz Hospital, Goes, The Netherlands; [2]Department of Medical Oncology, Waikato Hospital, Hamilton, New Zealand; [3]Medical Oncology and Breast Unit, University Hospital of Parma, Parma; [4]Department of Medical Oncology, University of Parma, Parma, Italy; [5]Department of Medical Oncology, National Cancer Institute of Lithuania, Faculty of Medicine, Vilnius University, Vilnius, Lithuania; [6]Department of Medical Oncology, Hospital Foch, Suresnes, France; [7]Department of Medical Oncology, General University Hospital of Valencia, Valencia, Spain; [8]Department of Medical Oncology, GZA Hospitals, Wilrijk, Belgium

Available online xxx

Background: The European Society for Medical Oncology (ESMO) 2021 conference provided a high number of randomized phase III trial reports, many of which were claimed to be practice changing. Given the short time available for conference presentations, results and conclusions tend to have greatest priority with less time remaining for study background and study methodology.

Purpose: On behalf of the ESMO Practicing Oncologists Working Group, 11 potentially practice-changing reports were selected and screened for three main questions: (i) Did the investigators provide sufficient details with regard to Patients and Methods to make the results comprehensible? (ii) Were there any reasons to consider bias? (iii) To which extent did the results presented translate to clinical benefit?

Results: In 2 out of 11 trials, the study design presented differed considerably from the study design described at ClinicalTrials.gov. Allocation concealment was not carried out in 6 out of 11 trials. In none of the trials reporting progression-free survival was informative censoring considered an issue. In none of the trials reporting overall survival was desirable crossover considered an issue. Defined trial outcome measures depicted at ClinicalTrials.gov, which could boost or weaken the ESMO-Magnitude of Clinical Benefit Scale score, were often lacking in the presentation. Study success was claimed in a heterogeneous manner, which was often not clearly linked to overall clinical benefit.

Conclusion: ESMO conference presentations can inform the scientific community and catalyze further research but cannot replace the full papers in peer-reviewed journals, which are needed to estimate the thoroughness of the results, the overall impact on clinical benefit and the consequences for future treatment guidelines.

Key words: phase III, solid cancers, curative treatment, palliative treatment, methodology, clinical benefit

INTRODUCTION

During the European Society for Medical Oncology (ESMO) 2021 conference, the outcome of a great number of phase III studies testing new interventions was reported, many of which could be regarded as potentially practice changing. A conference report usually precedes publication in a peer-reviewed paper. The period in-between is often utilized by key opinion leaders and scientific organizations to prepare the working floor for the awaited new European Medicines Agency/Food and Drug Administration approval. But drawing firm conclusions from reported abstracts has however certain caveats. Firstly, the outcome measures presented may be based on early data. Early reported differences in progression-free survival (PFS) and/or overall survival (OS) between comparator and intervention group could be affected by an imbalance in point censoring. Point censoring occurs when a patient withdraws from the study, is lost to follow-up or does not experience an event within the study duration.[1,2] Secondly, conference abstracts and presentations tend to focus mainly on results and conclusions and materials and methods are generally under-reported, whereas an interested conference attendee should be given the opportunity to follow the line from materials and methods toward results in order to reflect on the conclusions made. Specific issues to address are: (i) Was treatment allocation blinded in order to ameliorate performance bias? Performance bias happens when one group

*Correspondence to: Henk K. van Halteren, Department of Medical Oncology, Adrz Hospital, PO Box 15, 4460 AA Goes, The Netherlands. Tel: 0031613063783

E-mail: hk.vanhalteren@adrz.nl (H. K. van Halteren).

of subjects in an experiment gets more attention from investigators than another group. The difference in care levels may lead to systematic differences between groups, making it difficult or impossible to conclude that a drug or other intervention caused an effect, as opposed to standard level of care.[3] (ii) Is the treatment given in the experimental arm regularly available in later treatment lines or could the treatment given in the control arm be considered potentially beneficial for refractory patients in the experimental arm? In both cases, crossover should be regarded as desirable. Investigators should provide crossover data, because crossover imbalance could affect later outcome measures, such as OS.[4] Thirdly, the presentation of the study outline may not correspond with the original outline, described at ClinicalTrials.gov (methodology changes, not reported outcome measures, *post hoc* subgroup analysis, etc.). Previous research has revealed inconsistencies between abstract and full paper in quite a proportion of studies (median percentage 39%, range 14%-54%).[5] Fourthly, the definition of clinical benefit used by reporting investigators should ideally correspond with the principles of the ESMO-Magnitude of Clinical Benefit Scale (MCBS).[6]

We carried out a snapshot survey of 11 selected ESMO 2021 phase III study presentations, which could be regarded as potentially practice-shaping breakthroughs, in order to find out how the issues mentioned above were dealt with, i.e. (i) Did the investigators provide sufficient details with regard to Patients and Methods to make the results comprehensible? (ii) Were there any reasons to consider bias? (iii) To which extent did the results presented translate to clinical benefit?

METHODS

All authors were requested to go through the complete conference program and to provide their top 10 list of reports. Based on these individual rankings, a general ranking was made, out of which the studies ranking 1-11 were selected for the survey. For each report, we firstly derived study information published on the ClinicalTrials.gov website. Secondly, we scrutinized the abstract and PowerPoint presentation for overall content, but with specific emphasis on:

Was treatment allocation blinded?
Could the comparator arm be regarded as standard?
Should desirable crossover be anticipated and—if yes—were crossover data made available?
Were there any inconsistencies between study design depicted at ClinicalTrials.gov and study design presented?
Were the data made available based on intention to treat (ITT)?
Were point censoring data per group made available in presentations reporting PFS in order to exclude imbalance?
How did the pre-planned outcome measures compare to the reported outcome measures?
What was the clinical benefit in terms of treatment toxicity, quality of life (QoL) and OS?

Table 1. Abbreviations used for outcome measures

Overall survival	OS
Progression-free survival	PFS
Recurrence-free survival	RFS
Disease-free survival	DFS
Distant metastasis-free survival	DMFS
Metastasis-free survival	MFS
Event-free survival	EFS
Pathological complete response	pCR
Quality of life	QoL
Treatment-related adverse events	TRAE
Performance status	PS

The abbreviations of the outcome measures reported are shown in Table 1.

OVERVIEW OF THE 11 ABSTRACT PRESENTATIONS SELECTED

Breast cancer

The Destiny-Breast03 trial is an open-label, randomized, phase III study in which the comparative efficacy of trastuzumab emtansine (TE, comparator arm) versus trastuzumab deruxtecan (TD, experimental arm) was evaluated in the second-line setting of patients with advanced Her2-positive breast cancer, who had progressed after treatment with a taxane and trastuzumab.[7] Due to the absence of allocation concealment, performance bias should be anticipated. The comparator arm could be regarded as a current standard. Two hundred and sixty-three patients were allocated to TE and 261 patients to TD. A total of 60.1% of patients treated with TE and 62.1% of patients treated with TD had also received pertuzumab in first line. PFS was chosen as primary endpoint, and OS as secondary endpoint; QoL was not assessed. PFS and OS analysis was carried out at an ITT basis. PFS was superior for TD compared to TE [hazard ratio (HR) of progression 0.28, 95% confidence interval (CI) 0.22-0.37]. With regard to PFS, censoring data were not shown and an imbalance in point censoring cannot be excluded. Patients treated with TD also achieved an OS benefit (HR of death 0.56, 95% CI 0.36-0.87) compared to patients treated with TE. The number of patients with grade $\geq$3 treatment-related adverse event (TRAE) rate was higher [116 (45.1%) versus 104 (39.8%) patients] in the experimental group. There were no treatment-related deaths in either study arms.

The KEYNOTE-522 trial is a quadruple-blind, placebo-controlled, randomized, phase III study, which evaluates the additive value of pembrolizumab in the neoadjuvant and adjuvant setting of triple-negative early breast cancer.[8] The neoadjuvant chemotherapy schedule consisted of four 3-weekly cycles of carboplatin and paclitaxel, followed by four 3-weekly cycles of doxorubicin/epirubicin + cyclophosphamide. In the experimental arm, four 3-weekly cycles of pembrolizumab were added to the last four chemotherapy cycles and another nine cycles were given after surgery. Pathological complete response (pCR) and event-free survival (EFS) were taken as primary outcome

measures. EFS according to programmed death-ligand 1 (PD-L1) status, OS and QoL [European Organisation for Research and Treatment of Cancer (EORTC) Breast Cancer-Specific QoL Questionnaire (QLQ-BR23) score; EORTC Quality of Life Core 30 Questionnaire (QLQ-C30) score] were secondary outcome measures. A total of 1174 patients were enrolled and randomized at a 2 : 1 ratio to either the experimental ($n = 784$) or the comparator ($n = 390$) arm. The pCR rate was significantly higher in the experimental arm (64.8% versus 56.2%, $P = 0.00055$), but it is unclear why the analysis was carried out with a comparator group of 201 patients and an experimental group of 401 patients and based on the presented data selection bias cannot be excluded. When the patient sample does not accurately reflect the entire treatment group, the association between treatment given and pCR rate could differ between selected patients and not selected patients, which is regarded as selection bias. EFS was estimated for the entire group of 1174 patients and the experimental arm appeared superior (HR of event 0.63, 95% CI 0.48-0.82). Grade $\geq$3 TRAE occurred in 77.1% of patients treated in the experimental arm and in 73.3% of patients treated in the comparator arm. There were three toxic deaths, two in the experimental arm and one toxic death in the comparator arm. OS and QoL outcome data were not reported.

Cancer of the uterine cervix

The KEYNOTE-826 trial is a double-blind, placebo-controlled, randomized, phase III study, which evaluates the benefit of adding pembrolizumab to standard first-line systemic therapy for advanced cancer of the uterine cervix.[9] The treatment regimen assigned to patients in the comparator arm (paclitaxel, cisplatin or carboplatin, placebo, with or without the addition of bevacizumab) could be regarded as a current standard. Both OS and PFS were primary outcome measures. QoL (assessed by means of the EQ-5D-5L Visual Analogue Score, which is a self-reporting scale ranging from 0% to 100%) was a secondary outcome measure. Intended subset analysis and/or stratification was not mentioned in the study outline described on ClinicalTrials.gov. *Post hoc* stratification factors reported were synchronous versus metachronous metastases, PD-L1 status and bevacizumab versus no bevacizumab. Three hundred and nine patients were allocated to the comparator arm and 308 patients to the experimental arm. PFS and OS analysis was carried out at an ITT basis. With regard to PFS, the experimental arm turned out to be superior (HR of progression 0.65, 95% CI 0.53-0.79). The presented PFS data did not make clear whether there was an imbalance in point censoring. OS proved superior in the experimental group as well (HR of death 0.67, 95% CI 0.54-0.84). In the *post hoc* subset analysis, pembrolizumab failed to show a significant survival benefit in the subset of patients, who did not receive bevacizumab, but selection bias with regard to bevacizumab treatment should be anticipated. The majority of patients (251 out of 307 in the experimental arm and 232 out of 309 in the comparator arm) experienced grade $\geq$3 TRAE and there were 14 toxic deaths in each treatment arm. Loss of QoL (defined as a decrease of at least 10 points in the EuroQol-5D-5L Visual Analogue Score over a period of 30 weeks) occurred in 47.7% of patients in the comparator group and 39.5% of patients in the experimental group. Considering the high TRAE rate in both comparator and experimental arms, data with regard to absolute QoL changes over time would be informative.

The EMPOWER-Cervical 1/Gynecologic Oncology Group-3016/ENGOT-cx9 trial is an open-label, phase III trial, which evaluates the efficacy of cemiplimab (an immunoglobulin G4monoclonal antibody to the programmed cell death protein 1 receptor) administered intravenously every 3 weeks compared to investigators' choice chemotherapy in the second-line setting of platinum-refractory advanced cancer of the uterine cervix.[10] Due to the absence of allocation concealment, performance bias should be anticipated. Patients with squamous cell cancer and adenocarcinoma were both eligible. Primary outcome measure was OS. PFS and QoL (EORTC QLQ-C30) were secondary outcome measures. Six hundred and eight patients were randomized at a 1 : 1 ratio. No information was provided with regard to chemotherapy given to patients in the comparator arm; heterogeneity with regard to the agents given could have resulted in heterogeneity in treatment response and treatment toxicity. OS turned out to be higher for patients in the experimental arm (HR of death 0.69, 95% CI 0.56-0.84). Similar figures were seen for 477 patients with squamous cell cancer (HR of death 0.73, 95% CI 0.58-0.91) and 131 patients with adenocarcinoma (HR of death 0.56, 95% CI 0.36-0.85). Grade $\geq$3 TRAEs were seen in 45% of patients treated in the experimental arm and in 53.4% of patients treated in the control arm. There were two toxic deaths in the control arm. QoL was better maintained in patients treated in the experimental arm. PFS data were not reported.

Melanoma

The KEYNOTE-716 trial is a double-blind, placebo-controlled, randomized, phase III study in which patients who had undergone a complete resection of stage II high-risk melanoma (stage IIB and IIC) were randomized to 17 3-weekly cycles of placebo (comparator arm) or pembrolizumab (experimental arm).[11] The comparator arm can be regarded as standard, as adjuvant systemic therapy is currently not recommended for stage II melanoma. Recurrence-free survival (RFS) was chosen as primary outcome measure; distant metastasis-free survival, OS and QoL (assessed by means of the EORTC QLQ-C30 global health status/QoL scale) were chosen as secondary outcome measures. Intended subset analysis and/or stratification was not mentioned in the study outline described on ClinicalTrials.gov.

RFS and OS analysis was carried out at an ITT basis. Four hundred and eighty-seven patients were allocated to the pembrolizumab arm and 489 patients to the placebo arm. RFS appeared to improve after adjuvant pembrolizumab

(HR of recurrence 0.65, 95% CI 0.46-0.92). OS data were not reported. Grade ≥3 TRAE occurred more often in the experimental group (16.1% versus 4.3%). There were no toxic deaths in either study group and there was a trend towards a better sustained QoL in the experimental group compared with the comparator group.

Prostate cancer

The PEACE-1 trial is an open-label, randomized, phase III study in the first-line setting of advanced prostate cancer in which patients were randomized to androgen deprivation therapy (ADT) and docetaxel (D; comparator arm, $n = 296$); ADT, D, abiraterone (Abi) and prednisone (P, first experimental arm, $n = 292$); ADT, D and radiotherapy (R, second experimental arm, $n = 293$); and ADT, D, Abi, P and R (third experimental arm, $n = 292$).[12] Primary outcome measures were PFS and OS. QoL (assessed by the QLQ-C30 questionnaire) was a secondary outcome measure. Because the experimental drugs had already proven benefit in later lines of therapy, crossover should be regarded as desirable. Stratification for performance status and disease extent was specified in the study outline. Pre-planned subset analysis was not mentioned. The reported PFS and OS analysis did not encompass the previously specified study arms. Instead, the four different arms were mixed in order to create a *post hoc* subset of 355 (out of 584) patients treated with abiraterone and a subset of 355 (out of 587) patients, who were not treated with Abi. The methodology used for this selection process, depicted as a 2 × 2 factorial design, was not clarified. While comparing the two groups, Abi appeared to significantly improve PFS and OS, but the full paper should provide more insight into the reshaping process of the study arms, potential imbalance in point censoring and details with regard to crossover after disease progression.

Systemic Therapy in Advancing or Metastatic Prostate Cancer: Evaluation of Drug Efficacy (STAMPEDE) is a multi-arm, multi-stage, open-label, randomized controlled trial platform, which was initiated in 2005.[13] The STAMPEDE consortium assesses novel approaches (experimental arms) for the treatment of men with hormone-naïve prostate cancer who are starting long-term ADT (comparator arm). Due to the absence of allocation concealment, performance bias should be anticipated. In the experimental arm G of STAMPEDE (accrual 2012-2014), the benefit of adding abiraterone and prednisolone to ADT was evaluated, and in the experimental arm J (accrual 2014-2016), the benefit of adding abiraterone, prednisolone and enzalutamide to ADT was evaluated.

With regard to the presented analysis, it was decided to merge the comparator and experimental arms of the J and the G trial and to split the study population into patients with M0 and M1 disease. It could be comprehended that this decision was made to test the additive value of abiraterone in a larger patient set with M0 disease. But by doing so, two consecutive datasets with different follow-up duration were merged and a higher level of point censoring could be anticipated. Patients with M0 disease treated in the merged experimental arms appeared to have an improved metastasis-free survival (HR of metastasis 0.53, 95% CI 0.44-0.66), but at 84 months' follow-up the percentage of the patients censored was considerably higher in the experimental arm (62.4% versus 52.8%). Addition of abiraterone also appeared to result in an improved OS (HR of death 0.66, 95% CI 0.48-0.73). But again, as the survival curves started to diverge at 84 months, imbalance in point censoring started to increase as well (% point censoring at 84/108 months in the merged experimental arms 65%/80% and in the merged control arms 57.5%/65%). This imbalance makes the survival gain claimed less reliable. From a clinical perspective, it is difficult to draw conclusions as well. Testing new regimens compared to ADT over a period spanning >15 years increases the chance of the treatment assigned to the comparator arm (ADT only) to become out of date (substandard). In this report, the presented data do not provide insight into the incremental benefit of adding enzalutamide to abiraterone/prednisolone. In order to answer this question, it would have been more appropriate to assign the comparator group to treatment with ADT, abiraterone and prednisolone, instead of ADT only.

Pleural mesothelioma

The CheckMate 743 trial is an open-label, phase III, randomized study, which compares the efficacy of the experimental nivolumab/ipilimumab combination with standard chemotherapy in the first-line setting of unresectable pleural mesothelioma.[14] Due to the absence of allocation concealment, performance bias should be anticipated. Primary outcome measure was OS, and secondary outcome measures were PFS and OS according to PD-L1 status. QoL was not assessed. Three hundred and three patients were randomized to 2-weekly cycles of nivolumab and ipilimumab for a period of up to 2 years and 302 patients were randomized to six cycles of cisplatin or carboplatin plus pemetrexed. The only pre-planned subset analysis variable mentioned at ClinicalTrials.gov was PD-L1 status. In the ITT analysis, OS was superior in the experimental arm (HR of death 0.73, 95% CI 0.61-0.87). In the pre-planned subset analysis, patients with PD-L1 <1% did not appear to benefit. In the *post hoc* subset analysis, patients with non-epithelioid histology appeared to derive more benefit from checkpoint inhibitor treatment compared to chemotherapy than patients with epithelioid histology. The grade ≥3 TRAE rate was comparable in the experimental (31%) and the comparator study arm (32%). It is unknown if and how often patients in the experimental group crossed over to chemotherapy at the moment of disease progression. Such desirable crossover could have influenced OS analysis.

Ovarian cancer

The Olaparib Maintenance Retreatment in Patients With Epithelial Ovarian Cancer (OREO) trial is a double-blind, placebo-controlled, phase III study, which evaluates the efficacy of olaparib treatment (experimental arm) compared to placebo (comparator arm) in patients with advanced

ovarian cancer, who had progressed following olaparib maintenance and who had achieved a partial or complete response after subsequent platinum-based chemotherapy.[15] Patients were stratified for BRCA mutational status [112 BRCA-mutated (BRCAm) patients and 108 BRCA-wild-type (BRCAwt) patients] and subsequently randomized in a 2 : 1 ratio to olaparib maintenance or placebo. PFS was a primary outcome measure and OS a secondary. QoL was not assessed. With regard to PFS, olaparib maintenance displayed superior efficacy in BRCAm patients (HR of progression 0.57, 95% CI 0.37-0.87) and BRCAwt patients (HR of progression 0.43, 95% CI 0.26-0.71). Based on the high event rate, the PFS data could be regarded as very solid. PFS censoring data were however not shown and an imbalance in informative censoring cannot be excluded. Grade $\geq$3 TRAE occurred in 26 out of 146 patients treated with olaparib versus 5 out of 74 patients in the comparator group.

Colorectal cancer

The AtezoTRIBE trial is an open-label, randomized, phase II study, which evaluates the additive efficacy of atezolizumab if added to first-line systemic therapy for advanced colorectal cancer.[16] The decision to refrain from allocation concealment could have led to performance bias. Patients were randomized at a 1 : 2 ratio between treatment with FOLFOXIRI/bevacizumab (comparator arm) and treatment with FOLFOXIRI/bevacizumab and atezolizumab (experimental arm). Primary outcome measure was PFS. OS and QoL were not defined as outcome measures. PD-L1 status and/or deficient mismatch repair (dMMR) was not included in the stratification criteria. In spite of regimen intensity, patients with a performance status of 2 were eligible. The ITT population encompassed 73 patients treated in the comparator arm and 145 patients treated in the experimental arm. Thirteen out of 218 patients turned out to have a dMMR tumor. PFS appeared to be superior for the experimental arm (HR of progression 0.69, 95% CI 0.56-0.85). The number of patients censored in both study arms was not reported and imbalance in point censoring cannot be excluded. The overall $\geq$3 TRAE rate was not reported, neither whether there were toxic deaths. This information is essential considering the intensity of the treatment regimen used. Several *post hoc* subset analyses were reported, including one which stated that the eight dMMR patients treated in the experimental arm fared significantly better with regard to PFS than the five dMMR patients treated in the comparator arm.

Esophageal cancer

The Orient 15 trial is a quadruple-blind, placebo-controlled, phase III study, which evaluated the additive efficacy of sintilimab, if added to either a 3-weekly schedule of cisplatin and paclitaxel or a 3-weekly schedule of 5-fluorouracil and cisplatin in patients with advanced squamous cell cancer of the esophagus.[17] Primary outcome measure was OS, and PFS was a secondary outcome measure. QoL was not estimated. Subgroup analysis was not predefined. The treatments assigned to the comparator group could be regarded as standard regimens. The experimental arm encompassed 327 patients and the control arm 332 patients. In the ITT analysis, PFS was significantly superior in the experimental group (HR of progression 0.558, 95% CI 0.462-0.676). The number of patients censored in both study arms was not reported and imbalance in point censoring cannot be excluded. OS was significantly superior in the experimental group (HR of death 0.628, 95% CI 0.508-0.777). For both PFS and OS, *post hoc* subset analysis results were reported. The overall $\geq$3 TRAE rate was higher in the experimental group (59.9% versus 54.5%), as well as the rate of serious adverse events (20.8% versus 12.3%) and the number of toxic deaths (9 versus 6).

DISCUSSION

The ESMO 2021 conference provided data from a large number of randomized prospective trials. Given the fact that the study message should be conveyed in an abstract or by means of a short oral presentation, information should be cut short. As such, providing all essential information is a real challenge. But conference news spreads fast and conclusions tend to be spinned. Our paper focused on three basic questions: (i) Is the conference attendee provided with sufficient details regarding Patients and Methods to comprehend the results presented? (ii) Does the presentation give reason for considering bias? (iii) To which extent do the results presented translate to clinical benefit?

In order to elucidate these questions, we selected 11 study reports to retrieve the answers.

We have summarized the key points of each trial in Table 2. Three trials were carried out in the curative setting and eight trials were carried out in the palliative setting.

With regard to study methodology, there were several issues. Firstly, major alterations of study design were made in 2 out of 11 trials. The scientific soundness of these alterations and the reliability of the forthcoming results cannot be judged based on the information given. Secondly, 6 out of 11 trials were carried out without allocation concealment, which could be regarded as undesirable, given the risk of bias. Thirdly, in all six trials which reported PFS data, no information was presented to rule out imbalance in point censoring. Fourthly, in two out of seven trials, which reported OS data, desirable crossover could be anticipated, but crossover data were not presented. Finally, based on the overall survey, *post hoc* subset analysis appears to be regarded as 'business as usual', whereas it has a much higher rate of false positives than primary research with multiple tests being carried out on the same dataset. Stratified subgroup analysis should ideally be predefined in the trial protocol and *post hoc* subset analysis as a source of data mining should be discouraged.

Estimated clinical benefit of the experimental treatments evaluated in the reported trials should ideally be based on calculated ESMO-MCBS scores, but conference data do not enable such calculations. According to the principles of clinical benefit, we should strive for longer OS, less (grade $\geq$3)

Table 2. **Characteristics of the 11 practice-shaping ESMO 2021 conference reports discussed**

Study title	DestinyBreast03	KEYNOTE-522	KEYNOTE-826	Empower	KEYNOTE-716	PEACE-1	STAMPEDE	CheckMate 743	OREO	AtezoTRIBE	Orient 15
Cancer type	Breast	Breast	Uterine cervix	Uterine cervix	Melanoma	Prostate	Prostate	Mesothelioma	Ovary	Colorectal	Esophageal
Was randomization blinded?	No	Yes	Yes	No	Yes	No	No	No	Yes	No	Yes
Risk of performance bias?	Yes	No	No	Yes	No	Yes	Yes	Yes	No	Yes	No
Comparator arm to be regarded as standard?	Yes	Yes	Yes	Yes	Yes	Yes	Debatable	Yes	Yes	Yes	Yes
Were experimental and comparator arm redefined after enrolment and randomization?	No	No	No	No	No	Yes	Yes	No	No	No	No
Were ITT data made available?	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Curative or advanced care setting?	Advanced	Curative	Advanced	Advanced	Curative	Advanced	Curative	Advanced	Advanced	Advanced	Advanced
EFS/RFS/MFS/DMFS data available?	n.a.	Yes	n.a.	n.a.	Yes	n.a.	Yes	n.a.	n.a.	n.a.	n.a.
Experimental arm superior?	n.a.	Yes	n.a.	n.a.	Yes	n.a.	Yes	n.a.	n.a.	n.a.	n.a.
PFS data available?	Yes	n.a.	Yes	No	n.a.	Yes	n.a.	No	Yes	Yes	Yes
Experimental arm superior?	Yes	n.a.	Yes	n.a.	n.a.	Yes	n.a.	n.a.	Yes	Yes	Yes
Were PFS censoring data made available?	No	n.a.	No	n.a.	n.a.	No	n.a.	n.a.	No	No	No
Imbalance in informative censoring?	Unknown	n.a.	Unknown	n.a.	n.a.	Unknown	n.a.	n.a.	Unknown	Unknown	Unknown
OS data available?	Yes	No	Yes	Yes	No	Yes	Yes	Yes	No	No	Yes
Experimental arm superior?	Yes	n.a.	Yes	Yes	n.a.	Yes	Yes	Yes	n.a.	n.a.	Yes
Should desirable crossover be anticipated?	No	n.a.	No	No	n.a.	Yes	No	Yes	n.a.	n.a.	No
Were crossover data made available?	n.a.	n.a.	n.a.	n.a.	n.a.	No	n.a.	No	n.a.	n.a.	n.a.
Toxicity data available?	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	No	Yes
TRAE $\geq$3 rate experimental versus comparator arm	Higher	Higher	Higher	Lower	Higher	Higher	Higher	Equal	Higher	Unknown	Higher
QoL data available?	No	No	Yes	Yes	Yes	No	No	No	No	No	No
Experimental arm superior?	n.a.	n.a.	Yes	Yes	No	n.a.	n.a.	n.a.	n.a.	n.a.	n.a.

DMFS, distant metastasis-free survival; EFS, event-free survival; ITT, intention to treat; MFS, metastasis-free survival; n.a., not applicable; OS, overall survival; PFS, progression-free survival; QoL, quality of life; RFS, recurrence-free survival; TRAE, treatment-related adverse event.

TRAEs and sustained or improved QoL. And we should ascertain whether reported surrogate endpoints, such as PFS, have been proved to significantly relate to OS. It should be noted that the majority of the 11 study reports did not encompass all the clinically meaningful outcome measures specified at ClinicalTrials.gov. OS data were lacking in 4 out of 11 trials. In 1 out of 11 trials, information on TRAEs was lacking despite the intensive treatment regimen in both comparator and experimental groups. QoL data were presented for 3 out of 11 trials, but it is difficult to translate the figures shown in clinical benefit.

In conclusion, there is no doubt about the breakthrough potency of the 11 ESMO 2021 presentations discussed. But the information conveyed appeared incomplete for the purpose of safely estimating the true value of the experimental treatments. Therefore, conference presentations should be regarded as a means to inform the oncological community and to catalyze subsequent research. More guidance in regard to structure and items to be addressed could maybe lead to an even more fruitful spin-off during the conference. Finally, a full paper in a peer-reviewed journal should remain the gold standard for evidence-based oncology; both trial investigators and the oncological community should be enabled to draw solid conclusions and to define the related therapeutical consequences.

ACKNOWLEDGEMENTS

The authors acknowledge the work of Klizia Marinoni and Delanie Young, from the Scientific and Medical Division at ESMO, for the editorial assistance and George Pentheroudakis, Florian Lordick, Karin Jordan and Isabelle Ray-Coquard for the scientific feedback.

FUNDING

None declared.

DISCLOSURE

AT reports financial interests: Onco Care Limited, Member of Board of Directors, personal; practices in a private oncology setting; belongs to a company (three medical oncologists) who treat in private in Hamilton, New Zealand. Non-financial interests: Waikato Hospital, leadership role, current Director of Training for Medical Oncology at Waikato hospital. BP received honoraria from Novartis and BMS; was supported by ESMO with a Translational Research Fellowship aid supported by Roche. Any views, opinions, findings, conclusions or recommendations expressed in this material are those solely of the authors and do not necessarily reflect those of ESMO or Roche. BB reports invited speaker, expert honoraria: AstraZeneca, Sanofi, MSD, Swixxbiopharma, Janssen; Member of Steering Committee 'European Consensus: BRCA/Homologous Recombination Deficiency Testing in First-Line Ovarian Cancer' (Meridian HealthComms). Non-financial interest: member of NSGO and NSGO-CTU scientific committee. JB reports honorarium for advisory boards and educational symposium: AstraZeneca, Bristol-Myers Squibb, F. Hoffmann—La Roche Ltd, MSD, Novartis, Daiichi. AJC-T reports speaker's fees: BMS, Pierre-Fabre, Sanofi; meeting expenses: MSD, SANOFI, Novartis. MS reports speaker's fee: Roche, Novartis, Merck, BMS, Janssen; advisory role: Merck, BMS, Janssen; research grants: Janssen, BMS. HKvH has declared no conflicts of interest.

REFERENCES

1. Templeton AJ, Amir E, Tannock IF. Informative censoring—a neglected cause of bias in oncology trials. *Nat Rev Clin Oncol*. 2020;17:327-328.
2. Watt DC, Aitchison TC, Mackie RM, Sirel JM. Survival analysis, the importance of censored observations. *Melanoma Res*. 1996;6:379-385.
3. Viera AJ, Bangdiwala SI. Eliminating bias in randomized controlled trials: importance of allocation concealment and masking. *Fam Med*. 2007;39(2):132-137.
4. Haslam A, Prasad V. When is crossover desirable in cancer drug trials and when is it problematic? *Ann Oncol*. 2018;29(5):1079-1081.
5. Li G, Abbade LPF, Nwosu I, et al. A scoping review of comparisons between abstracts and full reports in biomedical research. *BMC Medl Res Methodol*. 2017;17:181.
6. Gyawali B, de Vries EGE, Dafni U, et al. Biases in study design, implementation and data analysis that distort the appraisal of clinical benefit and ESMO-Magnitude of Clinical Benefit Scale (ESMO-MCBS) scoring. *ESMO Open*. 2021;6:100117.
7. LBA1—Trastuzumab deruxtecan (T-DXd) vs trastuzumab emtansine (T-DM1) in patients (Pts) with HER2+ metastatic breast cancer (mBC): results of the randomized phase III DESTINY-Breast03 study. Available at https://oncologypro.esmo.org/meeting-resources/esmo-congress-2021/trastuzumab-deruxtecan-t-dxd-vs-trastuzumab-emtansine-t-dm1-in-patients-pts-with-her2-metastatic-breast-cancer-mbc-results-of-the-randomi. Accessed November 30, 2021.
8. VP7_2021—KEYNOTE-522: phase 3 study of pembrolizumab + chemotherapy vs placebo + chemotherapy as neoadjuvant treatment, followed by pembrolizumab vs placebo as adjuvant treatment for early triple-negative breast cancer (TNBC). Available at https://oncologypro.esmo.org/meeting-resources/esmo-congress-2021/vp7_2021-keynote-522-phase-3-study-of-pembrolizumab-chemotherapy-vs-placebo-chemotherapy-as-neoadjuvant-treatment-followed-by-pembrolizumab. Accessed November 30, 2021.
9. LBA2_PR—Pembrolizumab plus chemotherapy versus placebo plus chemotherapy for persistent, recurrent, or metastatic cervical cancer: randomized, double-blind, phase III KEYNOTE-826 study. Available at https://oncologypro.esmo.org/meeting-resources/esmo-congress-2021/pembrolizumab-plus-chemotherapy-versus-placebo-plus-chemotherapy-for-persistent-recurrent-or-metastatic-cervical-cancer-randomized-double-blind. Accessed November 30, 2021.
10. VP4_2021—EMPOWER-Cervical 1/GOG-3016/ENGOT-cx9: interim analysis of phase III trial of cemiplimab vs investigator's choice (IC) chemotherapy (chemo) in recurrent/metastatic (R/M) cervical carcinoma. Available at https://oncologypro.esmo.org/meeting-resources/esmo-congress-2021/vp4_2021-empower-cervical-1-gog-3016-engot-cx9-interim-analysis-of-phase-iii-trial-of-cemiplimab-vs-investigator-s-choice-ic-chemotherapy-che. Accessed November 30, 2021.
11. LBA3_PR—Pembrolizumab versus placebo after complete resection of high-risk stage II melanoma: efficacy and safety results from the KEYNOTE-716 double-blind phase III trial. Available at https://oncologypro.esmo.org/meeting-resources/esmo-congress-2021/pembrolizumab-versus-placebo-after-complete-resection-of-high-risk-stage-ii-melanoma-efficacy-and-safety-results-from-the-keynote-716-double-blind. Accessed November 30, 2021.
12. LBA5_PR—A phase III trial with a 2x2 factorial design in men with de novo metastatic castration-sensitive prostate cancer: overall survival with abiraterone acetate plus prednisone in PEACE-1. Available at https://oncologypro.esmo.org/meeting-resources/esmo-congress-2021/a-phase-iii-trial-with-a-2x2-factorial-design-in-men-with-de-novo-metastatic-castration-sensitive-prostate-cancer-overall-survival-with-abirateron. Accessed November 30, 2021.

13. LBA4_PR—Abiraterone acetate plus prednisolone (AAP) with or without enzalutamide (ENZ) added to androgen deprivation therapy (ADT) compared to ADT alone for men with high-risk non-metastatic (M0) prostate cancer (PCa): combined analysis from two comparisons in the STAMPEDE platform protocol. Available at https://oncologypro.esmo.org/meeting-resources/esmo-congress-2021/abiraterone-acetate-plus-prednisolone-aap-with-or-without-enzalutamide-enz-added-to-androgen-deprivation-therapy-adt-compared-to-adt-alone-fo. Accessed November 30, 2021.
14. LBA65—First-line nivolumab (NIVO) plus ipilimumab (IPI) vs chemotherapy (chemo) in patients (pts) with unresectable malignant pleural mesothelioma (MPM): 3-year update from CheckMate 743. Available at https://oncologypro.esmo.org/meeting-resources/esmo-congress-2021/first-line-nivolumab-nivo-plus-ipilimumab-ipi-vs-chemotherapy-chemo-in-patients-pts-with-unresectable-malignant-pleural-mesothelioma-mpm. Accessed November 30, 2021.
15. LBA33—Maintenance olaparib rechallenge in patients (pts) with ovarian carcinoma (OC) previously treated with a PARP inhibitor (PARPi): phase IIIb OReO/ENGOT Ov-38 trial. Available at https://oncologypro.esmo.org/meeting-resources/esmo-congress-2021/maintenance-Olaparib-rechallenge-in-patients-pts-with-ovarian-carcinoma-oc-previously-treated-with-a-parp-inhibitor-parpi-phase-iiib-oreo-en. Accessed November 30, 2021.
16. LBA20—FOLFOXIRI plus bevacizumab (bev) plus atezolizumab (atezo) versus FOLFOXIRI plus bev as first-line treatment of unresectable metastatic colorectal cancer (mCRC) patients: results of the phase II randomized AtezoTRIBE study by GONO. Available at https://oncologypro.esmo.org/meeting-resources/esmo-congress-2021/folfoxiri-plus-bevacizumab-bev-plus-atezolizumab-atezo-versus-folfoxiri-plus-bev-as-first-line-treatment-of-unresectable-metastatic-colorectal. Accessed November 30, 2021.
17. ESMO Congress 2021. Available at https://oncologypro.esmo.org/meeting-resources/esmo-congress-2021?event_resources_filter_form%5Bsearch%5D=orient%2015. Accessed November 30, 2021.